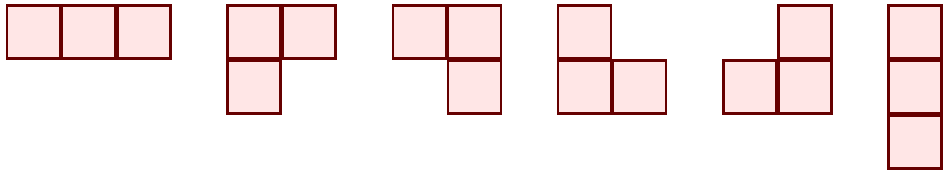
Problems That Philippe Would Have Loved

dedicated to
Philippe Patrick Michel Flajolet
(1948–2011)

Don Knuth — AofA 2014 — Paris

Problem 1:

Fixed Polyominoes



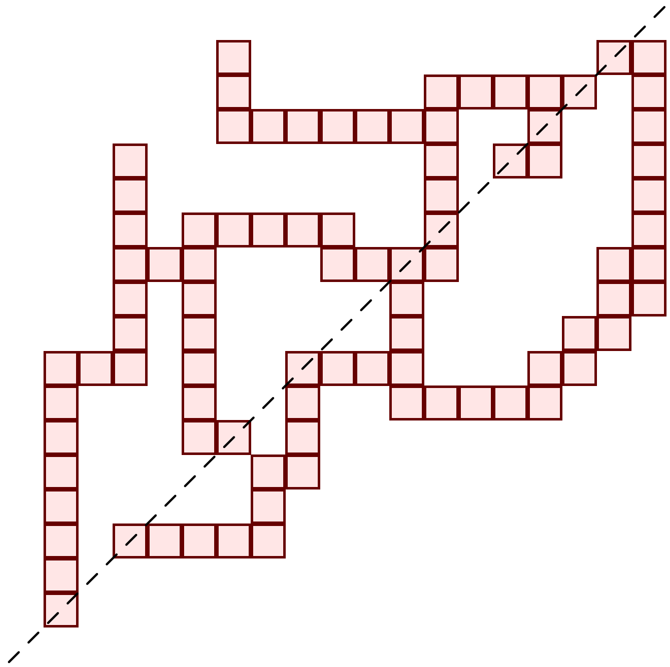
Rookwise connected patterns of n square cells:

1, 2, 6, 19, 63, 216, 760, 2725, 9910, 36446, 135268, 505861,
$\ldots,$
27268084442494384061453863 4, $(n = 47)$
$\ldots,$
6915071456253289693657442548 0218, $(n = 56)$
$\ldots,$
$t(n), \ldots$

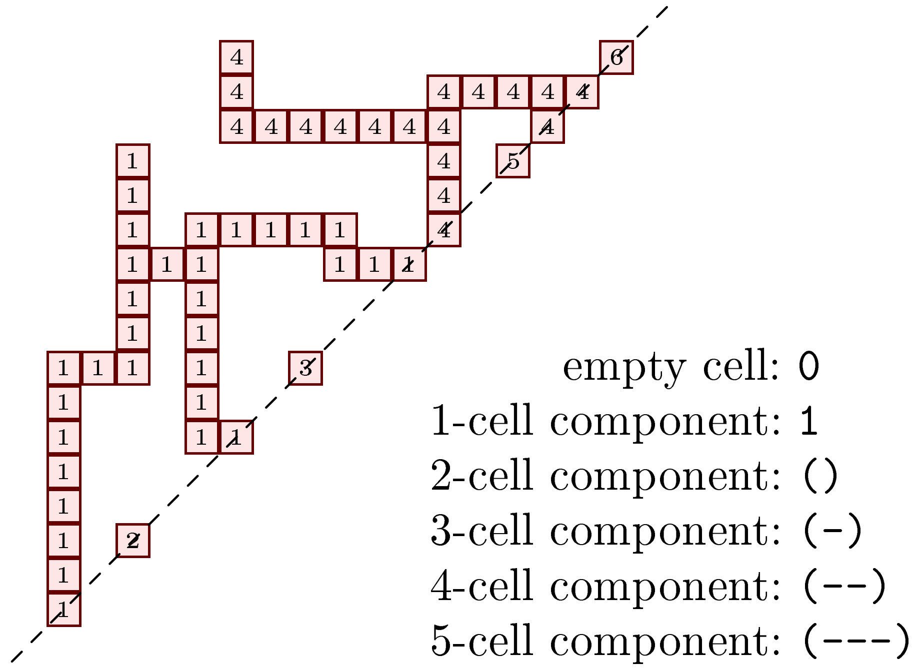
Reason to Hope:

$$t(n) = \theta^n \Big(\frac{.3169}{n} - \frac{.275}{n^2} + \frac{.33}{n^{5/2}} - \frac{.22}{n^3} + \cdots \Big)$$

