Catalan Numbers

Murray Bremner

University of Saskatchewan, Canada

University of Cape Town, Monday 30 November 2015

PART 1

PART 1

Catalan numbers arise from a simple enumeration problem:

PART 1

Catalan numbers arise from a simple enumeration problem:

> *In how many different ways can we compute the product of n factors which don't commute?*

PART 1

Catalan numbers arise from a simple enumeration problem:

> *In how many different ways can we compute the product of n factors which don't commute?*

With three factors abc there are two multiplications to perform and we can do them either in one order $(ab)c$ or the other $a(bc)$.

PART 1

Catalan numbers arise from a simple enumeration problem:

> *In how many different ways can we compute the product of n factors which don't commute?*

With three factors abc there are two multiplications to perform and we can do them either in one order $(ab)c$ or the other $a(bc)$.

We can represent these possibilities as rooted planar binary trees:

PART 1

Catalan numbers arise from a simple enumeration problem:

> *In how many different ways can we compute the product of n factors which don't commute?*

With three factors abc there are two multiplications to perform and we can do them either in one order $(ab)c$ or the other $a(bc)$.

We can represent these possibilities as rooted planar binary trees:

PART 1

Catalan numbers arise from a simple enumeration problem:

> *In how many different ways can we compute the product of n factors which don't commute?*

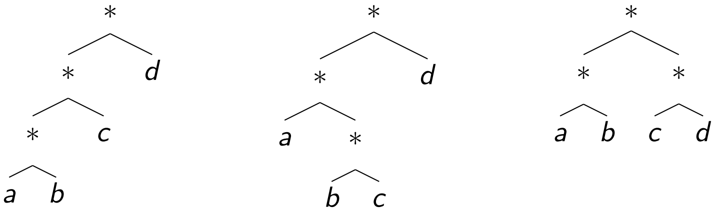
With three factors abc there are two multiplications to perform and we can do them either in one order $(ab)c$ or the other $a(bc)$.

We can represent these possibilities as rooted planar binary trees:

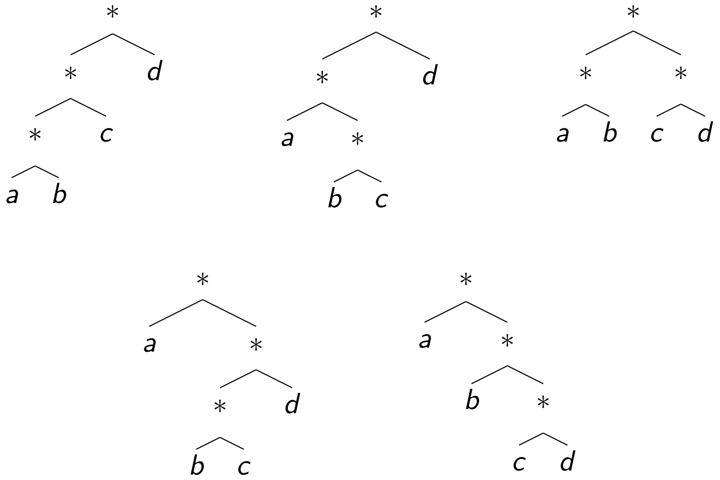
The asterisks represent multiplications, and the two subtrees below each asterisk represent the two factors being multiplied.

With four factors, there are five ways to compute the product:

With four factors, there are five ways to compute the product:



With four factors, there are five ways to compute the product:



Using parentheses these sequences of multiplications look like this:

Using parentheses these sequences of multiplications look like this:

$$((ab)c)d, \qquad (a(bc))d, \qquad (ab)(cd),$$

Using parentheses these sequences of multiplications look like this:

$$((ab)c)d, \qquad (a(bc))d, \qquad (ab)(cd),$$

$$a((bc)d), \qquad a(b(cd)).$$

Using parentheses these sequences of multiplications look like this:

$$((ab)c)d, \qquad (a(bc))d, \qquad (ab)(cd),$$

$$a((bc)d), \qquad a(b(cd)).$$

For 5, 6, 7, 8, 9, 10, . . . factors the number of possibilities is

$$14, \quad 42, \quad 132, \quad 429, \quad 1430, \quad 4862, \quad \ldots$$

Using parentheses these sequences of multiplications look like this:

$$((ab)c)d, \qquad (a(bc))d, \qquad (ab)(cd),$$

$$a((bc)d), \qquad a(b(cd)).$$

For 5, 6, 7, 8, 9, 10, ... factors the number of possibilities is

$$14, \quad 42, \quad 132, \quad 429, \quad 1430, \quad 4862, \quad \ldots$$

Is there a simple compact formula for the terms of this sequence?

Using parentheses these sequences of multiplications look like this:

$$((ab)c)d, \qquad (a(bc))d, \qquad (ab)(cd),$$

$$a((bc)d), \qquad a(b(cd)).$$

For 5, 6, 7, 8, 9, 10, ... factors the number of possibilities is

$$14, \quad 42, \quad 132, \quad 429, \quad 1430, \quad 4862, \quad \ldots$$

Is there a simple compact formula for the terms of this sequence?

If there is, how do we go about finding it?

Using parentheses these sequences of multiplications look like this:

$$((ab)c)d, \qquad (a(bc))d, \qquad (ab)(cd),$$

$$a((bc)d), \qquad a(b(cd)).$$

For 5, 6, 7, 8, 9, 10, . . . factors the number of possibilities is

$$14, \quad 42, \quad 132, \quad 429, \quad 1430, \quad 4862, \quad \ldots$$

Is there a simple compact formula for the terms of this sequence?

If there is, how do we go about finding it?

First, let's consider some other applications of Catalan numbers.

Asymptotics:

$$C_n \sim \frac{4^n}{n^{3/2}\sqrt{\pi}} \quad \text{as} \quad n \to \infty$$

Asymptotics:

$$C_n \sim \frac{4^n}{n^{3/2}\sqrt{\pi}} \quad \text{as} \quad n \to \infty$$

This follows from the formula we will prove and Stirling's formula:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \quad \text{as} \quad n \to \infty$$

Asymptotics:

$$C_n \sim \frac{4^n}{n^{3/2}\sqrt{\pi}} \quad \text{as} \quad n \to \infty$$

This follows from the formula we will prove and Stirling's formula:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \quad \text{as} \quad n \to \infty$$

The Catalan numbers appear as the solution to a very large number of different combinatorial problems.

Asymptotics:

$$C_n \sim \frac{4^n}{n^{3/2}\sqrt{\pi}} \quad \text{as} \quad n \to \infty$$

This follows from the formula we will prove and Stirling's formula:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \quad \text{as} \quad n \to \infty$$

The Catalan numbers appear as the solution to a very large number of different combinatorial problems.

Richard Stanley's *Enumerative Combinatorics: Volume 2* (Cambridge U. Press, 1999) has an exercise which gives 66 different interpretations of the Catalan numbers.

Asymptotics:

$$C_n \sim \frac{4^n}{n^{3/2}\sqrt{\pi}} \quad \text{as} \quad n \to \infty$$

This follows from the formula we will prove and Stirling's formula:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \quad \text{as} \quad n \to \infty$$

The Catalan numbers appear as the solution to a very large number of different combinatorial problems.

Richard Stanley's *Enumerative Combinatorics: Volume 2* (Cambridge U. Press, 1999) has an exercise which gives 66 different interpretations of the Catalan numbers.

The Catalan numbers appear as sequence A000108 in the OEIS (*On-Line Encyclopedia of Integer Sequences*, `oeis.org`):

Asymptotics:

$$C_n \sim \frac{4^n}{n^{3/2}\sqrt{\pi}} \quad \text{as} \quad n \to \infty$$

This follows from the formula we will prove and Stirling's formula:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \quad \text{as} \quad n \to \infty$$

The Catalan numbers appear as the solution to a very large number of different combinatorial problems.

Richard Stanley's *Enumerative Combinatorics: Volume 2* (Cambridge U. Press, 1999) has an exercise which gives 66 different interpretations of the Catalan numbers.

The Catalan numbers appear as sequence A000108 in the OEIS (*On-Line Encyclopedia of Integer Sequences*, `oeis.org`):
"This is probably the longest entry in the OEIS, and rightly so."

Dyck words: C_n is the number of Dyck words of length $2n$, where a Dyck word is a string of n a's and n b's such that no initial segment of the string has more b's than a's. For example:

$$n = 1: \quad ab$$
$$n = 2: \quad aabb, \quad abab$$
$$n = 3: \quad aaabbb, \quad aababb, \quad aabbab, \quad abaabb, \quad ababab$$

Dyck words: C_n is the number of Dyck words of length $2n$, where a Dyck word is a string of n a's and n b's such that no initial segment of the string has more b's than a's. For example:

$$n = 1: \quad ab$$
$$n = 2: \quad aabb, \quad abab$$
$$n = 3: \quad aaabbb, \quad aababb, \quad aabbab, \quad abaabb, \quad ababab$$

This is equivalent to another parentheses problem: if we replace a by (, and b by), we obtain these placements of parentheses:

$$n = 1: \quad ()$$
$$n = 2: \quad (()), \quad ()()$$
$$n = 3: \quad ((())), \quad (()()), \quad (())(), \quad ()(()), \quad ()()()$$

Dyck words: C_n is the number of Dyck words of length $2n$, where a Dyck word is a string of n a's and n b's such that no initial segment of the string has more b's than a's. For example:

$$n=1: \quad ab$$
$$n=2: \quad aabb, \quad abab$$
$$n=3: \quad aaabbb, \quad aababb, \quad aabbab, \quad abaabb, \quad ababab$$

This is equivalent to another parentheses problem: if we replace a by (, and b by), we obtain these placements of parentheses:

$$n=1: \quad ()$$
$$n=2: \quad (()), \quad ()()$$
$$n=3: \quad ((())), \quad (()()), \quad (())(), \quad ()(()), \quad ()()()$$

Exercise: Find a bijection between these parenthesizations and those in our original formulation of the enumeration problem.

