Notes on the Catalan problem

Daniele Paolo Scarpazza
Politecnico di Milano
<scarpazz@elet.polimi.it>

March 16th, 2004

An overview of Catalan problems

- Catalan numbers appear as the solution of a variety of problems;
- they were first described in the 18th century by Leonhard Euler (working on polygon triangulation);
- they are named after Eugene Catalan, a belgian mathematician which found their expression (working on parenthesizations).

A Catalan Problem: Balanced Parentheses

“Determine the number of balanced strings of parentheses of length $2n$”.

A string of parentheses is an ordered collection of symbols “(” and “)”.

Balanced: same number of open and close parentheses, and every prefix of the string has at least as many open parentheses as close parentheses;

Example: ( ) (( ) ( )) is balanced; strings) (( )) and (( ) ( ) are not.

n					C(n)
0	empty string				1
1	( )				1
2	( ) ( )	(( ))			2
3	( ) ( ) ( )	( ) (( ))	(( )) ( )	(( ) ( ))	5
	((( )))				
4	( ) ( ) ( ) ( )	( ) ( ) (( ))	( ) (( )) ( )	( ) (( ) ( ))	14
	( ) ((( )))	(( )) ( ) ( )	(( )) (( ))	(( ) ( )) ( )	
	(( ) ( ) ( ))	(( ) (( )))	((( ))) ( )	((( )) ( ))	
	((( ) ( )))	(((( ))))			
5	( ) ( ) ( ) ( ) ( )	( ) ( ) ( ) (( ))	( ) ( ) (( )) ( )	( ) ( ) (( ) ( ))	42
	( ) ( ) ((( )))	( ) (( )) ( ) ( )	( ) (( )) (( ))	( ) (( ) ( )) ( )	
	( ) (( ) ( ) ( ))	( ) (( ) (( )))	( ) ((( ))) ( )	( ) ((( )) ( ))	
	( ) ((( ) ( )))	( ) (((( ))))	(( )) ( ) ( ) ( )	(( )) ( ) (( ))	
	(( )) (( )) ( )	(( )) (( ) ( ))	(( )) ((( )))	(( ) ( )) ( ) ( )	
	(( ) ( )) (( ))	(( ) ( ) ( )) ( )	(( ) ( ) ( ) ( ))	(( ) ( ) (( )))	
	(( ) (( ))) ( )	(( ) (( )) ( ))	(( ) (( ) ( )))	(( ) ((( ))))	
	((( ))) ( ) ( )	((( ))) (( ))	((( )) ( )) ( )	((( )) ( ) ( ))	
	((( )) (( )))	((( ) ( ))) ( )	((( ) ( )) ( ))	((( ) ( ) ( )))	
	((( ) (( ))))	(((( )))) ( )	(((( ))) ( ))	(((( )) ( )))	
	(((( ) ( ))))	((((( )))))			

Another one: Mountain Ranges

"Determine the number of *mountain landscapes* which can be formed with n upstrokes and n downstrokes."

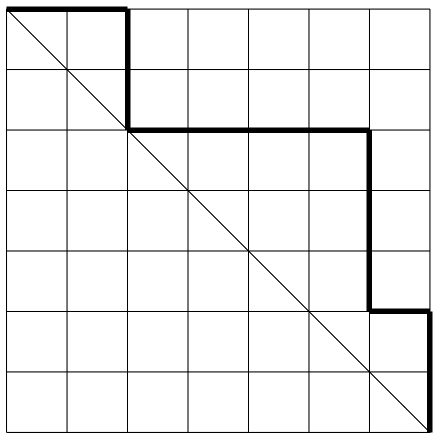
Mountain range: polyline of upstrokes "/" and downstrokes "\"; its extreme points lie on the same horizontal line, and no segments cross it.

```
n                                                                    C(n)
--------------------------------------------------------------------------
0       *                                                               1
--------------------------------------------------------------------------
1       /\                                                              1
--------------------------------------------------------------------------
2                    /\                                                 2
        /\/\        /  \
--------------------------------------------------------------------------
3                                                           /\          5
                       /\          /\          /\/\        /  \
        /\/\/\      /\/  \      /  \/\      /    \      /    \
```

Another one: Diagonal-avoiding paths on a lattice

"Given a $n \times n$ lattice, determine the number of paths of length $2n$ which do not cross the diagonal."

In a finite lattice $(i, j) : 1 \leq i \leq n, 1 \leq j \leq n$, a path is a connected sequence of "west" or "south" segments from node $(1, 1)$ to node (n, n).



Sample path in a 7 $\times$ 7 lattice, corresponding to string (( )) (((( ))) ( )).