$\theta \approx 4.062570$ [Klarner's constant]



diagonal slice pattern `(0100-0100)(01-)1`



diagonal slice pattern `(0100-0100)(01-)1`

(the parentheses in a pattern must be properly nested)

$$S \to X \mid SZX,\ X \to \mathtt{1} \mid \mathtt{(}I\mathtt{)},\ I \to \epsilon \mid \mathtt{0}I \mid XI \mid \mathtt{-}I,\ Z \to \epsilon \mid \mathtt{0}Z$$

$t_\alpha(n)$ = number of partial polyominoes built on slice α,
with n cells on and above the diagonal

$g_\alpha = \sum_n t_\alpha(n) z^n$

initial slice = all cells disconnected: `1`, `11`, `101`, `111`, `1001`, ...

final slice = all cells connected: `1`, `()`, `(0)`, `(-)`, `(00)`, ...

$$g = \sum_n t(n) z^n = \sum_{\alpha\,\text{final}} g_\alpha.$$

Infinite linear system $g_\alpha = z^{m(\alpha)}\bigl(\delta(\alpha) + \sum_\beta c(\alpha,\beta)\,g_\beta\bigr)$

$$\begin{aligned} m(\alpha) &= \text{number of (nonzero) cells in slice } \alpha;\\ \delta(\alpha) &= 1 \text{ if } \alpha \text{ is initial, otherwise } 0;\\ c(\alpha,\beta) &= \text{number of ways } \alpha \text{ can follow } \beta. \end{aligned}$$

$$\begin{aligned} g_{\texttt{1}} &= z(1 + 2g_{\texttt{1}} + g_{\texttt{11}} + 3g_{\texttt{()}} + g_{\texttt{101}} + 3g_{\texttt{(0)}} + \cdots)\\ g_{\texttt{11}} &= z^2(1 + 2g_{\texttt{1}} + 2g_{\texttt{()}} + g_{\texttt{101}} + 2g_{\texttt{(0)}} + 2g_{\texttt{(-)}} + \cdots)\\ g_{\texttt{()}} &= z^2(g_{\texttt{1}} + 2g_{\texttt{11}} + 2g_{\texttt{()}} + 3g_{\texttt{(0)}} + g_{\texttt{111}} + 3g_{\texttt{(-)}} + \cdots)\\ g_{\texttt{101}} &= z^2(1 + 4g_{\texttt{1}} + 3g_{\texttt{11}} + 2g_{\texttt{101}} + 4g_{\texttt{(0)}} + 2g_{\texttt{111}} + \cdots)\\ g_{\texttt{(0)}} &= z^2(g_{\texttt{()}} + 2g_{\texttt{(0)}} + 2g_{\texttt{(-)}} + 2g_{\texttt{1()}} + 2g_{\texttt{()1}} + \cdots) \end{aligned}$$

Conjecture. *All g_α have the same singularities.*

$$t_\alpha(n) \sim c_\alpha \theta^n / n.$$

Empirically, most of the contributions to $t(n)$ come from g_1:

$$\begin{aligned}
t_{\mathtt{1}}(15) &= 20112826 \quad (73.4\%)\\
t_{\mathtt{()}}(15) &= 4322819 \quad (15.8\%)\\
t_{\mathtt{(0)}}(15) &= 1085779 \quad (4.0\%)\\
t_{\mathtt{(-)}}(15) &= 608461 \quad (2.2\%)\\
t_{\mathtt{(00)}}(15) &= 393425 \quad (1.4\%)\\
t_{\mathtt{(0-)}}(15) = t_{\mathtt{(-0)}}(15) &= 163613 \quad (0.6\%)\\
t_{\mathtt{(000)}}(15) &= 139624 \quad (0.5\%)\\
t_{\mathtt{(---)}}(15) &= 84949 \quad (0.3\%)\\
t_{\mathtt{(-00)}}(15) = t_{\mathtt{(00-)}}(15) &= 51788 \quad (0.2\%)
\end{aligned}$$

and we also have $t_{\mathtt{1}}(14) = 5301097$ (73.6%),
$t_{\mathtt{()}}(14) = 1146353$ (15.9%), $t_{\mathtt{(0)}}(14) = 285841$ (4.0%), etc.

What slices α can follow a given slice β?

```
          (  0  0  -  (  )  0  -  1  )  0  0  0  1
```
$$\ldots x_{-1} x_0\ x_1\ x_2\ x_3\ x_4\ x_5\ x_6\ x_7\ x_8\ x_9\ x_{10} x_{11} x_{12} x_{13} x_{14} \ldots$$

Answer: One α for every solution to the Boolean formula

$$(x_0 \vee x_1 \vee x_3 \vee x_4 \vee x_7 \vee x_8 \vee x_9 \vee x_{10}) \\ \wedge (x_4 \vee x_5 \vee x_6) \wedge (x_8 \vee x_9) \wedge (x_{13} \vee x_{14})$$

in the infinite set of variables x_k for $-\infty < k < \infty$ (one constraint for every component of β), because every component of β must survive in α.