Lattice paths: Given a positive integer n, imagine that you must follow a path in the (x, y) plane satisfying these conditions:

- You must start at the point $(0, 0)$.
- You must take steps of length 1, east $(1, 0)$ or north $(0, 1)$.
- You must not pass above the diagonal line $y = x$.
- You must end at the point (n, n).

Lattice paths: Given a positive integer n, imagine that you must follow a path in the (x, y) plane satisfying these conditions:

- You must start at the point $(0, 0)$.
- You must take steps of length 1, east $(1, 0)$ or north $(0, 1)$.
- You must not pass above the diagonal line $y = x$.
- You must end at the point (n, n).

How many different paths are there which satisfy these conditions?

Lattice paths: Given a positive integer n, imagine that you must follow a path in the (x, y) plane satisfying these conditions:

- You must start at the point $(0, 0)$.
- You must take steps of length 1, east $(1, 0)$ or north $(0, 1)$.
- You must not pass above the diagonal line $y = x$.
- You must end at the point (n, n).

How many different paths are there which satisfy these conditions?

The answer is the Catalan number C_n.

Lattice paths: Given a positive integer n, imagine that you must follow a path in the (x, y) plane satisfying these conditions:

- You must start at the point $(0, 0)$.
- You must take steps of length 1, east $(1, 0)$ or north $(0, 1)$.
- You must not pass above the diagonal line $y = x$.
- You must end at the point (n, n).

How many different paths are there which satisfy these conditions?

The answer is the Catalan number C_n.

In fact, there is a simple bijection between Dyck words and these lattice paths: interpret a as "go east", and b as "go north".

Lattice paths: Given a positive integer n, imagine that you must follow a path in the (x, y) plane satisfying these conditions:

- You must start at the point $(0, 0)$.
- You must take steps of length 1, east $(1, 0)$ or north $(0, 1)$.
- You must not pass above the diagonal line $y = x$.
- You must end at the point (n, n).

How many different paths are there which satisfy these conditions?

The answer is the Catalan number C_n.

In fact, there is a simple bijection between Dyck words and these lattice paths: interpret a as "go east", and b as "go north".

The condition that a Dyck word has the same number of a's and b's guarantees that the lattice path ends at (n, n).

Lattice paths: Given a positive integer n, imagine that you must follow a path in the (x, y) plane satisfying these conditions:

- You must start at the point $(0, 0)$.
- You must take steps of length 1, east $(1, 0)$ or north $(0, 1)$.
- You must not pass above the diagonal line $y = x$.
- You must end at the point (n, n).

How many different paths are there which satisfy these conditions?

The answer is the Catalan number C_n.

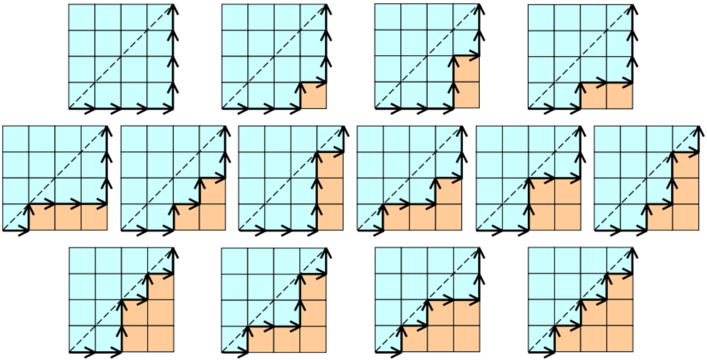
In fact, there is a simple bijection between Dyck words and these lattice paths: interpret a as "go east", and b as "go north".

The condition that a Dyck word has the same number of a's and b's guarantees that the lattice path ends at (n, n).

The condition that no initial segment of a Dyck word has more b's than a's guarantees that the lattice path stays below the diagonal.

The 14 lattice paths in a 4×4 grid from the Wikipedia article on "Catalan number":

The 14 lattice paths in a 4×4 grid from the Wikipedia article on "Catalan number":



Permutations with excluded subsequences: We write S_n for the symmetric group of all permutations of the set $X = \{1, 2, \ldots, n\}$.

Permutations with excluded subsequences: We write S_n for the symmetric group of all permutations of the set $X = \{1, 2, \ldots, n\}$.

We think of an element $p \in S_n$ as a sequence $p_1, p_2, \ldots, p_n$ representing a bijective function $p\colon X \to X$.

Permutations with excluded subsequences: We write S_n for the symmetric group of all permutations of the set $X = \{1, 2, \dots, n\}$.

We think of an element $p \in S_n$ as a sequence $p_1, p_2, \dots, p_n$ representing a bijective function $p \colon X \to X$.

By a *subsequence* of p of length k we mean a subset of k elements of p which are in order but *not necessarily consecutive*:

$$p_{i_1}, \quad p_{i_2}, \quad , \dots, \quad p_{i_k} \qquad (i_1 < i_2 < \cdots < i_k).$$

Permutations with excluded subsequences: We write S_n for the symmetric group of all permutations of the set $X = \{1, 2, \ldots, n\}$.

We think of an element $p \in S_n$ as a sequence $p_1, p_2, \ldots, p_n$ representing a bijective function $p \colon X \to X$.

By a *subsequence* of p of length k we mean a subset of k elements of p which are in order but *not necessarily consecutive*:

$$p_{i_1}, \quad p_{i_2}, \quad , \ldots, \quad p_{i_k} \qquad (i_1 < i_2 < \cdots < i_k).$$

How many permutations in S_n are there with no increasing subsequence of length 3?

Permutations with excluded subsequences: We write S_n for the symmetric group of all permutations of the set $X = \{1, 2, \ldots, n\}$.

We think of an element $p \in S_n$ as a sequence $p_1, p_2, \ldots, p_n$ representing a bijective function $p \colon X \to X$.

By a *subsequence* of p of length k we mean a subset of k elements of p which are in order but *not necessarily consecutive*:

$$p_{i_1}, \quad p_{i_2}, \quad , \ldots, \quad p_{i_k} \qquad (i_1 < i_2 < \cdots < i_k).$$

How many permutations in S_n are there with no increasing subsequence of length 3?

Trivially, for n = 1 there is 1, and for n = 2 there are 2. (There are no subsequences of length 3 to exclude.)

For $n = 3$, we have to exclude the permutation 123, leaving 5:

$$132, \quad 213, \quad 231, \quad 312, \quad 321.$$

For $n = 3$, we have to exclude the permutation 123, leaving 5:

$$132, \quad 213, \quad 231, \quad 312, \quad 321.$$

For $n = 4$, there are 24 permutations, and the increasing subsequences could occur in positions 123, 124, 134 or 234.

For $n = 3$, we have to exclude the permutation 123, leaving 5:

$$132, \quad 213, \quad 231, \quad 312, \quad 321.$$

For $n = 4$, there are 24 permutations, and the increasing subsequences could occur in positions 123, 124, 134 or 234. Here are all 24 permutations with the excluded ones underlined:

$\underline{1234}$	$\underline{1243}$	$\underline{1324}$	$\underline{1342}$	$\underline{1423}$	1432
$\underline{2134}$	2143	$\underline{2314}$	$\underline{2341}$	2413	2431
$\underline{3124}$	3142	3214	3241	3412	3421
$\underline{4123}$	4132	4213	4231	4312	4321

There are 10 excluded permutations, so 14 remain.

For $n = 1, 2, 3, 4$ we have 1, 2, 5, 14, . . . : the Catalan numbers!

For $n = 1, 2, 3, 4$ we have 1, 2, 5, 14, . . . : the Catalan numbers!

Theorem

The number of permutations in S_n which have no increasing subsequence of length 3 is the Catalan number C_n.

For $n = 1, 2, 3, 4$ we have 1, 2, 5, 14, . . . : the Catalan numbers!

Theorem

The number of permutations in S_n which have no increasing subsequence of length 3 is the Catalan number C_n.

For permutations with no increasing subsequences of length 4:

1, 2, 6, 23, 103, 513, 2761, 15767, 94359, 586590, 3763290, . . .

For $n = 1, 2, 3, 4$ we have 1, 2, 5, 14, . . . : the Catalan numbers!

Theorem

The number of permutations in S_n which have no increasing subsequence of length 3 is the Catalan number C_n.

For permutations with no increasing subsequences of length 4:

1, 2, 6, 23, 103, 513, 2761, 15767, 94359, 586590, 3763290, . . .

For permutations with no increasing subsequences of length 5:

1, 2, 6, 24, 119, 694, 4582, 33324, 261808, 2190688, 19318688, . . .

For $n = 1, 2, 3, 4$ we have 1, 2, 5, 14, . . . : the Catalan numbers!

Theorem

The number of permutations in S_n which have no increasing subsequence of length 3 is the Catalan number C_n.

For permutations with no increasing subsequences of length 4:

1, 2, 6, 23, 103, 513, 2761, 15767, 94359, 586590, 3763290, . . .

For permutations with no increasing subsequences of length 5:

1, 2, 6, 24, 119, 694, 4582, 33324, 261808, 2190688, 19318688, . . .

For permutations with no increasing subsequences of length 6:

1, 2, 6, 24, 120, 719, 5003, 39429, 344837, 3291590, 33835114, . . .

The last three sequences are from the OEIS.

Polygon triangulation: Let n be a positive integer, and consider a regular polygon with $n + 2$ sides.
So $n = 1$ gives a triangle, $n = 2$ gives a square, $n = 3$ gives a pentagon, $n = 4$ gives hexagon, etc.