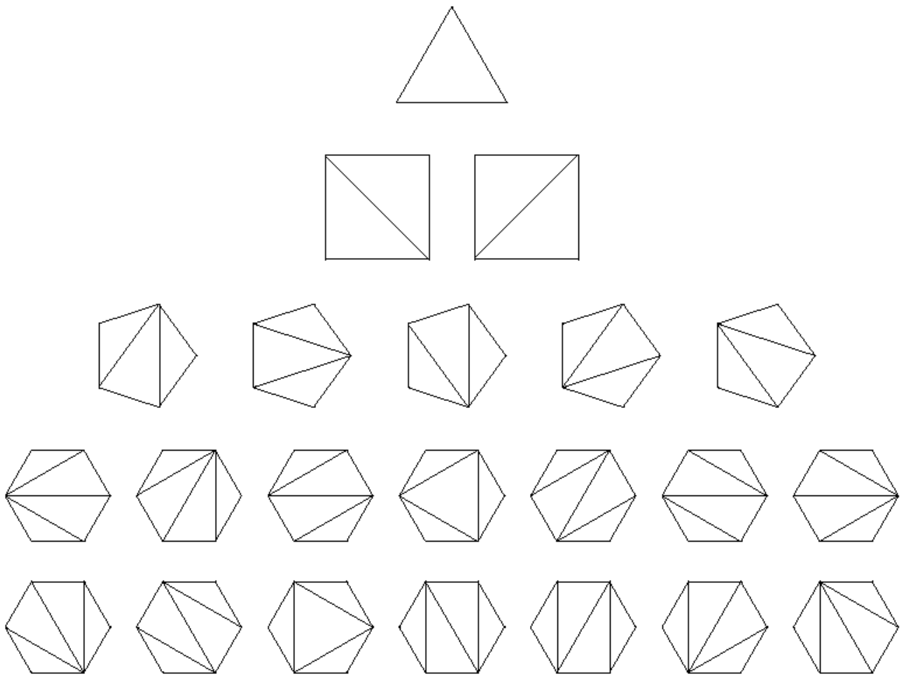
Another one: Multiplication precedence

"Determine the number of ways in which $n+1$ factors can be multiplied together, according to the precedence of multiplications."

n			C(n)
0	a		1
1	$a \cdot b$		1
2	$(a \cdot b) \cdot c$	$a \cdot (b \cdot c)$	2
3	$((a \cdot b) \cdot c) \cdot d$ $(a \cdot (b \cdot c)) \cdot d$ $a \cdot (b \cdot (c \cdot d))$	$(a \cdot b) \cdot (c \cdot d)$ $a \cdot ((b \cdot c) \cdot d)$	5
4	$(((a \cdot b) \cdot c) \cdot d) \cdot e$ $((a \cdot b) \cdot (c \cdot d)) \cdot e$ $(a \cdot b) \cdot (c \cdot (d \cdot e))$ $(a \cdot (b \cdot c)) \cdot (d \cdot e)$ $(a \cdot (b \cdot (c \cdot d))) \cdot e$ $a \cdot ((b \cdot c) \cdot (d \cdot e))$ $a \cdot (b \cdot ((c \cdot d) \cdot e))$	$((a \cdot b) \cdot c) \cdot (d \cdot e)$ $(a \cdot b) \cdot ((c \cdot d) \cdot e)$ $((a \cdot (b \cdot c)) \cdot d) \cdot e$ $(a \cdot ((b \cdot c) \cdot d)) \cdot e$ $a \cdot (((b \cdot c) \cdot d) \cdot e)$ $a \cdot ((b \cdot (c \cdot d)) \cdot e)$ $a \cdot (b \cdot (c \cdot (d \cdot e)))$	14

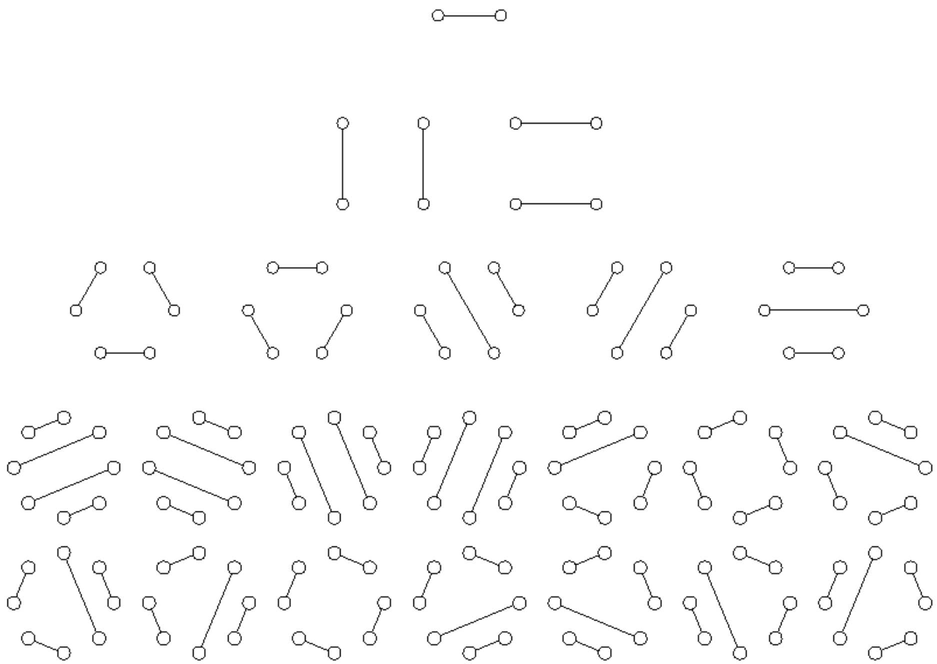
Another one: Convex polygon triangulation

"Determine the number of ways in which a convex polygon with $n+2$ edges can be triangulated."



Another one: Handshakes across a table

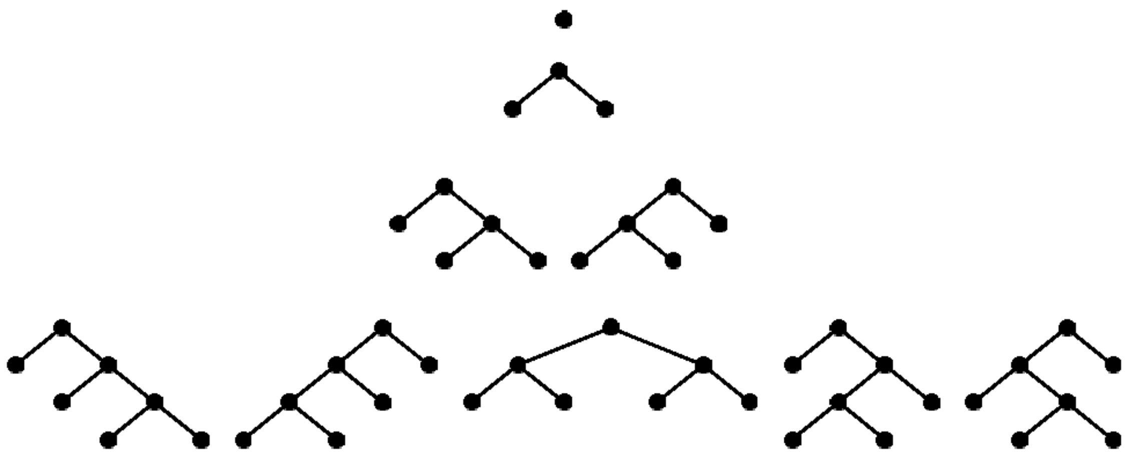
"Determine the number of ways in which $2n$ people sitting around a table can shake hands without crossing their arms".