Although $g_{\texttt{10000001}} = z^2 + 4z^3 + \cdots$, the actual contribution of $g_{\texttt{10000001}}$ to g is only $132z^{15} + O(z^{16})$, because at least 13 more cells are needed to connect the slice 10000001 with other cells below the diagonal (and there are 132 ways to do that with 13).

It's easy to compute the minimum "cost" k that's needed to connect a slice α below the diagonal. Slices of cost k essentially lose a factor of θ^k before they contribute to g, although that's somewhat compensated (for large k) by a large number of ways to connect them up. (132 is a Catalan number.)

Here are the 47-ominoes enumerated by (height, width), rounded to units of 10^{22}:

	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
6											1	2	3	5	6	6	6	4	3	1	1
7									2	7	15	25	33	36	33	26	18	11	5	3	1
8							2	10	28	57	87	107	109	95	71	47	27	14	6	3	1
9						5	23	66	129	191	226	224	190	141	92	53	27	13	5	2	1
10					6	35	105	208	307	362	356	301	223	147	86	46	22	9	4	1	
11				5	35	121	259	396	475	474	407	306	205	123	67	33	15	6	2	1	
12			2	23	105	259	428	539	555	488	376	259	160	90	46	21	9	3	1		
13			10	66	208	396	539	583	531	423	299	191	111	59	28	13	5	2	1		
14		2	28	129	307	475	555	531	439	321	212	127	70	35	16	7	3	1			
15		7	57	191	362	474	488	423	321	219	136	77	40	19	8	3	1				
16	1	15	87	226	356	407	376	299	212	136	80	43	21	10	4	2	1				
17	2	25	107	224	301	306	259	191	127	77	43	22	10	4	2	1					
18	3	33	109	190	223	205	160	111	70	40	21	10	5	2	1						
19	5	36	95	141	147	123	90	59	35	19	10	4	2	1							
20	6	33	71	92	86	67	46	28	16	8	4	2	1								
21	6	26	47	53	46	33	21	13	7	3	2	1									
22	6	18	27	27	22	15	9	5	3	1	1										
23	4	11	14	13	9	6	3	2	1												
24	3	5	6	5	4	2	1	1													
25	1	3	3	2	1	1															
26	1	1	1	1																	

References

Iwan Jensen and Anthony J. Guttmann, "Statistics of lattice animals (polyominoes) and polygons," *Journal of Physics* **A33** (2000), L257–L262.

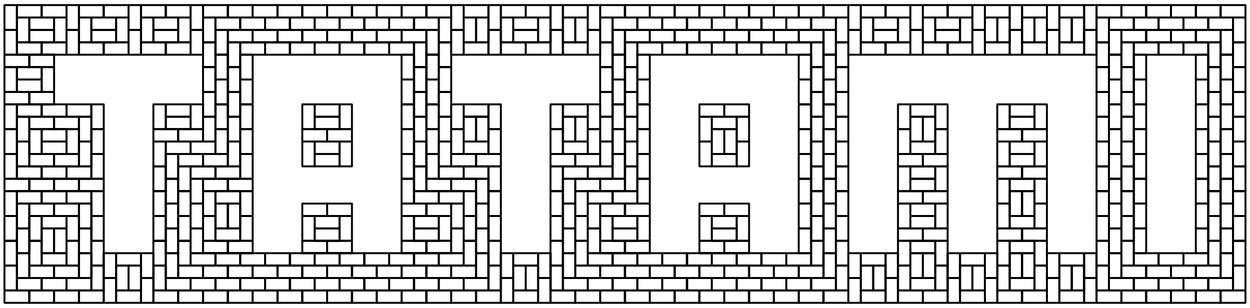
Iwan Jensen, "Enumeration of lattice animals and trees," *Journal of Statistical Physics* **102** (2001), 865–881.

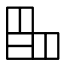
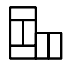
Anthony J. Guttmann (editor), *Polygons, Polyominoes, and Polycubes: Lecture Notes in Physics* **775** (2009), 490 pages.

On-Line Encyclopedia of Integer Sequences, sequence A001168 [`http://oeis.org/A001168`].

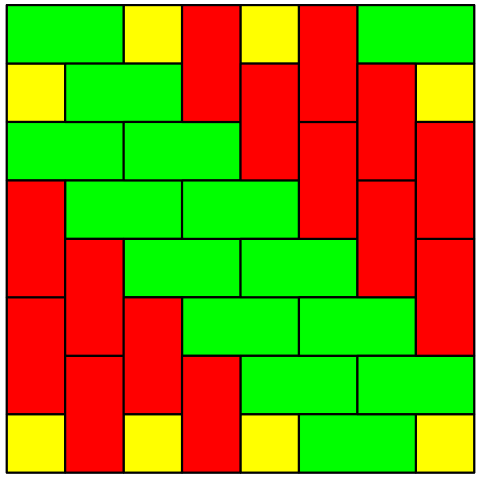
Problem 2 (A. Erickson, F. Ruskey):

Tatami Polynomials

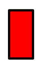


(Notice that  is disallowed, but  would be fine.)