Polygon triangulation: Let n be a positive integer, and consider a regular polygon with $n + 2$ sides.
So $n = 1$ gives a triangle, $n = 2$ gives a square, $n = 3$ gives a pentagon, $n = 4$ gives hexagon, etc.
Join the vertices of the polygon by non-intersecting straight lines so that the polygon is divided into triangles, and count how many distinct ways there are of doing this.

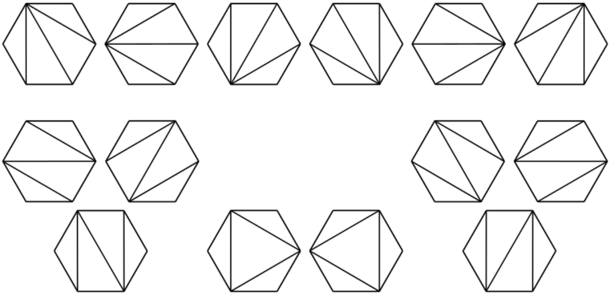
Polygon triangulation: Let n be a positive integer, and consider a regular polygon with $n+2$ sides.
So $n=1$ gives a triangle, $n=2$ gives a square, $n=3$ gives a pentagon, $n=4$ gives hexagon, etc.
Join the vertices of the polygon by non-intersecting straight lines so that the polygon is divided into triangles, and count how many distinct ways there are of doing this.
$n=1$: Polygon is a triangle, nothing to do, only one possibility.

Polygon triangulation: Let n be a positive integer, and consider a regular polygon with $n + 2$ sides.
So $n = 1$ gives a triangle, $n = 2$ gives a square, $n = 3$ gives a pentagon, $n = 4$ gives hexagon, etc.
Join the vertices of the polygon by non-intersecting straight lines so that the polygon is divided into triangles, and count how many distinct ways there are of doing this.
$n = 1$: Polygon is a triangle, nothing to do, only one possibility.
$n = 2$: Square can be divided by one diagonal or the other, giving two possibilities.

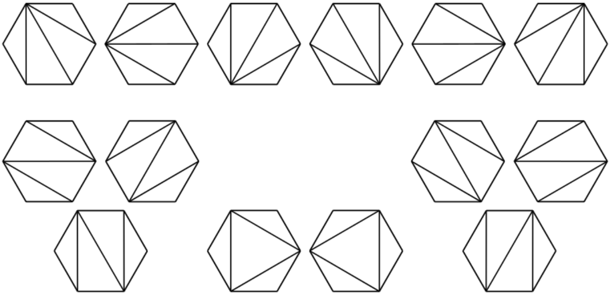
Polygon triangulation: Let n be a positive integer, and consider a regular polygon with $n+2$ sides.
So $n=1$ gives a triangle, $n=2$ gives a square, $n=3$ gives a pentagon, $n=4$ gives hexagon, etc.
Join the vertices of the polygon by non-intersecting straight lines so that the polygon is divided into triangles, and count how many distinct ways there are of doing this.
$n=1$: Polygon is a triangle, nothing to do, only one possibility.
$n=2$: Square can be divided by one diagonal or the other, giving two possibilities.
$n=3$: Clockwise from the top, label the vertices of the pentagon 1 (top), 2 (right), 3, 4 (base), 5 (left). From any vertex, draw two lines from that vertex to the endpoints of the opposite edge. Every decomposition has this form, giving five possibilities.

$n = 4$: The 14 triangulations of a hexagon from the Wikipedia article on "Catalan number":

$n = 4$: The 14 triangulations of a hexagon from the Wikipedia article on “Catalan number”:



$n = 4$: The 14 triangulations of a hexagon from the Wikipedia article on "Catalan number":



Theorem

The number of ways in which the regular polygon with n sides can be triangulated is the Catalan number C_{n-2}.

Young tableaux: Let n be a positive integer, and consider a *partition* of n:

$$n = n_1 + n_2 + \cdots + n_k, \qquad n_1 \geq n_2 \geq \cdots \geq n_k \geq 1,$$

A *Young diagram* consists of k rows of (left-justified) empty squares, with n_i squares in row i.

Young tableaux: Let n be a positive integer, and consider a *partition* of n:

$$n = n_1 + n_2 + \cdots + n_k, \qquad n_1 \geq n_2 \geq \cdots \geq n_k \geq 1,$$

A *Young diagram* consists of k rows of (left-justified) empty squares, with n_i squares in row i.
A *Young tableau* consists of some placement of the numbers $1, 2, \ldots, n$ into the n boxes.

Young tableaux: Let n be a positive integer, and consider a *partition* of n:

$$n = n_1 + n_2 + \cdots + n_k, \qquad n_1 \geq n_2 \geq \cdots \geq n_k \geq 1,$$

A *Young diagram* consists of k rows of (left-justified) empty squares, with n_i squares in row i.
A *Young tableau* consists of some placement of the numbers $1, 2, \ldots, n$ into the n boxes.
A Young tableau is *standard* if the numbers increase from left to right along the rows, and from top to bottom along the columns.

Young tableaux: Let n be a positive integer, and consider a *partition* of n:

$$n = n_1 + n_2 + \cdots + n_k, \qquad n_1 \geq n_2 \geq \cdots \geq n_k \geq 1,$$

A *Young diagram* consists of k rows of (left-justified) empty squares, with n_i squares in row i.
A *Young tableau* consists of some placement of the numbers $1, 2, \ldots, n$ into the n boxes.
A Young tableau is *standard* if the numbers increase from left to right along the rows, and from top to bottom along the columns.
The irreducible representations of the symmetric group S_n are labelled by the partitions of n, and the dimension of each representation is the number of standard Young tableaux for the corresponding Young diagram.

Young tableaux: Let n be a positive integer, and consider a *partition* of n:

$$n = n_1 + n_2 + \cdots + n_k, \qquad n_1 \geq n_2 \geq \cdots \geq n_k \geq 1,$$

A *Young diagram* consists of k rows of (left-justified) empty squares, with n_i squares in row i.
A *Young tableau* consists of some placement of the numbers $1, 2, \ldots, n$ into the n boxes.
A Young tableau is *standard* if the numbers increase from left to right along the rows, and from top to bottom along the columns.
The irreducible representations of the symmetric group S_n are labelled by the partitions of n, and the dimension of each representation is the number of standard Young tableaux for the corresponding Young diagram.
For $n = 2m$, partition $n = m + m$ (Young diagram is $2 \times m$ array), the number of standard tableaux is the Catalan number C_n.

It is the number of ways 1, 2, ..., 2n can be arranged in a 2-by-n rectangle so that each row and each column is increasing:

It is the number of ways 1, 2, ..., 2n can be arranged in a 2-by-n rectangle so that each row and each column is increasing:

$$n = 1: \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad n = 2: \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

$$n = 3: \begin{bmatrix} 1\,2\,3 \\ 4\,5\,6 \end{bmatrix} \begin{bmatrix} 1\,2\,4 \\ 3\,5\,6 \end{bmatrix} \begin{bmatrix} 1\,2\,5 \\ 3\,4\,6 \end{bmatrix} \begin{bmatrix} 1\,3\,4 \\ 2\,5\,6 \end{bmatrix} \begin{bmatrix} 1\,3\,5 \\ 2\,4\,6 \end{bmatrix}$$

$$n = 4: \begin{bmatrix} 1\,2\,3\,4 \\ 5\,6\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,3\,5 \\ 4\,6\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,3\,6 \\ 4\,5\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,3\,7 \\ 4\,5\,6\,8 \end{bmatrix}$$

$$\begin{bmatrix} 1\,2\,4\,5 \\ 3\,6\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,4\,6 \\ 3\,5\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,4\,7 \\ 3\,5\,6\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,5\,6 \\ 3\,4\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,5\,7 \\ 3\,4\,6\,8 \end{bmatrix}$$

$$\begin{bmatrix} 1\,3\,4\,5 \\ 2\,6\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,3\,4\,6 \\ 2\,5\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,3\,4\,7 \\ 2\,5\,6\,8 \end{bmatrix} \begin{bmatrix} 1\,3\,5\,6 \\ 2\,4\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,3\,5\,7 \\ 2\,4\,6\,8 \end{bmatrix}$$

It is the number of ways 1, 2, ..., 2n can be arranged in a 2-by-n rectangle so that each row and each column is increasing:

$$n = 1 : \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad n = 2 : \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

$$n = 3 : \begin{bmatrix} 1\,2\,3 \\ 4\,5\,6 \end{bmatrix} \begin{bmatrix} 1\,2\,4 \\ 3\,5\,6 \end{bmatrix} \begin{bmatrix} 1\,2\,5 \\ 3\,4\,6 \end{bmatrix} \begin{bmatrix} 1\,3\,4 \\ 2\,5\,6 \end{bmatrix} \begin{bmatrix} 1\,3\,5 \\ 2\,4\,6 \end{bmatrix}$$

$$n = 4 : \begin{bmatrix} 1\,2\,3\,4 \\ 5\,6\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,3\,5 \\ 4\,6\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,3\,6 \\ 4\,5\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,3\,7 \\ 4\,5\,6\,8 \end{bmatrix}$$

$$\begin{bmatrix} 1\,2\,4\,5 \\ 3\,6\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,4\,6 \\ 3\,5\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,4\,7 \\ 3\,5\,6\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,5\,6 \\ 3\,4\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,2\,5\,7 \\ 3\,4\,6\,8 \end{bmatrix}$$

$$\begin{bmatrix} 1\,3\,4\,5 \\ 2\,6\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,3\,4\,6 \\ 2\,5\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,3\,4\,7 \\ 2\,5\,6\,8 \end{bmatrix} \begin{bmatrix} 1\,3\,5\,6 \\ 2\,4\,7\,8 \end{bmatrix} \begin{bmatrix} 1\,3\,5\,7 \\ 2\,4\,6\,8 \end{bmatrix}$$

Again we see the sequence 1, 2, 5, 14, . . .

Now, let's return to finding a formula for the Catalan numbers.