Another one: Binary rooted trees

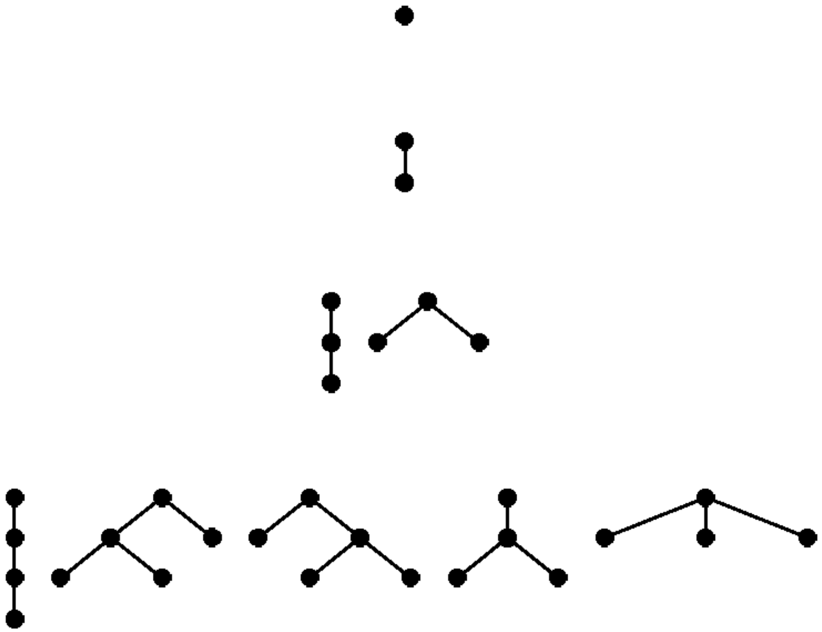
"Determine the number of binary rooted trees with n internal nodes."

Each non-leaf is an *internal node*. Binary rooted trees with n internal nodes and n ranging from 0 to 3:



Another one: Plane trees

"Determine the number of plane trees with n edges." A plane tree is such that it is possible to draw it on a plane with no edges crossing each other.

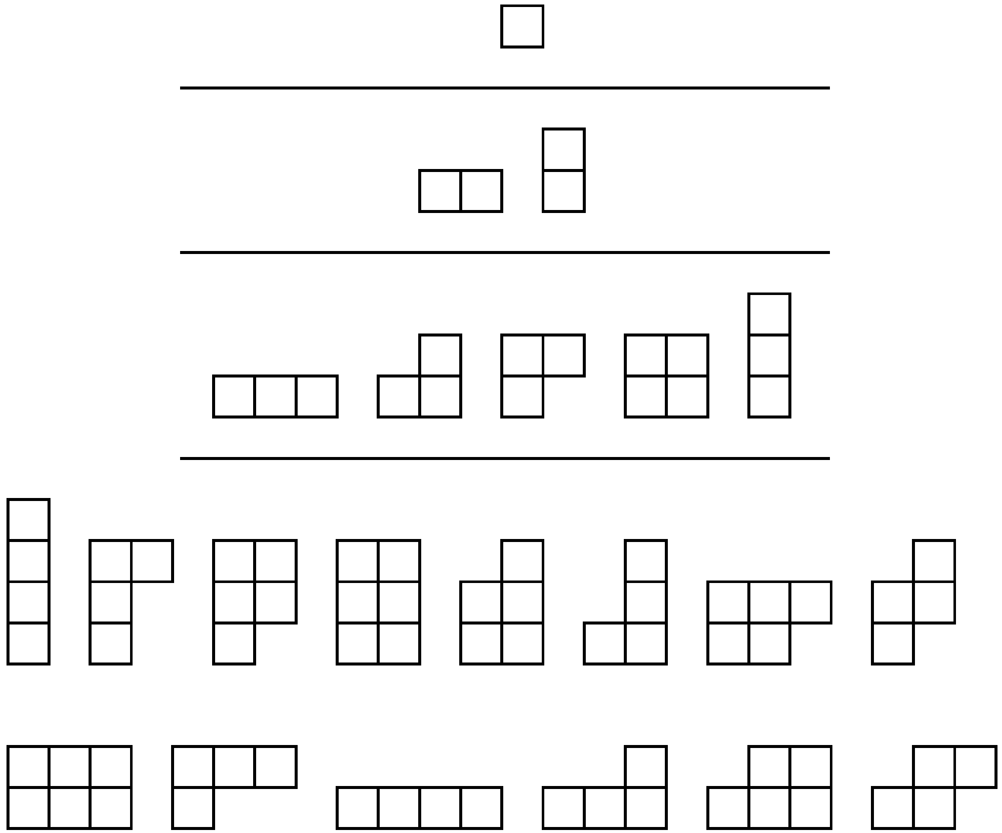


Another one: Skew Polyominos

“Determine the number of skew polyominos with perimeter $2n + 2$.”

Polyomino: figure composed by squares connected by their edges.

Skew polyomino: successive columns of squares from left to right increase in height: the bottom of the column to the left is always lower or equal to the bottom of the column to the right. Similarly, the top of the column to the left is always lower than or equal to the top of the column to the right.



Derivations of the Catalan numbers

Catalan numbers derived with generating functions

Any string containing $n > 0$ pairs of parentheses can be decomposed as:

$$(A)B$$

where, if A contains k pairs of parentheses, B must contain $n - k - 1$.