$n \times n$ tilings with n monominoes, $\binom{n}{2}$ dominoes:



There are $n2^{n-1}$ solutions; the generating function for $n = 8$ is

$g_8($  $) \;=\; \sum_{k=0}^{28}$  ^{8}  ^{k}  $^{28-k}$;

$$g_8(z) \;=\; 2 + 4z + 6z^2 + \cdots + 6z^{26} + 4z^{27} + 2z^{28}.$$

Let $S_n(z) = (1+z)(1+z^2)\ldots(1+z^n)$.

Theorem (Erickson, Ruskey, 2013)**:**

If $n \geq 2$ *we have* $g_n(z) = 2h_n(z) + 2z^{n(n-1)/2}h_n(z^{-1})$, *where*

$$\begin{aligned} h_n(z) &= 2\sum_{k=1}^{\lfloor (n-1)/2 \rfloor} S_{n-k-2}(z)S_{k-1}(z)z^{n-k-1} + S_{\lfloor (n-2)/2 \rfloor}(z)^2 \\ &= P_n(z)\prod_{j=1}^{\infty} S_{\lfloor (n-2)/2^j \rfloor}(z), \end{aligned}$$

and $P_n(z)$ *is a (mysterious) polynomial.*

For example, $P_3(z) = 1 + 2z$, $P_4(z) = 1 + z + 2z^2$,
$P_5(z) = 1 + z + 2z^2 + 4z^3 + 2z^5$, $\ldots$.

If $n \geq 2$, $\deg(P_n(z)) = \sum_{k=1}^{n-2}$ (number of odd divisors of k).

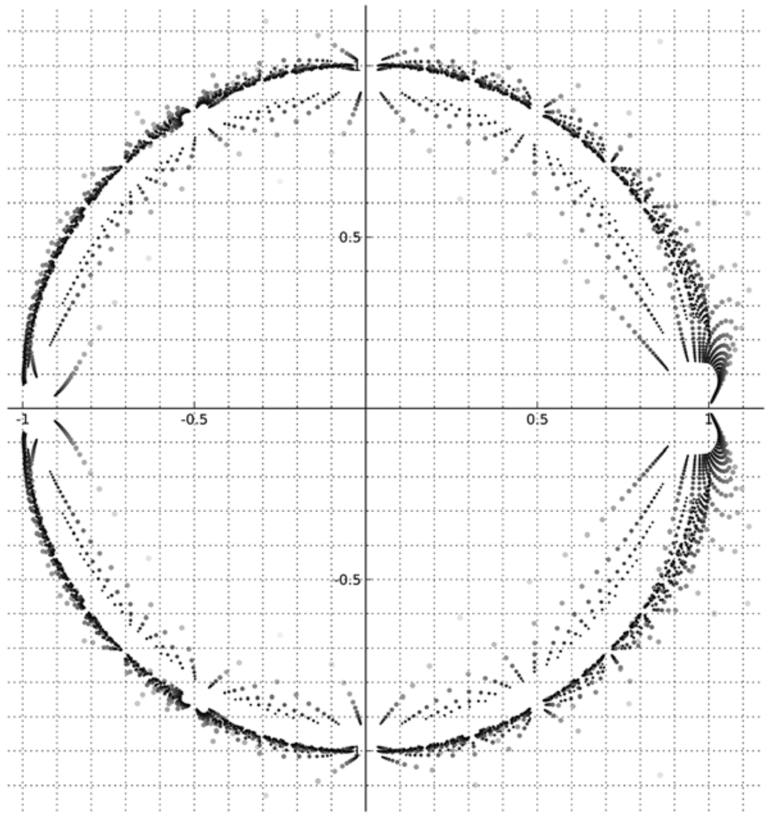
$P_n(1) = n\,2^{\nu(n-2)-1}$, where νk denotes the sideways sum of k.

Conjectures:

- $P_n(z)$ is irreducible over the integers.
- $$\sum_{n \geq 2} P_n(-1) z^{n-2} = \frac{1+z}{1-2z^2}\sqrt{\frac{1-2z}{1+2z}}.$$

(Both have been verified for $n < 200$. The degree of $P_{199}(z)$ is 13022, and its largest coefficient has 55 decimal digits.)