Observation 1: Every product z of degree n has the form $z = x \cdot y$ where x, y have degrees i, j with $i + j = n$ and $1 \leq i, j \leq n - 1$.

Now, let's return to finding a formula for the Catalan numbers.

Observation 1: Every product z of degree n has the form $z = x \cdot y$ where x, y have degrees i, j with $i + j = n$ and $1 \leq i, j \leq n - 1$.

Using trees, x and y are the left and right subtrees of the root.

Now, let's return to finding a formula for the Catalan numbers.

Observation 1: Every product z of degree n has the form $z = x \cdot y$ where x, y have degrees i, j with $i + j = n$ and $1 \leq i, j \leq n - 1$.

Using trees, x and y are the left and right subtrees of the root.

Conclusion: We can give an algorithm to construct inductively all products of degree n once and only once:

- Set $P[1] \leftarrow [\,X\,]$ (the list containing the single element X)
- For n from 2 to MAXDEG do
  - Set $P[n] \leftarrow []$ (empty list)
  - For j from 1 to $n - 1$ do:
    - For x in $P[n - j]$ do for y in $P[j]$ do:
      Adjoin $[x, y]$ to the list $P[n]$.

Now, let's return to finding a formula for the Catalan numbers.

Observation 1: Every product z of degree n has the form $z = x \cdot y$ where x, y have degrees i, j with $i + j = n$ and $1 \leq i, j \leq n - 1$.

Using trees, x and y are the left and right subtrees of the root.

Conclusion: We can give an algorithm to construct inductively all products of degree n once and only once:

- Set $P[1] \leftarrow [\,X\,]$ (the list containing the single element X)
- For n from 2 to MAXDEG do
  - Set $P[n] \leftarrow [\,]$ (empty list)
  - For j from 1 to $n - 1$ do:
    - For x in $P[n-j]$ do for y in $P[j]$ do:
      Adjoin $[x, y]$ to the list $P[n]$.

The symbol X is a placeholder for all the factors; in each element of $P[n]$ we need to replace the n X's by the variables $a_1, \ldots, a_n$.

The number we are looking for is $P_n = |\,P[n]\,|$, the number of ways of placing balanced parantheses into a sequence of n factors.

The number we are looking for is $P_n = |\,P[n]\,|$, the number of ways of placing balanced parantheses into a sequence of n factors.

Observation 2: The previous observation can be reformulated as:

$$P_n = P_{n-1}P_1 + P_{n-2}P_2 + \cdots + P_2P_{n-2} + P_1P_{n-1},$$

$$P_n = \sum_{i+j=n} P_iP_j, \qquad P_n = \sum_{j=1}^{n-1} P_{n-j}P_j.$$

The number we are looking for is $P_n = |\,P[n]\,|$, the number of ways of placing balanced parantheses into a sequence of n factors.

Observation 2: The previous observation can be reformulated as:

$$P_n = P_{n-1}P_1 + P_{n-2}P_2 + \cdots + P_2P_{n-2} + P_1P_{n-1},$$

$$P_n = \sum_{i+j=n} P_iP_j, \qquad P_n = \sum_{j=1}^{n-1} P_{n-j}P_j.$$

Observation 3: It will probably be useful to consider the so-called *ordinary generating function*

$$P(t) = \sum_{n=1}^{\infty} P_n t^n,$$

which is a formal power series in the indeterminate t.

The terms appearing in Observation 2 also appear in $P(t)^2$:

$$\begin{aligned}
P(t)^2 &= \left(\sum_{i=1}^{\infty} P_i t^i\right)\left(\sum_{j=1}^{\infty} P_j t^j\right) \\
&= \sum_{n=2}^{\infty}\left(\sum_{i+j=n} P_i P_j\right) t^{i+j} \qquad \text{(outer sum starts at } n=2\text{)} \\
&= \sum_{n=2}^{\infty} P_n t^n \qquad \text{(Observation 2)} \\
&= P(t) - t \qquad \text{(since } P_1 = 1\text{)}
\end{aligned}$$

The terms appearing in Observation 2 also appear in $P(t)^2$:

$$\begin{aligned} P(t)^2 &= \left(\sum_{i=1}^{\infty} P_i t^i\right)\left(\sum_{j=1}^{\infty} P_j t^j\right) \\ &= \sum_{n=2}^{\infty}\left(\sum_{i+j=n} P_i P_j\right) t^{i+j} \qquad \text{(outer sum starts at } n=2) \\ &= \sum_{n=2}^{\infty} P_n t^n \qquad \text{(Observation 2)} \\ &= P(t) - t \qquad (\text{since } P_1 = 1) \end{aligned}$$

Conclusion: We have $P(t)^2 - P(t) + t = 0$. This is a quadratic equation for the function $P(t)$ whose coefficients are polynomials in t, so we can use the quadratic formula to solve for $P(t)$.

$$P(t)^2 - P(t) + t = 0 \quad \implies \quad a = 1, \quad b = -1, \quad c = t$$

$$P(t)^2 - P(t) + t = 0 \quad \Longrightarrow \quad a = 1, \quad b = -1, \quad c = t$$

$$P(t) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{1 - 4t}}{2}$$

$$P(t)^2 - P(t) + t = 0 \quad \Longrightarrow \quad a = 1, \quad b = -1, \quad c = t$$

$$P(t) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{1 - 4t}}{2}$$

There is no constant term in the formal power series for $P(t)$, hence $P(0) = 0$, hence we must take the minus sign:

$$P(t)^2 - P(t) + t = 0 \quad \Longrightarrow \quad a = 1, \quad b = -1, \quad c = t$$

$$P(t) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{1 - 4t}}{2}$$

There is no constant term in the formal power series for $P(t)$, hence $P(0) = 0$, hence we must take the minus sign:

$$P(t) = \frac{1 - \sqrt{1 - 4t}}{2}$$

$$P(t)^2 - P(t) + t = 0 \quad \Longrightarrow \quad a = 1, \quad b = -1, \quad c = t$$

$$P(t) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{1-4t}}{2}$$

There is no constant term in the formal power series for $P(t)$, hence $P(0) = 0$, hence we must take the minus sign:

$$P(t) = \frac{1 - \sqrt{1-4t}}{2}$$

Calculating the first few terms on a computer algebra system gives

$$P(t)^2 - P(t) + t = 0 \quad \Longrightarrow \quad a = 1, \quad b = -1, \quad c = t$$

$$P(t) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{1 - 4t}}{2}$$

There is no constant term in the formal power series for $P(t)$, hence $P(0) = 0$, hence we must take the minus sign:

$$P(t) = \frac{1 - \sqrt{1 - 4t}}{2}$$

Calculating the first few terms on a computer algebra system gives

$$t + t^2 + 2t^3 + 5t^4 + 14t^5 + 42t^6 + 132t^7 + 429t^8 + 1430t^9 + \cdots ,$$

$$P(t)^2 - P(t) + t = 0 \quad \Longrightarrow \quad a = 1, \quad b = -1, \quad c = t$$

$$P(t) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{1 - 4t}}{2}$$

There is no constant term in the formal power series for $P(t)$, hence $P(0) = 0$, hence we must take the minus sign:

$$P(t) = \frac{1 - \sqrt{1 - 4t}}{2}$$

Calculating the first few terms on a computer algebra system gives

$$t + t^2 + 2t^3 + 5t^4 + 14t^5 + 42t^6 + 132t^7 + 429t^8 + 1430t^9 + \cdots ,$$

so it looks like we're on the right track!

Binomial Theorem:

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

Binomial Theorem:

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

Generalized Binomial Coefficients:

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-(k-1))}{k!}$$

Binomial Theorem:

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

Generalized Binomial Coefficients:

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-(k-1))}{k!}$$

There are k factors in the numerator.

Binomial Theorem:

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

Generalized Binomial Coefficients:

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-(k-1))}{k!}$$

There are k factors in the numerator.
This is defined for any real (or complex) number n, and any nonnegative integer k.

Binomial Theorem:

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

Generalized Binomial Coefficients:

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-(k-1))}{k!}$$

There are k factors in the numerator.
This is defined for any real (or complex) number n, and any nonnegative integer k.

In particular, we can use this to generalize the Binomial Theorem to fractional exponents, obtaining formal power series.

Newton's Binomial Theorem:

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

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-(k-1))}{k!}$$

Newton's Binomial Theorem:

$$(x+y)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^{\alpha-k} y^k,$$
$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-(k-1))}{k!}$$

We are interested in $\sqrt{1-4t}$ so we set $x=1$, $y=-4t$, $\alpha=1/2$:

$$\sqrt{1-4t} = \sum_{k=0}^{\infty} \binom{1/2}{k} (-4t)^k = \sum_{k=0}^{\infty} (-1)^k 2^{2k} \binom{1/2}{k} t^k.$$

Newton's Binomial Theorem:

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

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-(k-1))}{k!}$$

We are interested in $\sqrt{1-4t}$ so we set $x=1$, $y=-4t$, $\alpha=1/2$:

$$\sqrt{1-4t} = \sum_{k=0}^{\infty} \binom{1/2}{k} (-4t)^k = \sum_{k=0}^{\infty} (-1)^k 2^{2k} \binom{1/2}{k} t^k.$$

What is $\binom{1/2}{k}$ for $k \geq 0$? (No combinatorial interpretation.)