All configurations of n parenthesis pairs are the ones where A is empty and B contains $n - 1$ pairs, plus the ones where A contains 1 pair and B contains $n - 2$, and so on:

$$\begin{aligned}
C_1 &= C_0C_0 \\
C_2 &= C_0C_1 + C_1C_0 \\
C_3 &= C_0C_2 + C_1C_1 + C_2C_0 \\
C_4 &= C_0C_3 + C_1C_2 + C_2C_1 + C_3C_0 \\
... &= ...
\end{aligned}$$

which can be rewritten in the form of a recurrence relation:

$$C_0 = 1, \quad C_1 = 1, \quad C_n = \sum_{i=0}^{n-1} C_i C_{n-1-i}$$

We will now solve the above recurrences with the use of generating functions.

$$C(x) = C_0 + C_1 \cdot x + C_2 \cdot x^2 + ... = \sum_{i=0}^{+\infty} C_i \cdot x^i$$

Let's now examine the expression of $[C(x)]^2 = C(x)C(x)$, as follows:

$$\begin{array}{ccccccc} C_0C_0 & + & (C_0C_1 + C_1C_0) & x & +(C_0C_2 + C_1C_1 + C_2C_0) & x^2 + ... = \\ || & & || & & || & \\ C_1 & + & C_2 & x & + \quad C_3 & x^2 + ... \end{array}$$

still a generating function with Catalan coefficients, shifted one position left:

$$[C(x)]^2 = C_1 + C_2x + C_3x^2 = \sum_{i=0}^{+\infty} C_{i+1}x^i \quad .$$

Therefore if we multiply the whole series by x and add C_0, the original Catalan series is obtained:

$$C(x) = C_0 + x[C(x)]^2.$$

A quadratic equation, which could be put into the more familiar form:

$$xC^2 - C + C_0 = 0,$$

where C is the unknown and x, C_0 are constant coefficients. Replacing C_0 with its value (i.e., 1), the solution is trivially given by:

$$C = \frac{1 \pm \sqrt{1 - 4x}}{2x}.$$

Only the $-$ solution is acceptable, being $C_0 = 1$:

$$C = \frac{1 - \sqrt{1 - 4x}}{2x}. \tag{1}$$

The solution contains the power of a binomial with fractional exponent:

$$\sqrt{1 - 4x} = (1 - 4x)^{1/2} = \sum_{n \geq 0} \binom{1/2}{n} (-4x)^n,$$

which can be expanded as:

$$\begin{aligned}(1-4x)^{1/2} &= 1-\frac{1/2}{1}4x+\frac{(1/2)(-1/2)}{2\cdot 1}(4x)^2+\\ &+ \frac{(1/2)(-1/2)(-3/2)}{3\cdot 2\cdot 1}(4x)^3+\\ &+ \frac{(1/2)(-1/2)(-3/2)(-5/2)}{4\cdot 3\cdot 2\cdot 1}(4x)^4+...\end{aligned}$$

which can be simplified as follows:

$$(1-4x)^{1/2}=1-\frac{1}{1!}2x+\frac{1}{2!}4x^2-\frac{3\cdot 1}{3!}8x^3+\frac{5\cdot 3\cdot 1}{4!}16x^4+...$$

Now, substituting we obtain:

$$C(x) = 1 - \frac{1}{2!}2x + \frac{3 \cdot 1}{3!}4x^2 + \frac{5 \cdot 3 \cdot 1}{4!}8x^3 + \frac{7 \cdot 5 \cdot 3 \cdot 1}{5!}16x^4 + ...$$

We can get rid of terms like $7 \cdot 5 \cdot 3 \cdot 1$ (factorials missing the even factors), by considering that:

$$\begin{aligned}
2^2 \cdot 2! &= 4 \cdot 2 \\
2^3 \cdot 3! &= 6 \cdot 4 \cdot 2 \\
2^4 \cdot 4! &= 8 \cdot 6 \cdot 4 \cdot 2 \\
... &= ... \\
2^n \cdot n! &= \prod_{i=1}^{n} 2i
\end{aligned}$$

Consequently:

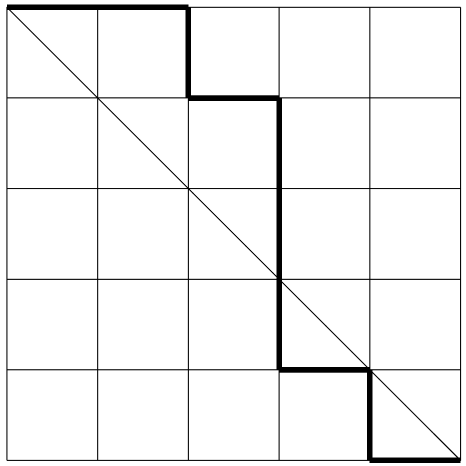
$$C(x) = 1 + \frac{1}{2}(\frac{2!}{1!1!})x + \frac{1}{3}(\frac{4!}{2!2!})x^2 + \frac{1}{4}(\frac{6!}{3!3!})x^2 = $$
$$= \sum_{i=0}^{+\infty} \frac{1}{1+i}\binom{2i}{i} z^i$$

Therefore, the i^{th} Catalan number is:

$$C_i = \frac{1}{1+i}\binom{2i}{i}.$$

The simplest proof

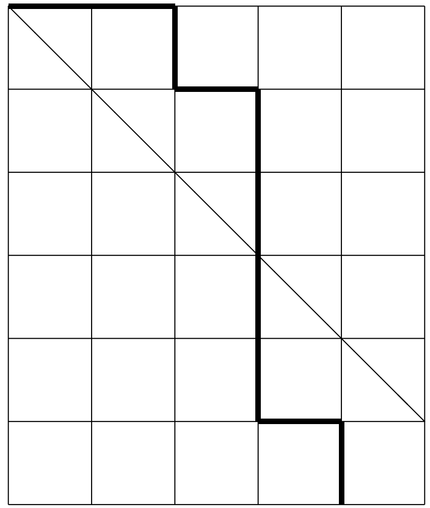
Solution based on considerations on the count of diagonal-avoiding paths on a lattice: determining C_n equals counting the total paths through the grid and subtracting the number of invalid ones.