Here are the roots of $P_n(z)$, when n is even and less than 50:



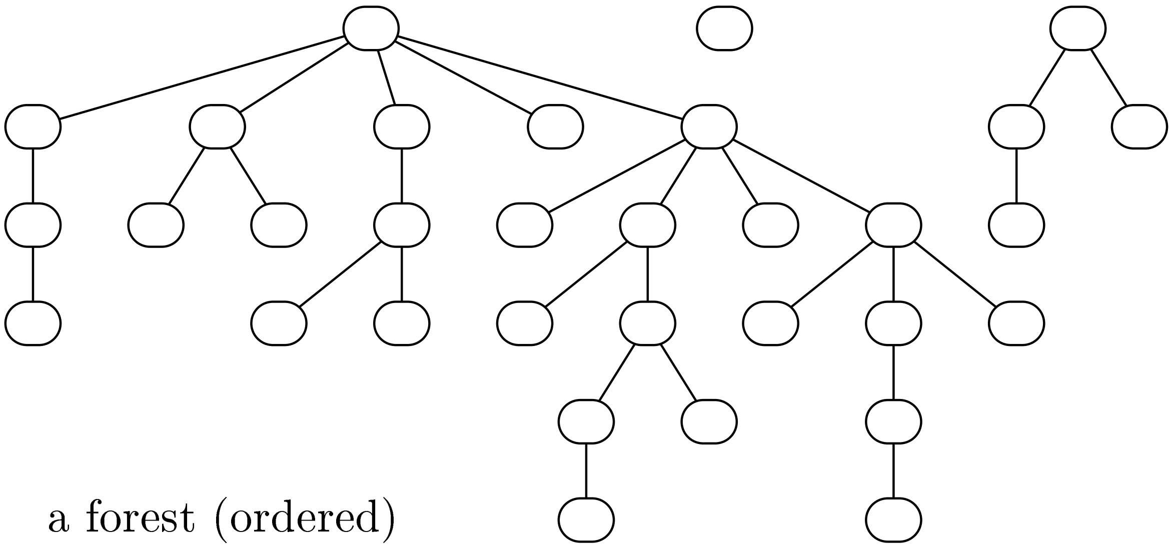
References

Alejandro Erickson, Frank Ruskey, Mark Schurch, and Jennifer Woodcock, "Monomer-dimer tatami tilings of rectangular regions," *The Electronic Journal of Combinatorics* **18** (2011), #P109, 24 pages.

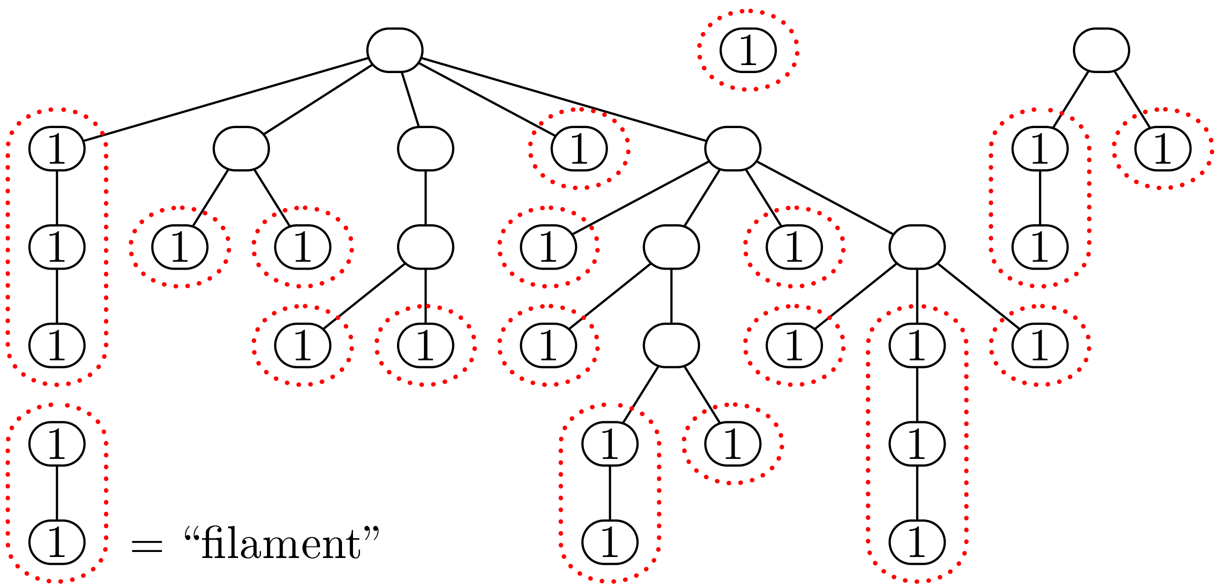
Alejandro Erickson and Frank Ruskey, "Enumerating maximal tatami mat coverings of square grids with v vertical dominoes," `arXiv:1304.0070` [math.CO] (2013), 23 pages.