$$\binom{1/2}{k} = \frac{1}{k!} \cdot \frac{1}{2}\left(\frac{1}{2} - 1\right)\left(\frac{1}{2} - 2\right)\cdots\left(\frac{1}{2} - (k-1)\right)$$

$$= \frac{1}{2} \cdot \frac{1}{k!} \cdot (-1)^{k-1} \cdot \left(1 - \frac{1}{2}\right)\left(2 - \frac{1}{2}\right)\cdots\left((k-1) - \frac{1}{2}\right)$$

$$= \frac{1}{2} \cdot \frac{1}{k!} \cdot (-1)^{k-1} \cdot \left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\cdots\left(\frac{2k-3}{2}\right)$$

$$= (-1)^{k-1} \cdot \frac{1}{2^k} \cdot \frac{1}{k!} \cdot (1 \cdot 3 \cdots (2k-3))$$

$$= (-1)^{k-1} \cdot \frac{1}{2^k} \cdot \frac{1}{k!} \cdot (1 \cdot 3 \cdots (2k-3)) \cdot \frac{2 \cdot 4 \cdots (2k-2)}{2^{k-1}(k-1)!}$$

$$= (-1)^{k-1} \cdot \frac{1}{2^{2k-1}} \cdot \frac{(2k-2)!}{k!(k-1)!}$$

Lemma:

$$\binom{1/2}{k} = \frac{(-1)^{k-1}}{2^{2k-1}} \cdot \frac{(2k-2)!}{k!(k-1)!}$$

Lemma:

$$\binom{1/2}{k} = \frac{(-1)^{k-1}}{2^{2k-1}} \cdot \frac{(2k-2)!}{k!(k-1)!}$$

Recall:

$$\sqrt{1-4t} = \sum_{k=0}^{\infty} (-1)^k 2^{2k} \binom{1/2}{k} t^k$$

Lemma:

$$\binom{1/2}{k} = \frac{(-1)^{k-1}}{2^{2k-1}} \cdot \frac{(2k-2)!}{k!(k-1)!}$$

Recall:

$$\sqrt{1-4t} = \sum_{k=0}^{\infty} (-1)^k 2^{2k} \binom{1/2}{k} t^k$$

The coefficient of t^k is

$$(-1)^k 2^{2k} \cdot \frac{(-1)^{k-1}}{2^{2k-1}} \cdot \frac{(2k-2)!}{k!(k-1)!} = -2\,\frac{(2k-2)!}{k!(k-1)!}$$

Lemma:

$$\binom{1/2}{k} = \frac{(-1)^{k-1}}{2^{2k-1}} \cdot \frac{(2k-2)!}{k!(k-1)!}$$

Recall:

$$\sqrt{1-4t} = \sum_{k=0}^{\infty} (-1)^k 2^{2k} \binom{1/2}{k} t^k$$

The coefficient of t^k is

$$(-1)^k 2^{2k} \cdot \frac{(-1)^{k-1}}{2^{2k-1}} \cdot \frac{(2k-2)!}{k!(k-1)!} = -2\,\frac{(2k-2)!}{k!(k-1)!}$$

Conclusion:

$$P(t) = \frac{1}{2}(1-\sqrt{1-4t}) \quad \Longrightarrow \quad P_n = \frac{(2n-2)!}{n!(n-1)!}$$

Now we have a formula, so we can evaluate P_n for any n directly; for example, the number of ways to multiply $n = 100$ factors is:

Now we have a formula, so we can evaluate P_n for any n directly; for example, the number of ways to multiply $n = 100$ factors is:

227508830794229349661819540395688853956041682601541047340

Now we have a formula, so we can evaluate P_n for any n directly; for example, the number of ways to multiply $n = 100$ factors is:

227508830794229349661819540395688853956041682601541047340

The *Catalan numbers* are this sequence shifted one step left:

$$C_n = P_{n+1} = \frac{(2n)!}{(n+1)!n!} = \frac{1}{n+1}\binom{2n}{n}.$$

The n in C_n counts the number of *operations*, not *arguments*.

Now we have a formula, so we can evaluate P_n for any n directly; for example, the number of ways to multiply $n = 100$ factors is:

227508830794229349661819540395688853956041682601541047340

The *Catalan numbers* are this sequence shifted one step left:

$$C_n = P_{n+1} = \frac{(2n)!}{(n+1)!n!} = \frac{1}{n+1}\binom{2n}{n}.$$

The n in C_n counts the number of *operations*, not *arguments*.

The Catalan numbers are named after the Belgian-French mathematician Eugène Charles Catalan (1814-1894).

Now we have a formula, so we can evaluate P_n for any n directly; for example, the number of ways to multiply $n = 100$ factors is:

227508830794229349661819540395688853956041682601541047340

The *Catalan numbers* are this sequence shifted one step left:

$$C_n = P_{n+1} = \frac{(2n)!}{(n+1)!n!} = \frac{1}{n+1}\binom{2n}{n}.$$

The n in C_n counts the number of *operations*, not *arguments*.

The Catalan numbers are named after the Belgian-French mathematician Eugène Charles Catalan (1814-1894).

Catalan also gave his name to the conjecture (1844) that the only two consecutive integers which are powers of natural numbers are $8 = 2^3$ and $9 = 3^2$. This was proved by Mihăilescu (2002/2004).

PART 2

PART 2

m-ary Catalan numbers: The original Catalan numbers can be generalized from a binary operation (taking two factors) to an m-ary operation (taking m factors).

PART 2

m-ary Catalan numbers: The original Catalan numbers can be generalized from a binary operation (taking two factors) to an m-ary operation (taking m factors).

In this case, the binomial formula generalizes very nicely:

$$\frac{1}{(m-1)n+1}\binom{mn}{n}$$

Setting $m = 2$ gives our formula for the binary Catalan numbers.

PART 2

m-ary Catalan numbers: The original Catalan numbers can be generalized from a binary operation (taking two factors) to an m-ary operation (taking m factors).

In this case, the binomial formula generalizes very nicely:

$$\frac{1}{(m-1)n+1}\binom{mn}{n}$$

Setting $m = 2$ gives our formula for the binary Catalan numbers.

The proof that we gave for $m = 2$ does not generalize: it would require explicit solution of a polynomial of degree m.

PART 2

m-ary Catalan numbers: The original Catalan numbers can be generalized from a binary operation (taking two factors) to an m-ary operation (taking m factors).

In this case, the binomial formula generalizes very nicely:

$$\frac{1}{(m-1)n+1}\binom{mn}{n}$$

Setting $m = 2$ gives our formula for the binary Catalan numbers.

The proof that we gave for $m = 2$ does not generalize: it would require explicit solution of a polynomial of degree m.

There is a (relatively) elementary proof that works for all m using convolution of formal power series in *Concrete Mathematics* by Graham, Knuth, and Patashnik (Section 7.5); I will call this GKP.

Just as in the binary case, we can regard the m-ary case as enumerating rooted planar complete trees, but now the trees are m-ary instead of binary: every internal node has exactly m children.

Just as in the binary case, we can regard the m-ary case as enumerating rooted planar complete trees, but now the trees are m-ary instead of binary: every internal node has exactly m children.

We must remember that n is the number of internal nodes (m-ary multiplications from the algebraic point of view) not the number of leaf nodes (arguments or factors from the algebraic point of view).

Just as in the binary case, we can regard the m-ary case as enumerating rooted planar complete trees, but now the trees are m-ary instead of binary: every internal node has exactly m children.

We must remember that n is the number of internal nodes (m-ary multiplications from the algebraic point of view) not the number of leaf nodes (arguments or factors from the algebraic point of view).

Here is a table for $2 \leq m \leq 6$ and $1 \leq n \leq 9$:

Just as in the binary case, we can regard the m-ary case as enumerating rooted planar complete trees, but now the trees are m-ary instead of binary: every internal node has exactly m children.

We must remember that n is the number of internal nodes (m-ary multiplications from the algebraic point of view) not the number of leaf nodes (arguments or factors from the algebraic point of view).

Here is a table for $2 \leq m \leq 6$ and $1 \leq n \leq 9$:

${}_m\backslash{}^n$	1	2	3	4	5	6	7	8	9
2	1	2	5	14	42	132	429	1430	4862
3	1	3	12	55	273	1428	7752	43263	246675
4	1	4	22	140	969	7084	53820	420732	3362260
5	1	5	35	285	2530	23751	231880	2330445	23950355
6	1	6	51	506	5481	62832	749398	9203634	115607310

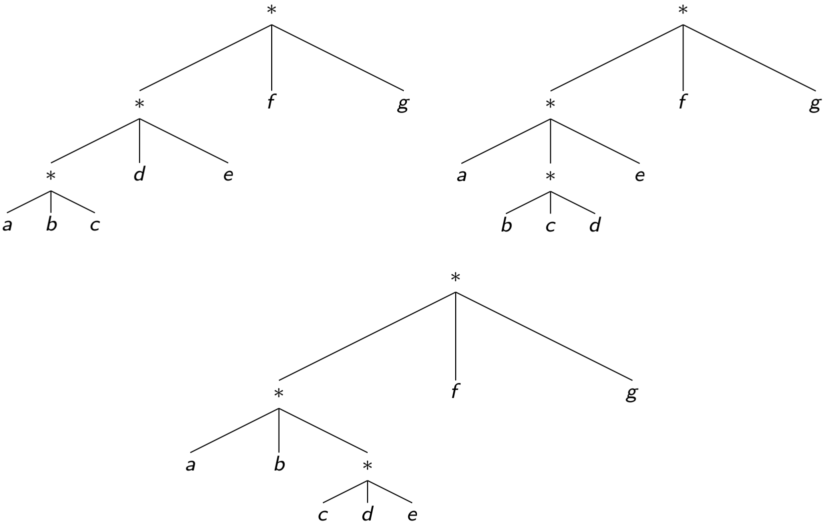
Just as in the binary case, we can regard the m-ary case as enumerating rooted planar complete trees, but now the trees are m-ary instead of binary: every internal node has exactly m children.