(example of an invalid path).

$P(i, i+1)$: first illegal reached point.

Transformation: from point P on, replace S $\leftrightarrow$ W segments.



Transformation starts at point $(i, i+1)$ and it causes the $n-i$ W segments to be replaced by S segments and the remaining $n-i-1$ S segments to be replaced by W segments, $\Rightarrow$ new ending coordinates $(i+(n-i-1),(i+1)+(n-i)) = (n-1, n+1)$.

By construction, every illegal path in a $n \times n$ lattice corresponds to exactly one non-constrained path in $(n-1) \times (n+1)$ lattice.

Paths in a $a \times b$ lattice: $\binom{a+b}{a} \Rightarrow$ in a $n \times n$: $\binom{2n}{n}$

Invalid paths in a $n \times n$ lattice $=$ paths in a $(n-1) \times (n+1)$ lattice: $\binom{2n}{n+1}$

$$C_n = \binom{2n}{n} - \binom{2n}{n+1} = \binom{2n}{n} - \frac{n}{n+1}\binom{2n}{n} = \frac{1}{n+1}\binom{2n}{n}.$$

A novel interpretation

A novel calculation of the Catalan numbers, inspired by formal language considerations.

Language of balanced parentheses (Van Dyck language).

$$
\begin{array}{rcl}
G & = & (\Sigma, N, S, R) \\
\Sigma & = & \{(,)\} \\
N & = & \{S\} \\
R & = & \{r_1, r_2\} \\
r_1 & : & S \rightarrow \varepsilon \\
r_2 & : & S \rightarrow (S)S
\end{array}
$$

Sentential form: sequence of terminal and nonterminal symbols which can be derived from the start symbol S. Strings are special sentential forms of terminal symbols only.

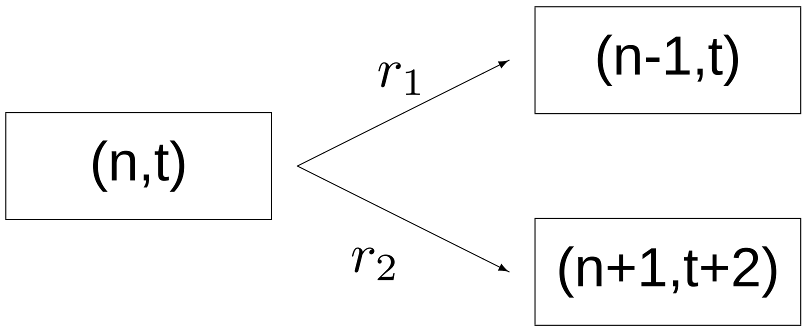
Labelling: (n, t)-label for a sentential form containing n nonterminals and t terminals. Strings $\in \mathcal{L}(G)$ will be labeled $(0, 2i), i \in \mathbb{N}$.

Theorem: the number of terminal symbols is even. Proof: the axiom contains no terminals, rules preserve parity.

Derivation step: a substitution replacing a single S symbol in a sentential form with the right-hand side of rule r_1 or r_2.

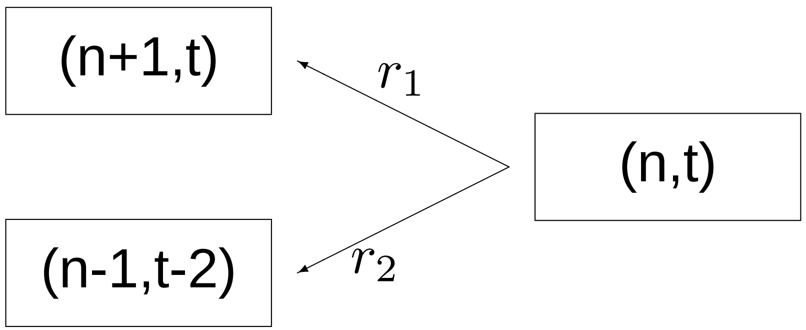
The derived form of a (n, t) form will have:

- one nonterminal less, same number of terminals (rule r_1 applied);
- one more non-terminal and two more terminals (rule r_2 applied).



Thus, a given sentential form derives either from:

- a $(n+1, t)$ form, through rule r_1; or
- a $(n-1, t-2)$ form, through rule r_2.



Through recursive application of above scheme, all predecessors of a given sentential form can be determined, up to the axiom, which has obviously label $(1, 0)$.

Theorem: axiom can only have label $(1, 0)$.

Theorem: label $(1, 0)$ corresponds to axiom only.

In general it is not true that each sentential form has exactly 2 predecessors:

- a $(1, 0)$ form has no predecessor by definition, being the axiom;
- $(0, t)$ and $(1, t)$ forms can only have a $(n + 1, t)$ predecessor. (Proof:

by contradiction, the $(n-1, t-2)$ predecessor would have zero or less nonterminals, therefore it could have no successors.)

- (n, t) forms with $n > t$ do not exist, apart from the axiom $(1, 0)$. (Proof: by induction. For each form (n, t), be $\delta = t - n$. The axiom has $\delta = -1$, both rules increment δ.)

Each derivation tree starts with a label corresponding to a string, $(0, 2i)$, and reaches leaf nodes which are either axioms or invalid nodes.

Number of axioms contained in the tree of a $(0, 2i)$-string

$$=$$

number of different ways in the derives a $(0, 2i)$-string

$$=$$

number of different strings of balanced parentheses of length $2i$ (since each derivation is unique).

Examples follow; axioms are marked with “!”, invalid nodes with “$\times$”.