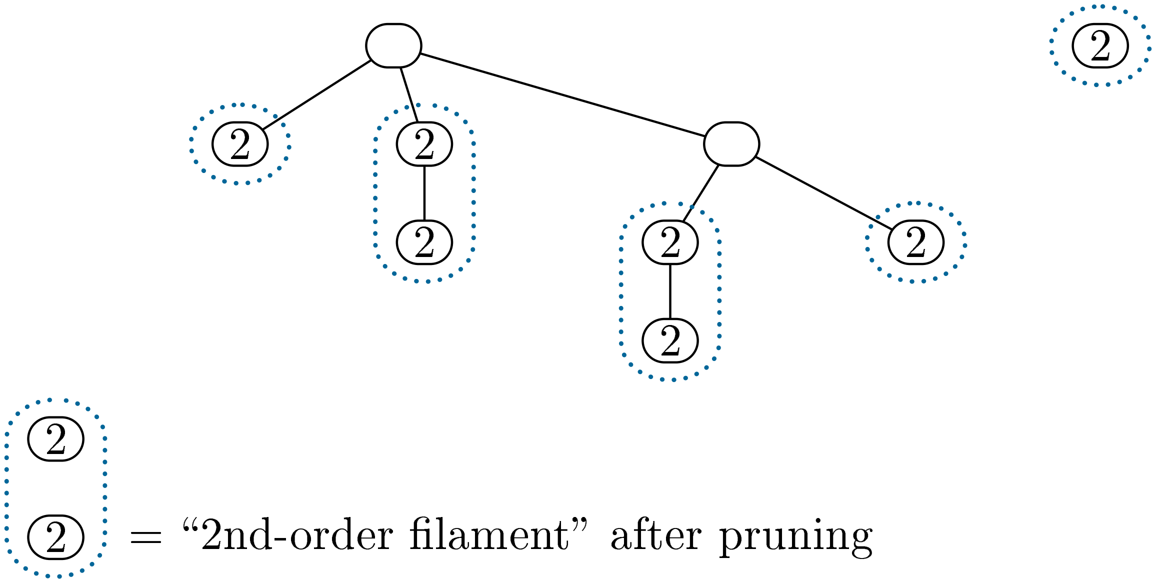
Problem 3 (G. M. F. X. Viennot):

Pruning a Forest

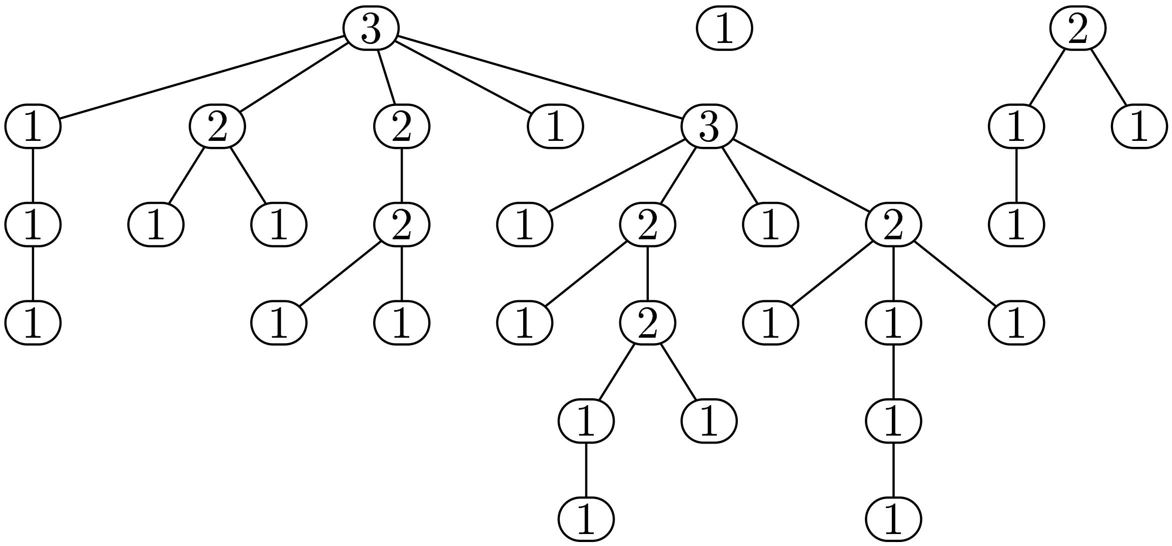


a forest (ordered)


1
= “filament”