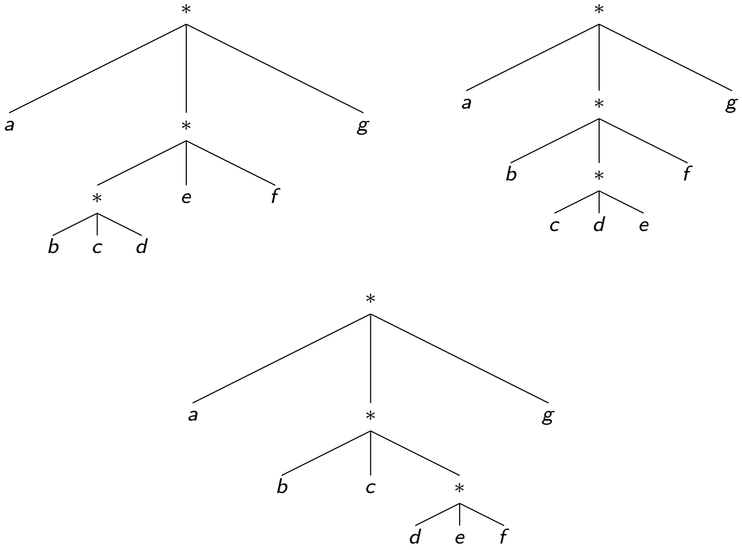
We must remember that n is the number of internal nodes (m-ary multiplications from the algebraic point of view) not the number of leaf nodes (arguments or factors from the algebraic point of view).

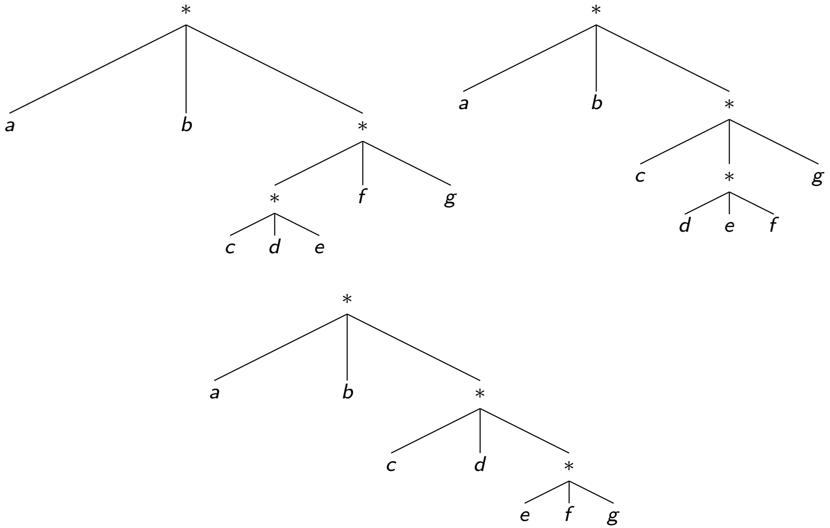
Here is a table for $2 \leq m \leq 6$ and $1 \leq n \leq 9$:

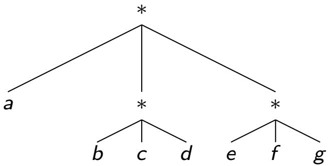
$_m\backslash^n$	1	2	3	4	5	6	7	8	9
2	1	2	5	14	42	132	429	1430	4862
3	1	3	12	55	273	1428	7752	43263	246675
4	1	4	22	140	969	7084	53820	420732	3362260
5	1	5	35	285	2530	23751	231880	2330445	23950355
6	1	6	51	506	5481	62832	749398	9203634	115607310

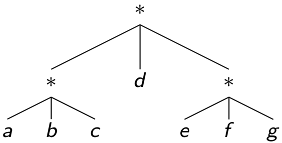
In particular, there should be 12 ternary trees with 3 internal nodes:

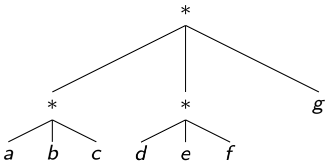

*
* f g
* d e
a b c
*
* f g
a * e
b c d
*
* f g
a b *
c d e


*
a * g
* e f
b c d
*
a * g
b * f
c d e
*
a * g
b c *
d e f


*
a
b
*
*
f
g
c
d
e
*
a
b
*
c
*
g
d
e
f
*
a
b
*
c
d
*
e
f
g


*
a
*
*
b
c
d
e
f
g


*
*
d
*
a
b
c
e
f
g


*
*
*
g
a
b
c
d
e
f

In the rest of this talk, I follow GKP very closely. Their proof of the m-ary Catalan formula starts with the notion of *Raney sequence*.

In the rest of this talk, I follow GKP very closely. Their proof of the m-ary Catalan formula starts with the notion of *Raney sequence*.

Lemma (Raney's Lemma)

Let $(a_0, a_1, \ldots, a_{2n})$ be a sequence of integers with $\sum_{i=0}^{2n} a_i = 1$. Then exactly one of the $2n+1$ cyclic shifts of the sequence has the property that all of its partial sums are positive:

$$(a_0, a_1, \ldots, a_{2n}), \quad (a_1, a_2, \ldots, a_0), \quad \ldots, \quad (a_{2n}, a_0, \ldots, a_{2n-1}).$$

In the rest of this talk, I follow GKP very closely. Their proof of the m-ary Catalan formula starts with the notion of *Raney sequence*.

Lemma (Raney's Lemma)

Let $(a_0, a_1, \ldots, a_{2n})$ be a sequence of integers with $\sum_{i=0}^{2n} a_i = 1$. Then exactly one of the $2n+1$ cyclic shifts of the sequence has the property that all of its partial sums are positive:

$$(a_0, a_1, \ldots, a_{2n}), \quad (a_1, a_2, \ldots, a_0), \quad \ldots, \quad (a_{2n}, a_0, \ldots, a_{2n-1}).$$

Definition

A *Raney sequence* is a sequence of integers such that

$$\sum_{i=0}^{2n} a_i = 1 \qquad \text{and} \qquad \sum_{i=0}^{k} a_i \geq 1 \ \ (k = 0, \ldots, 2n).$$

In particular, consider the Raney sequences $(a_0, a_1, \ldots, a_{2n})$ which contain only $+1$ and -1.

In particular, consider the Raney sequences $(a_0, a_1, \ldots, a_{2n})$ which contain only $+1$ and -1.

To get sum $+1$ we need $n+1$ copies of $+1$ and n copies of -1, which gives a total of $\binom{2n+1}{n}$ sequences.

In particular, consider the Raney sequences $(a_0, a_1, \ldots, a_{2n})$ which contain only $+1$ and -1.

To get sum $+1$ we need $n+1$ copies of $+1$ and n copies of -1, which gives a total of $\binom{2n+1}{n}$ sequences.

Raney's Lemma tells us that exactly $\frac{1}{2n+1}$ of these have all their partial sums positive.

In particular, consider the Raney sequences $(a_0, a_1, \ldots, a_{2n})$ which contain only $+1$ and -1.

To get sum $+1$ we need $n+1$ copies of $+1$ and n copies of -1, which gives a total of $\binom{2n+1}{n}$ sequences.

Raney's Lemma tells us that exactly $\frac{1}{2n+1}$ of these have all their partial sums positive.

Altogether we obtain this number of $+1/-1$ Raney sequences:

$$\frac{1}{2n+1}\binom{2n+1}{n} = \frac{1}{2n+1} \cdot \frac{(2n+1)!}{(n+1)!n!} = \frac{(2n)!}{(n+1)!n!}$$
$$= \frac{1}{n+1}\binom{2n}{n}$$

In particular, consider the Raney sequences $(a_0, a_1, \dots, a_{2n})$ which contain only $+1$ and -1.

To get sum $+1$ we need $n+1$ copies of $+1$ and n copies of -1, which gives a total of $\binom{2n+1}{n}$ sequences.

Raney's Lemma tells us that exactly $\frac{1}{2n+1}$ of these have all their partial sums positive.

Altogether we obtain this number of $+1/-1$ Raney sequences:

$$\frac{1}{2n+1}\binom{2n+1}{n} = \frac{1}{2n+1} \cdot \frac{(2n+1)!}{(n+1)!n!} = \frac{(2n)!}{(n+1)!n!}$$
$$= \frac{1}{n+1}\binom{2n}{n}$$

The Catalan number, C_n!

We give a bijection from placements of parentheses to $+1/-1$ Raney sequences.

We give a bijection from placements of parentheses to $+1/-1$ Raney sequences. We illustrate with an example from GKP:

$$a((bc)(de))$$

We give a bijection from placements of parentheses to $+1/-1$ Raney sequences. We illustrate with an example from GKP:

$$a((bc)(de))$$

Explicitly write in the multiplication symbols, and add an outermost pair of parentheses so that there are as many pairs of parentheses as multiplication symbols:

$$(a \cdot ((b \cdot c) \cdot (d \cdot e)))$$

We give a bijection from placements of parentheses to $+1/-1$ Raney sequences. We illustrate with an example from GKP:

$$a((bc)(de))$$

Explicitly write in the multiplication symbols, and add an outermost pair of parentheses so that there are as many pairs of parentheses as multiplication symbols:

$$(a \cdot ((b \cdot c) \cdot (d \cdot e)))$$

Erase the arguments and the left parentheses:

$$\cdot\cdot)\cdot\cdot)))$$

We give a bijection from placements of parentheses to $+1/-1$ Raney sequences. We illustrate with an example from GKP:

$$a((bc)(de))$$

Explicitly write in the multiplication symbols, and add an outermost pair of parentheses so that there are as many pairs of parentheses as multiplication symbols:

$$(a \cdot ((b \cdot c) \cdot (d \cdot e)))$$

Erase the arguments and the left parentheses:

$$\cdot\cdot)\cdot\cdot)))$$

Replace each multiplication by $+1$ and each right parenthesis by -1, and add an extra $+1$ at the beginning:

$$+1, +1, +1, -1, +1, +1, -1, -1, -1$$

What is the m-ary generalization?

What is the m-ary generalization?