(0,2)
|
(1,2)
|
(2,2)
|
(1,0) !

Figure 1: Derivation tree for $(0, 2)$, i.e., $2n = 2$

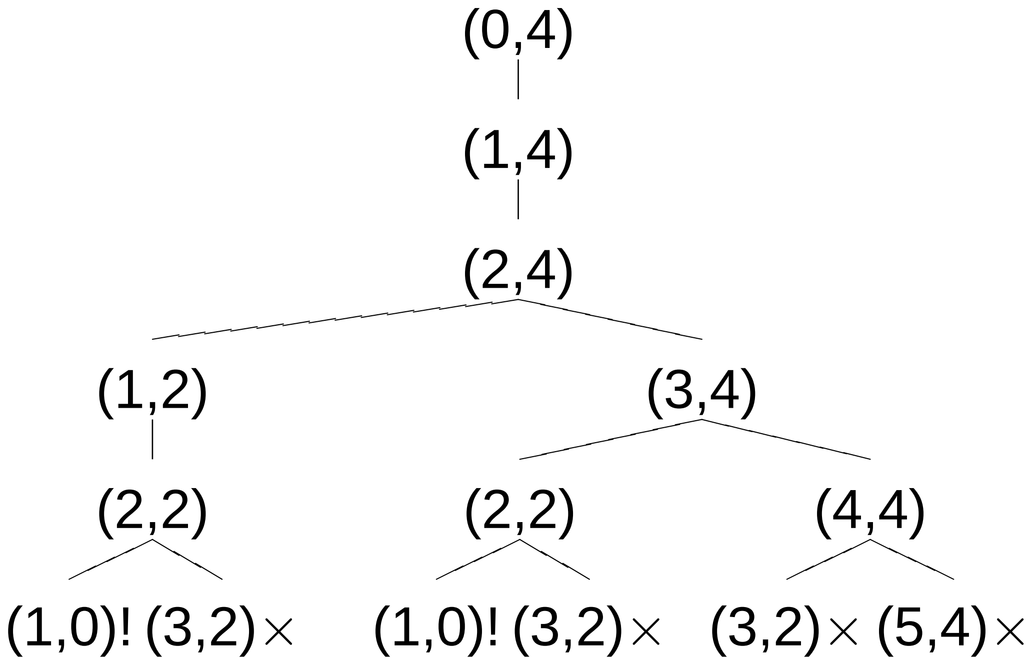


Figure 2: Derivation tree for $(0,4)$, i.e., $2n = 4$

The count of axiom nodes in each derivation tree, is the i^{th} Catalan number.

According to the above rules, the C_i is the number of axioms in the derivation tree of a $(0, t)$ form, with $t = 2i$, thus:

$$R(n,t) = \begin{cases} 1 & \text{if } (n = 1 \wedge t = 0) \\ 0 & \text{if } n > t \\ R(n+1,t) & \text{if } n <= 1 \\ R(n-1,t-2) + R(n+1,t) & \text{otherwise.} \end{cases}$$