2
2
2
2
2
2
2
2
2
2
= "2nd-order filament" after pruning



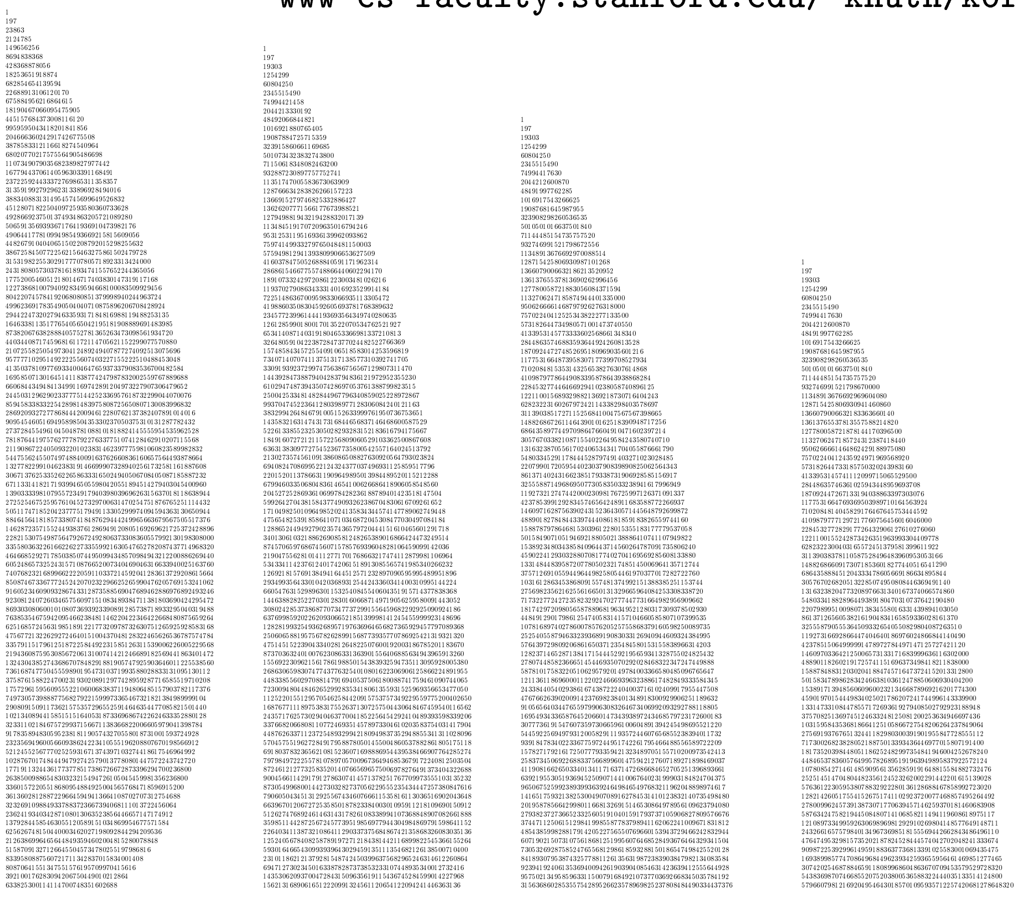
Theorem (Mireille Vauchaussade de Chaumont, 1985)**:**

The number of n-node binary trees with Strahler number k equals the number of n-node forests with pruning number k.

```
Checking all forests with 8 nodes...
After 0 prunings (cases with k leaves):
0 0 0 0 0 0 0 0 1430
After 1 pruning (cases with k 1s):
128 432 500 275 81 13 1 0 0
After 2 prunings (cases with k 2s):
1288 13 1 0 0 0 0 0 0
After 3 prunings (cases with k 3s):
14 0 0 0 0 0 0 0 0
```

1		
13		
81		
275		
500	1	
432	13	
128	1288	14

`www-cs-faculty.stanford.edu/~knuth/korsh-pruning.c`



($n = 100$: counting the 1s, 2s, 3s, and 4s)

(calculations by Jim Korsh (2007))

References

M. Vauchaussade de Chaumont and X. G. Viennot, “Enumeration of RNAs secondary structure by complexity,” *Lecture Notes in Biomathematics* **57** (1985), 360–365.

X. G. Viennot, “Trees everywhere,” *Lecture Notes in Computer Science* **431** (1990), 18–41.

D. Zeilberger, “A bijection from ordered trees to binary trees that sends the pruning order to the Strahler number,” *Discrete Mathematics* **82** (1990), 89–92.

D. E. Knuth, “Three Catalan bijections,” *Institut Mittag-Leffler Reports*, No. 04, 2004/2005, Spring (2005), 19 pages.

D. E. Knuth, “Trees, Rivers, and RNA,” (December 2007), a Webinar available at *Stanford University Online*.

Problem 4:

Lattice Paths of Slope 2/5

$$A[i,j] = \begin{cases} 0, & \text{if } j \geq 2i/5 + 2/5, \\ A[i-1,j] + A[i,j-1], & \text{if } j < 2i/5 + 2/5; \end{cases}$$

$$B[i,j] = \begin{cases} 0, & \text{if } j \geq 2i/5 + 1/5, \\ B[i-1,j] + B[i,j-1], & \text{if } j < 2i/5 + 1/5; \end{cases}$$

$A[i,0] = B[i,0] = 1$. When $0 \leq i \leq 4$ and $0 \leq j \leq 10$ we have:

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 0 & 0 & 0 & 0 & 0 & 4 & 9 & 15 & 22 & 30 & 39 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 15 & 37 & 67 & 106 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 106 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 0 & 0 & 0 & 0 & 0 & 3 & 7 & 12 & 18 & 25 & 33 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 18 & 43 & 76 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 76 \end{pmatrix}$$

Thus $A[x,y]$ enumerates lattice paths from $(0,0)$ that stay in the region $y < \frac{2}{5}x + \frac{2}{5}$, while $B[x,y]$ enumerates the paths that stay in the region $y < \frac{2}{5}x + \frac{1}{5}$.