Definition

For an integer $m \geq 2$, an *m-Raney sequence* is a sequence

$$(a_0, a_1, \ldots, a_{mn})$$

of the numbers 1 and $1 - m$ (so -1 corresponds to $m = 2$) whose total sum is 1 and whose partial sums are all positive.

What is the m-ary generalization?

Definition

For an integer $m \geq 2$, an *m-Raney sequence* is a sequence

$$(a_0, a_1, \ldots, a_{mn})$$

of the numbers 1 and $1 - m$ (so -1 corresponds to $m = 2$) whose total sum is 1 and whose partial sums are all positive.

If $1 - m$ occurs k times (so 1 occurs $mn + 1 - k$ times) then

$$k(1-m) + (mn + 1 - k) = 1 \implies -km + mn = 0 \implies k = n.$$

What is the m-ary generalization?

Definition

For an integer $m \geq 2$, an *m-Raney sequence* is a sequence

$$(a_0, a_1, \ldots, a_{mn})$$

of the numbers 1 and $1 - m$ (so -1 corresponds to $m = 2$) whose total sum is 1 and whose partial sums are all positive.

If $1 - m$ occurs k times (so 1 occurs $mn + 1 - k$ times) then

$$k(1-m) + (mn + 1 - k) = 1 \implies -km + mn = 0 \implies k = n.$$

So each sequence has n occurrences of $1 - m$ and $mn + 1 - n$ occurrences of 1, and each sequence has length $mn + 1$.

The total number of such sequences (without the positivity condition on partial sums) is this binomial coefficient:

$$\binom{mn+1}{n}$$

The total number of such sequences (without the positivity condition on partial sums) is this binomial coefficient:

$$\binom{mn+1}{n}$$

Using Raney's Lemma to include the positivity condition gives:

$$\begin{aligned}
\frac{1}{mn+1}\binom{mn+1}{n} &= \frac{1}{mn+1} \cdot \frac{(mn+1)!}{((m-1)n+1)!n!} \\
&= \frac{(mn)!}{((m-1)n+1)!n!} \\
&= \frac{1}{((m-1)n+1)}\binom{mn}{n}.
\end{aligned}$$

The total number of such sequences (without the positivity condition on partial sums) is this binomial coefficient:

$$\binom{mn+1}{n}$$

Using Raney's Lemma to include the positivity condition gives:

$$\begin{aligned}\frac{1}{mn+1}\binom{mn+1}{n} &= \frac{1}{mn+1}\cdot\frac{(mn+1)!}{((m-1)n+1)!n!}\\ &= \frac{(mn)!}{((m-1)n+1)!n!}\\ &= \frac{1}{((m-1)n+1)}\binom{mn}{n}.\end{aligned}$$

This is the formula we've seen before for m-ary Catalan numbers.

Why does this formula enumerate placements of parentheses for an m-ary operation, or complete rooted planar m-ary trees?

Why does this formula enumerate placements of parentheses for an m-ary operation, or complete rooted planar m-ary trees?

We need to construct a bijection between these placements of parentheses and the m-Raney sequences of the last definition.

Why does this formula enumerate placements of parentheses for an m-ary operation, or complete rooted planar m-ary trees?

We need to construct a bijection between these placements of parentheses and the m-Raney sequences of the last definition.

The trivial sequence 1 of length 1 is an m-Raney sequence.

Why does this formula enumerate placements of parentheses for an m-ary operation, or complete rooted planar m-ary trees?

We need to construct a bijection between these placements of parentheses and the m-Raney sequences of the last definition.

The trivial sequence 1 of length 1 is an m-Raney sequence.

If we form a sequence of m sequences which are m-Raney sequences and follow it with the number $1 - m$, then we get another m-Raney sequence: the partial sums (over the sequences) increase from 1 to m and then drop to 1.

Why does this formula enumerate placements of parentheses for an m-ary operation, or complete rooted planar m-ary trees?

We need to construct a bijection between these placements of parentheses and the m-Raney sequences of the last definition.

The trivial sequence 1 of length 1 is an m-Raney sequence.

If we form a sequence of m sequences which are m-Raney sequences and follow it with the number $1 - m$, then we get another m-Raney sequence: the partial sums (over the sequences) increase from 1 to m and then drop to 1.

Conversely, it can be shown with a little more work that all m-Raney sequences arise this way.

The recursive construction of m-Raney sequences we have just described corresponds to the equation

$$C_n^{(m)} = \sum_{n_1+n_2+\cdots+n_m+1=n} C_{n_1}^{(m)} C_{n_2}^{(m)} \cdots C_{n_m}^{(m)},$$

The recursive construction of m-Raney sequences we have just described corresponds to the equation

$$C_n^{(m)} = \sum_{n_1+n_2+\cdots+n_m+1=n} C_{n_1}^{(m)} C_{n_2}^{(m)} \cdots C_{n_m}^{(m)},$$

which is the same recursion that counts placements of parentheses for an m-ary operation.

The recursive construction of m-Raney sequences we have just described corresponds to the equation

$$C_n^{(m)} = \sum_{n_1+n_2+\cdots+n_m+1=n} C_{n_1}^{(m)} C_{n_2}^{(m)} \cdots C_{n_m}^{(m)},$$

which is the same recursion that counts placements of parentheses for an m-ary operation.

Note that the $+1$ under the summation sign comes from the fact that we are counting *operations* not *arguments*:

The recursive construction of m-Raney sequences we have just described corresponds to the equation

$$C_n^{(m)} = \sum_{n_1+n_2+\cdots+n_m+1=n} C_{n_1}^{(m)} C_{n_2}^{(m)} \cdots C_{n_m}^{(m)},$$

which is the same recursion that counts placements of parentheses for an m-ary operation.

Note that the $+1$ under the summation sign comes from the fact that we are counting *operations* not *arguments*: if we combine m factors involving respectively $n_1, \ldots, n_m$ operations then when we multiply those factors we introduce one more operation.

PART 3

PART 3

What happens if we assume that the binary operation is commutative?

PART 3

What happens if we assume that the binary operation is commutative?

In this case, it's convenient to think of a single generator x subject to a commutative but nonassociative operation.

PART 3

What happens if we assume that the binary operation is commutative?

In this case, it's convenient to think of a single generator x subject to a commutative but nonassociative operation.

We want to count the distinct "powers" of x in each degree:

PART 3

What happens if we assume that the binary operation is commutative?

In this case, it's convenient to think of a single generator x subject to a commutative but nonassociative operation.

We want to count the distinct "powers" of x in each degree:

$$x, \qquad x^2, \qquad x^2x = xx^2 \text{ (commutativity)}, \qquad x^2x^2,$$
$$(x^2x)x = (xx^2)x = x(x^2x) = x(xx^2) \text{ (commutativity)}.$$

PART 3

What happens if we assume that the binary operation is commutative?

In this case, it's convenient to think of a single generator x subject to a commutative but nonassociative operation.

We want to count the distinct "powers" of x in each degree:

$$x, \qquad x^2, \qquad x^2x = xx^2 \text{ (commutativity)}, \qquad x^2x^2,$$
$$(x^2x)x = (xx^2)x = x(x^2x) = x(xx^2) \text{ (commutativity)}.$$

Degrees $n = 1, 2, 3, 4$ have 1, 1, 1, 2 distinct n-th powers of x.

PART 3

What happens if we assume that the binary operation is commutative?

In this case, it's convenient to think of a single generator x subject to a commutative but nonassociative operation.

We want to count the distinct "powers" of x in each degree:

$$x, \qquad x^2, \qquad x^2x = xx^2 \ \text{(commutativity)}, \qquad x^2x^2,$$
$$(x^2x)x = (xx^2)x = x(x^2x) = x(xx^2) \ \text{(commutativity)}.$$

Degrees $n = 1, 2, 3, 4$ have 1, 1, 1, 2 distinct n-th powers of x.

In each equivalence class under commutativity, we need to choose one representative normal form; for example, x^2x and $(x^2x)x$.

PART 3

What happens if we assume that the binary operation is commutative?

In this case, it's convenient to think of a single generator x subject to a commutative but nonassociative operation.

We want to count the distinct "powers" of x in each degree:

$$x, \qquad x^2, \qquad x^2x = xx^2 \ \text{(commutativity)}, \qquad x^2x^2,$$
$$(x^2x)x = (xx^2)x = x(x^2x) = x(xx^2) \ \text{(commutativity)}.$$

Degrees $n = 1, 2, 3, 4$ have 1, 1, 1, 2 distinct n-th powers of x.

In each equivalence class under commutativity, we need to choose one representative normal form; for example, x^2x and $(x^2x)x$.

We also need to choose a total order on the normal forms that respects the degrees.

We can solve both problems with an algorithm that generates the normal forms by degree and by total order within each degree.

We can solve both problems with an algorithm that generates the normal forms by degree and by total order within each degree.

- Set $Q[1] \leftarrow [\,x\,]$ (*List containing the single element* x)
- For n from 2 to MAXDEG do
  - Set $Q[n] \leftarrow [\,]$ (*Empty list*)
  - For j from 1 to $\lfloor (n-1)/2 \rfloor$ do: (*Stop before we reach* $n/2$)
    - (*In this loop, left factor has higher degree than right factor*)
    - For x in $Q[n-j]$ do for y in $Q[j]$ do:
      Adjoin $[x, y]$ to the list $Q[n]$.
  - If n is even then (*Special case: two factors of same degree*)
    - For i to length($Q[n/2]$) do for j from i to length($Q[n/2]$) do
    - (*In this loop, the left factor precedes the right factor in the total order on degree* i)
      Adjoin $[Q[n/2][i], Q[n/2][j]]$ to the list $Q[n]$.