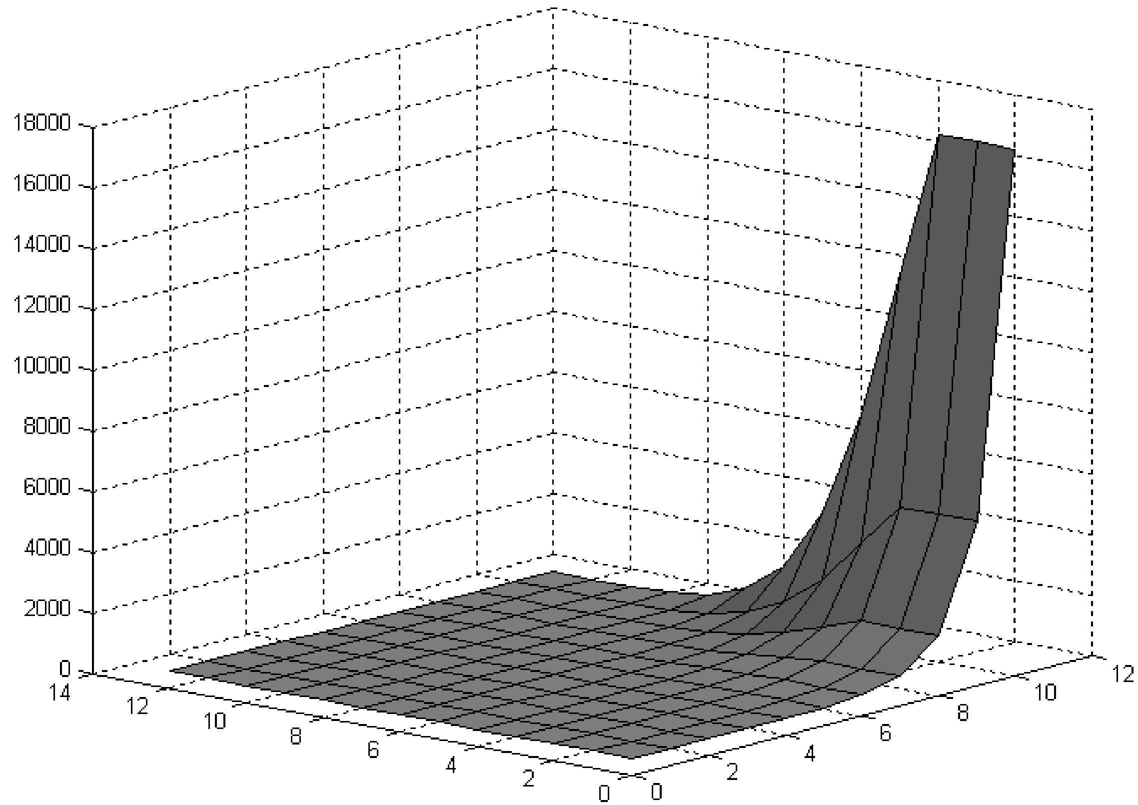
By construction,

$$C_i = R(0, 2i).$$

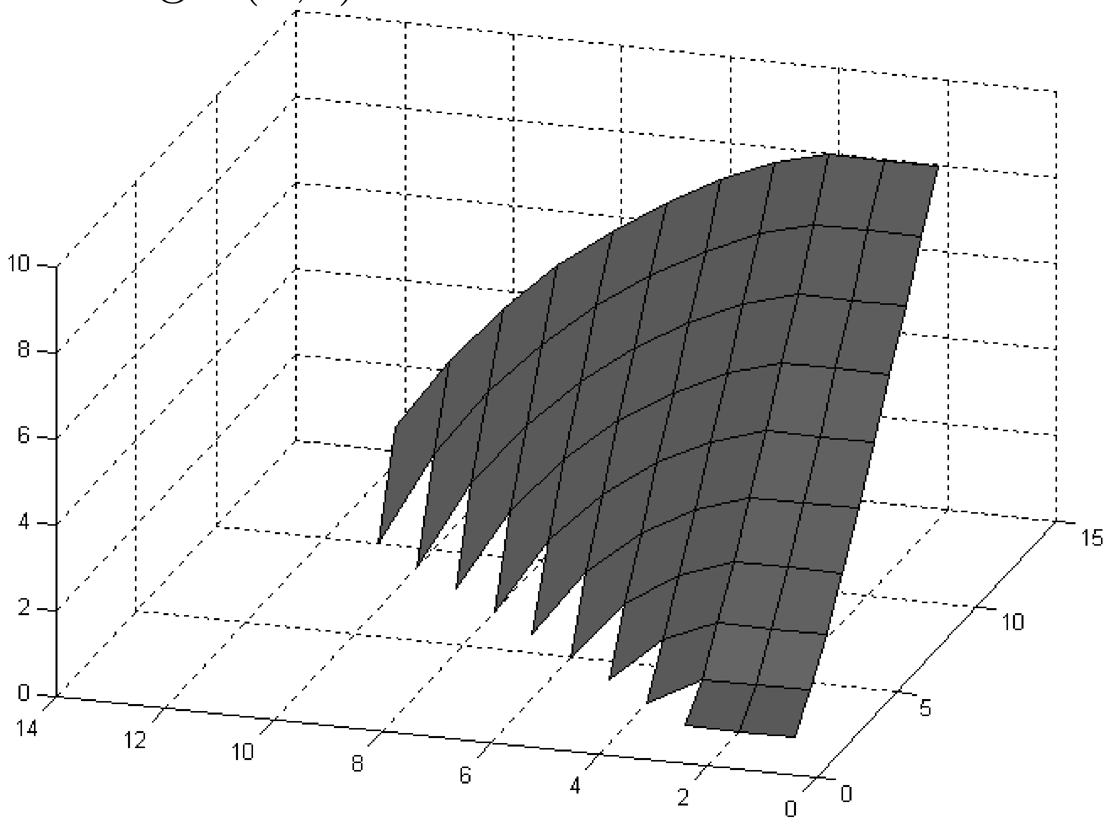
The above relation R was implemented in Tcl function and tested for correctness.

t/n	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	2	1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	5	5	5	3	1	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	14	14	14	9	4	1	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	42	42	42	28	14	5	1	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	132	132	132	90	48	20	6	1	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	429	429	429	297	165	75	27	7	1	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	1430	1430	1430	1001	572	275	110	35	8	1	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	4862	4862	4862	3432	2002	1001	429	154	44	9	1	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	16796	16796	16796	11934	7072	3640	1638	637	208	54	10	1	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	58786	58786	58786	41990	25194	13260	6188	2548	910	273	65	11	1	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	208012	208012	208012	149226	90440	48450	23256	9996	3808	1260	350	77	12	1

Plot of the surface $R(n,t)$. Points with odd values of t were not plotted.



Plot of the surface $\log R(n,t)$. Points with odd values of t were not plotted.



Future Developments

If D is the language defined by: $S \rightarrow \varepsilon \mid (S)S$, then D can be recursively written as: $D = \{\varepsilon\} + (D)D$ where $+$ denotes disjoint set union.

If an alphabet $A = \{a_1, a_2, ..., a_m\}$ is considered, A^* denotes the language of all the strings over alphabet A, and if $\alpha \in A^*$ is a string over A, $|\alpha|$ denotes the length of α.

For a language A^*, a function $w : A^* \rightarrow \mathbb{Z}[\![x]\!]$ can be defined:

$$w(\alpha) = x^{|\alpha|}, \text{with } \alpha \in A^*$$

and we set by convention that $w(\varepsilon) = 1$.

It is trivial to prove that function w exhibits the following property:

$$\forall \alpha, \beta \in A^*, w(\alpha \cdot \beta) = w(\alpha) w(\beta),$$

where $\cdot$ denotes string concatenation. w can be extended on languages, by defining $w(L) = \sum_{\alpha \in L} w(\alpha)$. Therefore

$$w(A^*) = \sum_{\alpha \in A*} w(\alpha) = \sum_{\alpha \in A*} x^{|\alpha|} = \sum_{n \geq 0} \left(\sum_{|\alpha| = n} 1 \right) x^n = \sum_{n \geq 0} m^n x^n = \frac{1}{1 - mx}.$$

The following equation can be set for language D:

$$w(D) = 1 + x^2 w(D)^2$$

which can be solved by replacing $y = w(D)$, thus obtaining $y = 1 + x^2y^2$, or in a more familiar form: $x^2y^2 - y + 1 = 0$. The solution is given by:

$$y = \frac{1 - \sqrt{1 - 4x^2}}{2x^2}$$

It can be shown that:

$$D(x) = w(D) = D_0 + D_2x^2 + D_4x^4 + D_6x^6 + ...$$

where $\forall i \in \mathbb{N}, D_{2i} = C_i$ and $D_{2i+1} = 0$, therefore

$$D(x) = w(D) = C_0 + C_1x^2 + C_2x^4 + C_3x^6 + ...$$

Our aim is to extend above considerations to the language E of *sentential forms* of grammar G.