Theorem (Nakamigawa, Tokushige, 2012)**:**

$$A[5t-1, 2t-1] + B[5t-1, 2t-1] = \frac{2}{7t-1}\binom{7t-1}{2t}, \quad \text{for all } t \geq 1.$$

Empirical observation:

$$\frac{A[5t-1, 2t-1]}{B[5t-1, 2t-1]} \;=\; a - \frac{b}{t} + O(t^{-2}),$$

where $a \approx 1.63026$ and $b \approx 0.159$ (I think).

Reference

Tomoki Nakamigawa and Norihide Tokushige, "Counting lattice paths via a new cycle lemma," *SIAM Journal on Discrete Mathematics* **26** (2012), 745–754.

News Flash!

Cyril Banderier announced on Friday, 20 June 2014, that he has solved this problem (and the series are algebraic)!

Problem 5:

The Principle of Negligible Perturbation

Random Structures & Algorithms

VOLUME 1 NUMBER 1 SPRING 1990

“The idea will be to change the transition probabilities between certain nodes, obtaining a ‘perturbed’ probability distribution $\Pr'$ on which it is relatively easy to compute the probability of some given event. Let t be a level of the tree, and let E be the set of all nodes at level t such that the given event is true. Let C be the set of all nodes α at level t whose probability has been perturbed somewhere along the path from the root to α; thus, $\Pr(\alpha) = \Pr'(\alpha)$ for all $\alpha \notin C$. Summing over all $\alpha \notin C$ and taking complements tells us that $\Pr(C) = \Pr'(C)$. If $\Pr(C)$ is small, then the perturbation will have a negligible effect on the probability of E, because

$$|\Pr(E) - \Pr'(E)| = \cdots \leq \cdots = 2\Pr(C).$$

Expected values can be estimated in a similar way.”

“(The principle of negligible perturbation seems almost absurdly simple, but we will see that it simplifies our analyses in surprisingly nontrivial ways. The idea is similar in spirit to Laplace’s method of asymptotic analysis, where integrals are estimated by changing the integrand in unimportant portions of the domain. Another kindred method is Wilkinson’s well-known technique of ‘backward error analysis,’ in which numerical errors are conveniently studied by assuming that exact answers have been obtained from approximate data; the actual situation, in which approximate answers are calculated from exact data, is more difficult to handle directly.)”

“The previous proof uses a convenient simplification, indicated by the words ‘According to Lemma 1, we may assume that … ’. The assumption we are making holds q.s., but it is not always true; moreover, it is a probabilistic assertion about time $N+1$. So we should be careful that we are not fallaciously using the future to influence probability calculations in the past. A rigorous justification can be made by appealing to the principle of negligible perturbation: We simply recompute the transition probabilities when the assumption $k_h \le 2n^\delta$ is invalid.

More precisely, if α is any node in the branching process, we let the perturbed transition probabilities from α to α_h^a and α_h^r be $1/(k_h'+1)n$ and $k_h'/(k_h'+1)n$, respectively, where

$$k_h' = \min(k_h, 2n^\delta)\,.$$

The proof of Lemma 3 is valid for the perturbed branching process, using k_h' in place of k_h in the formula for $P(\alpha, m, t)$.”

"Notice that the principle of negligible perturbations has permitted us, in this proof, to estimate probabilities of events that start at time t by using assumptions that might fail at some future time $> t$. (Thus, $\|A_p\|$ might be $\leq 2n^{2\delta}(\log n)^2$ at the beginning of a run but not at the end.) Arguments based on a weaker principle, which would require only that the assumptions hold at time t, would be more complicated; we would have to argue that $\|A_p\|$ cannot grow by more than 1 at each time step, and our upper bound would be $\bigl(\rho' + m/\bigl(n(2n^{\delta}+1)\bigr)\bigr)^m$ instead of $(\rho')^m$."

The basic idea is this. Suppose we're analyzing a rather complicated algorithm, $\mathcal{A}$. We replace it with a simpler algorithm, $\mathcal{A}'$, which does the same thing with high probability.

When $\mathcal{A}'$ differs from $\mathcal{A}$, it may do entirely stupid things: It might deliver results out of range, the invariants on its data structures might be violated, it might divide by zero or go haywire in any number of bizarre ways. It certainly need not compute a decent answer. But if the probability that $\mathcal{A}'$ and $\mathcal{A}$ behave differently is really small, we're allowed to pretend that the two algorithms have the same expected performance.

References

Svante Janson and Donald E. Knuth, "Shellsort with three increments," *Random Structures & Algorithms* **10** (1997), 125–142.

Donald E. Knuth, Rajeev Motwani, and Boris Pittel, "Stable husbands," *Random Structures & Algorithms* **1** (1990), 1–14.

✔ **applause**