The resulting sequence is the *Wedderburn-Etherington numbers*:

The resulting sequence is the *Wedderburn-Etherington numbers*:

$$1,\ 1,\ 1,\ 2,\ 3,\ 6,\ 11,\ 23,\ 46,\ 98,\ 207,\ 451,\ 983,\ 2179,\ 4850,\ \ldots$$

The resulting sequence is the *Wedderburn-Etherington numbers*:

$$1,\ 1,\ 1,\ 2,\ 3,\ 6,\ 11,\ 23,\ 46,\ 98,\ 207,\ 451,\ 983,\ 2179,\ 4850,\ \ldots$$

To explain the number 3 in degree 5, we have:

$$4+1: \quad (((x^2)x)x)x, \quad (x^2x^2)x; \quad 3+2: \quad (x^2x)x^2.$$

The resulting sequence is the *Wedderburn-Etherington numbers*:

1, 1, 1, 2, 3, 6, 11, 23, 46, 98, 207, 451, 983, 2179, 4850, ...

To explain the number 3 in degree 5, we have:

$$4+1: \quad (((x^2)x)x)x, \quad (x^2x^2)x; \quad 3+2: \quad (x^2x)x^2.$$

To explain the number 6 in degree 6, we have:

$$5+1: \quad ((((x^2)x)x)x)x, \quad ((x^2x^2)x)x, \quad ((x^2x)x^2)x;$$
$$4+2: \quad (((x^2)x)x)x^2, \quad (x^2x^2)x^2; \quad 3+3: \quad (x^2x)(x^2x).$$

The resulting sequence is the *Wedderburn-Etherington numbers*:

1, 1, 1, 2, 3, 6, 11, 23, 46, 98, 207, 451, 983, 2179, 4850, ...

To explain the number 3 in degree 5, we have:

$$4+1: \quad (((x^2)x)x)x, \quad (x^2x^2)x; \quad 3+2: \quad (x^2x)x^2.$$

To explain the number 6 in degree 6, we have:

$$5+1: \quad ((((x^2)x)x)x)x, \quad ((x^2x^2)x)x, \quad ((x^2x)x^2)x;$$
$$4+2: \quad (((x^2)x)x)x^2, \quad (x^2x^2)x^2; \quad 3+3: \quad (x^2x)(x^2x).$$

To explain the number 11 in degree 7, we have:

$$[6+1]\ 6\cdot 1 = 6, \quad [5+2]\ 3\cdot 1 = 3, \quad [4+3]\ 2\cdot 1 = 2: \quad \text{total } 11.$$

The resulting sequence is the *Wedderburn-Etherington numbers*:

1, 1, 1, 2, 3, 6, 11, 23, 46, 98, 207, 451, 983, 2179, 4850, ...

To explain the number 3 in degree 5, we have:

$$4+1: \quad (((x^2)x)x)x, \quad (x^2x^2)x; \quad 3+2: \quad (x^2x)x^2.$$

To explain the number 6 in degree 6, we have:

$$\begin{array}{ll} 5+1: & ((((x^2)x)x)x)x, \quad ((x^2x^2)x)x, \quad ((x^2x)x^2)x; \\ 4+2: & (((x^2)x)x)x^2, \quad (x^2x^2)x^2; \quad 3+3: \quad (x^2x)(x^2x). \end{array}$$

To explain the number 11 in degree 7, we have:

$$[6+1]\ \ 6\cdot 1 = 6, \quad [5+2]\ \ 3\cdot 1 = 3, \quad [4+3]\ \ 2\cdot 1 = 2: \quad \text{total}\ \ 11.$$

To explain the number 23 in degree 8, we have:

$$[7+1]\ 11\cdot 1,\ [6+2]\ 6\cdot 1,\ [5+3]\ 3\cdot 1,\ [4+4]\ \binom{2+1}{2}:\ \text{total}\ 23.$$

Write Q_n for the number of commutative nonassociative n-th powers of x.

Write Q_n for the number of commutative nonassociative n-th powers of x. The last algorithm gives this equation:

$$Q_n = \sum_{i=1}^{\lfloor (n-1)/2 \rfloor} Q_{n-i} Q_i \quad + \quad (n \text{ even}) \, \frac{1}{2} Q_{n/2}(Q_{n/2} + 1).$$

Write Q_n for the number of commutative nonassociative n-th powers of x. The last algorithm gives this equation:

$$Q_n = \sum_{i=1}^{\lfloor (n-1)/2 \rfloor} Q_{n-i} Q_i \quad + \quad (n \text{ even}) \, \frac{1}{2} Q_{n/2}(Q_{n/2}+1).$$

The term which only occurs for n even means that we choose two factors from degree $n/2$, allowing repetitions; it can also be written

$$\binom{Q_{n/2}}{2} + Q_{n/2} = \binom{Q_{n/2}+1}{2}.$$

Write Q_n for the number of commutative nonassociative n-th powers of x. The last algorithm gives this equation:

$$Q_n = \sum_{i=1}^{\lfloor (n-1)/2 \rfloor} Q_{n-i} Q_i \quad + \quad (n \text{ even})\, \frac{1}{2} Q_{n/2}(Q_{n/2} + 1).$$

The term which only occurs for n even means that we choose two factors from degree $n/2$, allowing repetitions; it can also be written

$$\binom{Q_{n/2}}{2} + Q_{n/2} = \binom{Q_{n/2} + 1}{2}.$$

As before, define the formal power series with Q_n as coefficients:

$$Q(t) = \sum_{n=1}^{\infty} Q_n t^n.$$

Let's see what happens when we expand $Q(t)^2$. We easily see,

$$Q(t)^2 = \Big(\sum_{i=1}^{\infty} Q_i t^i\Big)\Big(\sum_{j=1}^{\infty} Q_j t^j\Big) = \sum_{n=2}^{\infty}\Big(\sum_{i+j=n} Q_i Q_j\Big) t^n.$$

Let's see what happens when we expand $Q(t)^2$. We easily see,

$$Q(t)^2 = \Big(\sum_{i=1}^{\infty} Q_i t^i\Big)\Big(\sum_{j=1}^{\infty} Q_j t^j\Big) = \sum_{n=2}^{\infty}\Big(\sum_{i+j=n} Q_i Q_j\Big) t^n.$$

To relate $\sum_{i+j=n} Q_i Q_j$ to Q_n we must distinguish even and odd n.

Let's see what happens when we expand $Q(t)^2$. We easily see,

$$Q(t)^2 = \Big(\sum_{i=1}^{\infty} Q_i t^i\Big)\Big(\sum_{j=1}^{\infty} Q_j t^j\Big) = \sum_{n=2}^{\infty}\Big(\sum_{i+j=n} Q_i Q_j\Big)t^n.$$

To relate $\sum_{i+j=n} Q_i Q_j$ to Q_n we must distinguish even and odd n.
If n is odd, then

$$\begin{aligned}\sum_{i+j=n} Q_i Q_j &= \sum_{i=1}^{\lfloor (n-1)/2 \rfloor} Q_{n-i} Q_i \quad + \sum_{i=1}^{\lfloor (n-1)/2 \rfloor} Q_i Q_{n-i} \\ &= 2Q_n.\end{aligned}$$

Let's see what happens when we expand $Q(t)^2$. We easily see,

$$Q(t)^2 = \Big(\sum_{i=1}^{\infty} Q_i t^i\Big)\Big(\sum_{j=1}^{\infty} Q_j t^j\Big) = \sum_{n=2}^{\infty}\Big(\sum_{i+j=n} Q_i Q_j\Big) t^n.$$

To relate $\sum_{i+j=n} Q_i Q_j$ to Q_n we must distinguish even and odd n. If n is odd, then

$$\begin{aligned}\sum_{i+j=n} Q_i Q_j &= \sum_{i=1}^{\lfloor (n-1)/2\rfloor} Q_{n-i} Q_i \quad + \sum_{i=1}^{\lfloor (n-1)/2\rfloor} Q_i Q_{n-i} \\ &= 2Q_n.\end{aligned}$$

But if n is even, then

$$\begin{aligned}\sum_{i+j=n} Q_i Q_j &= \sum_{i=1}^{\lfloor (n-1)/2\rfloor} Q_{n-i} Q_i \quad + \sum_{i=1}^{\lfloor (n-1)/2\rfloor} Q_i Q_{n-i} + Q_{n/2}^2 \\ &= 2Q_n - Q_{n/2}.\end{aligned}$$

From this we see that $Q(t)$ satisfies this functional equation:

$$Q(t)^2 = 2(Q(t) - t) - Q(t^2)$$

From this we see that $Q(t)$ satisfies this functional equation:

$$Q(t)^2 = 2(Q(t) - t) - Q(t^2)$$

I am not aware of an exact solution of this equation.

From this we see that $Q(t)$ satisfies this functional equation:

$$Q(t)^2 = 2(Q(t) - t) - Q(t^2)$$

I am not aware of an exact solution of this equation.

Here are the two original papers by Wedderburn and Etherington:

From this we see that $Q(t)$ satisfies this functional equation:

$$Q(t)^2 = 2(Q(t) - t) - Q(t^2)$$

I am not aware of an exact solution of this equation.

Here are the two original papers by Wedderburn and Etherington:

J. H. M. Wedderburn:
The Functional Equation $g(x^2) = 2\alpha x + [g(x)]^2$.
Annals of Mathematics. Second Series.
Vol. 24, No. 2 (Dec., 1922), pp. 121-140.

From this we see that $Q(t)$ satisfies this functional equation:

$$Q(t)^2 = 2(Q(t) - t) - Q(t^2)$$

I am not aware of an exact solution of this equation.

Here are the two original papers by Wedderburn and Etherington:

J. H. M. Wedderburn:
The Functional Equation $g(x^2) = 2\alpha x + [g(x)]^2$.
Annals of Mathematics. Second Series.
Vol. 24, No. 2 (Dec., 1922), pp. 121-140.

I. M. H. Etherington:
Non-associate powers and a functional equation.
The Mathematical Gazette
21 (1937), 36-39; addendum 21 (1937), 153.