Each string in D corresponds to exactly 1 mountain range; each string in E corresponds to exactly one *extended* mountain range, with also horizontal strokes. Example: $(S)(S)((S)(S(S)))S$, and its corresponding extended mountain range:

```
             _
        _  _/ \
 _  _  / \/    \
/ \/ \/         \_
```

The language E of sentential forms of grammar G is a new language, can be defined by the a new grammar H, given as follows:

$$
\begin{array}{rcl}
H & = & (\Sigma', N, B, R') \\
\Sigma' & = & \{(,), \mathtt{S}\} \\
N & = & \{B\} \\
R & = & \{r_1, r_2, r_3\} \\
r_1 & : & B \longrightarrow \varepsilon \\
r_2 & : & B \longrightarrow \mathtt{S}B \\
r_3 & : & B \longrightarrow (B)B
\end{array}
$$

Note: S is a terminal symbol for grammar H. The above language is called the *Motzkin* language.

Now let us consider the production:

$$B \to \varepsilon \mid SB \mid (B)B,$$

and we give now a recursive definition of E:

$$E = \{\varepsilon\} \ + \ SE \ + \ (E)E.$$

It is then time to introduce a newer, more useful definition of $w(\alpha)$:

$$w(\alpha) = x^{p(\alpha)} y^{o(\alpha)} z^{|\alpha|}$$

where $p(\alpha) = |\alpha|_(+ |\alpha|_)$, and $o(\alpha) = |\alpha|_S$, therefore $p(\alpha) + o(\alpha) = |\alpha|$.

From the recursive definition of E, it is possible to set:

$$w(E) = 1 + yzw(E) + x^2 z^2 w(E)^2,$$

which, replacing $e = w(E)$, is:

$$x^2 z^2 e^2 + (yz - 1)e + 1 = 0,$$

which, solved by e yields:

$$e = \frac{1 - yz - \sqrt{1 - 2yz + y^2 z^2 - 4x^2 z^2}}{2x^2 z^2}.$$

Thus $e(x, y, z)$ can be written as a formal power series with coefficients

$E_{i,j,k}$:

$$
\begin{aligned}
e(x,y,z) &= E_{0,0,0}+\\
&+ E_{1,0,0}x+E_{0,1,0}y+E_{0,0,1}z+\\
&+ E_{2,0,0}x^2+E_{0,2,0}y^2+E_{0,0,2}z^2+E_{1,1,0}xy+E_{0,1,1}yz+E_{1,0,1}xz+\\
&+ \ldots
\end{aligned}
$$

It is now evident that the number of sentential forms with n nonterminals and t terminals, previously called $R(n,t)$ is given by :

$$R(n,t)=E_{n,t,n+t}=[x^ty^n]e(x,y,1),$$

where the notation $[...]$ has the following meaning:

$$[x^n]f(x) = f_n \Leftrightarrow f(x) = \sum_{n \geq 0} f_n x^n,$$

in particular

$$[x^i y^j z^k]e(x, y, z) = E_{i,j,k}.$$

Furthermore, an expression of $e(x, y, 1)$ can be obtained by restriction:

$$e(x, y, 1) = e(x, y, z)|_{z=1} = \frac{1 - y - \sqrt{1 - 2y + y^2 - 4x^2}}{2x^2}.$$

Incidentally, the i-th Catalan number, which was equal to $R(0, 2i)$ can be obtained by setting $y = 0$, thus:

$$e(x,0,1) = e(x,y,1)|_{y=0} = \frac{1-\sqrt{1-4x^2}}{2x^2},$$

which is identical to a previous equation and admits the same solutions.

To obtain an expression of $R(i,j)$, we can collect $(1-y)$ in the numerator and $(1-y)^2$ in the denominator, thus obtaining:

$$e = \frac{1-y}{(1-y)^2} \frac{1-\sqrt{1-4\frac{x^2}{(1-y)^2}}}{\frac{2x^2}{(1-y)^2}} = \frac{1}{1-y} \left. \frac{1-\sqrt{1-4q^2}}{2q^2} \right|_{q=\frac{x}{1-y}},$$

which can be solved by comparison with previous case, thus:

$$e = \frac{1}{1-y} D\left(\frac{x}{1-y}\right)$$

but since

$$D(q) = \frac{1 - \sqrt{1 - 4q^2}}{2q^2} = \sum_{k \geq 0} D_k q^k$$

then

$$e = \frac{1}{1-y} \sum_{k \geq 0} D_k \frac{x^k}{(1-y)^k} = \sum_{k \geq 0} D_k \frac{x^k}{(1-y)^{k+1}}$$

$$= \sum_{k \geq 0} D_k x^k \sum_{n \geq 0} \binom{n+k}{k} y^n$$

$$= \sum_{n,k \geq 0} \binom{n+k}{k} D_k x^k y^n$$

therefore it should be true that:

$$R(n,k) = \begin{cases} 0 & \text{if } k \text{ odd} \\ \binom{n+k}{k} C_{k/2} & \text{if } k \text{ even} \end{cases}$$

Problem: the above equation is incorrect!

Cause: the real language of sentential forms of E is smaller than $\mathcal{L}(H)$;

Proof: $S, SS, SSS, ...S(), S(S), ... \in \mathcal{L}(H) - E$.

Thus: above work need to be remade.

Our current efforts are devoted to finding a new correct and unambiguous grammar for language E, obtaining an appropriate recursive definition of E and a corresponding equation for $w(E)$, which, solved, would yield a formal power series for $w(E)$, thus, a closed form for $R(n, t)$ numbers.