On some combinatorial problems in lambda calculus

Katarzyna Grygiel

Computational Logic and Applications
Kraków 2008

Outline

- Motivation

Outline

- Motivation
- Lambda trees, lambda terms

Outline

- Motivation
- Lambda trees, lambda terms
- First attempts: hits and misses

Outline

- Motivation
- Lambda trees, lambda terms
- First attempts: hits and misses
- Some bounds

Outline

- Motivation
- Lambda trees, lambda terms
- First attempts: hits and misses
- Some bounds
- Analytic approach

Outline

- Motivation
- Lambda trees, lambda terms
- First attempts: hits and misses
- Some bounds
- Analytic approach
- Some other questions

Preliminary problem

Question

How many programs have the halting property?

Preliminary problem

Question

How many programs have the halting property?

In terms of lambda calculus: compute the following limit

$$\lim_{n\to\infty} \frac{SNTerms(n)}{Terms(n)} =?$$

where $SNTerms(n)$ denotes the number of closed lambda terms of length n that strongly normalize and $Terms(n)$ is the the number of all closed lambda terms of length n.

Catalan numbers

Binary tree

A binary tree is a rooted tree in which each vertex has up to two children and each child node is called left or right.

Catalan numbers

Binary tree

A binary tree is a rooted tree in which each vertex has up to two children and each child node is called left or right.

Catalan numbers

$C(n)$ — the number of binary trees with exactly n leaves

First values: $1, 1, 2, 5, 14, 42, \ldots$

Generating function: $c(x) = \dfrac{1 - \sqrt{1-4x}}{2x}$,

Exact formula: $C(n) = \dfrac{1}{n+1}\dbinom{2n}{n}$.

Motzkin numbers

Unary-binary tree

A unary-binary tree is a rooted tree in which each vertex has up to two children.

Motzkin numbers

Unary-binary tree

A unary-binary tree is a rooted tree in which each vertex has up to two children.

Motzkin numbers

$M(n)$ — the number of u-b trees with exactly n nodes

First values: $1, 1, 2, 4, 9, 21, 51, 127, 323, \ldots$

Generating function: $m(x) = \dfrac{1 - x - \sqrt{(1-3x)(1+x)}}{2x}$

Asymptotics: $M(n) \sim 3^n \sqrt{\dfrac{3}{4\pi n^3}}$

Specified Motzkin numbers

Specified Motzkin numbers

$M(n,k)$ — the number of rooted u-b trees with n vertices and exactly k leaves

Generating function: $m(x,y) = \dfrac{1-x-\sqrt{(1-x)^2-4x^2y}}{2x}$

Exact formula: $M(n,k) = C(k-1)\dbinom{n-1}{n-2k+1}$

Lambda trees

Lambda tree

A lambda tree is a unary-binary tree in which each leaf can be labelled in as many ways as many vertices with exactly one child occur on the path from the leaf to the root of the tree.

Lambda trees

Lambda tree

A lambda tree is a unary-binary tree in which each leaf can be labelled in as many ways as many vertices with exactly one child occur on the path from the leaf to the root of the tree.

Problem

How many lambda trees of a given size are there?

What are lambda trees actually?

Since we do not bother about the names of variables, we count lambda terms up to α-conversion.
We count the length of a lambda term as follows:

$$
\begin{aligned}
|x| &= 1 \\
|\lambda x.M| &= 1 + |M| \\
|MN| &= 1 + |M| + |N|
\end{aligned}
$$

What are lambda trees actually?

Since we do not bother about the names of variables, we count lambda terms up to α-conversion.
We count the length of a lambda term as follows:

$$\begin{aligned} |x| &= 1 \\ |\lambda x.M| &= 1 + |M| \\ |MN| &= 1 + |M| + |N| \end{aligned}$$

Correspondence

There is a one-to-one correspondence between closed lambda terms of length n and lambda trees of size n.

T_n sequence

First values of T_n sequence:

$$0, 1, 2, 4, 13, 42, 139, 506, 1915, 7558, 31092, 132170, 580466, \ldots$$

T_n sequence

First values of T_n sequence:

$0, 1, 2, 4, 13, 42, 139, 506, 1915, 7558, 31092, 132170, 580466, \ldots$

Already in the On-Line Encyclopedia of Integer Sequences (Christophe Raffalli, 9.02.2008)
`http://www.research.att.com/~njas/sequences/A135501`

T_n sequence

First values of T_n sequence:

$0, 1, 2, 4, 13, 42, 139, 506, 1915, 7558, 31092, 132170, 580466, \ldots$

Already in the On-Line Encyclopedia of Integer Sequences (Christophe Raffalli, 9.02.2008)
`http://www.research.att.com/~njas/sequences/A135501`

`...with a note:` *Is there a generating function?*

A good-looking recurrence

$T(n,k)$

Let $T(n,k)$ denote the number of lambda trees with n vertices and in which each leaf can be labelled either in the standard way or with an element from a set of k elements.

A good-looking recurrence

$T(n, k)$

Let $T(n, k)$ denote the number of lambda trees with n vertices and in which each leaf can be labelled either in the standard way or with an element from a set of k elements.

A nice [?] recurrence

$$T(0, k) = 0$$
$$T(1, k) = k$$
$$T(n, k) = T(n-1, k+1) + \sum_{i=1}^{n-2} T(i, k)T(n-i-1, k).$$

Let us define

$$\varphi_k(x) = \sum_{n \in \mathbb{N}} T(n,k)x^n$$

We get

$$\varphi_{k+1}(x) = \frac{\varphi_k(x)}{x} - (\varphi_k(x))^2 - k$$

Let us define

$$\varphi_k(x) = \sum_{n \in \mathbb{N}} T(n,k)x^n$$

We get

$$\varphi_{k+1}(x) = \frac{\varphi_k(x)}{x} - (\varphi_k(x))^2 - k$$

Well, $\varphi_0(x)$ is the generating function for T_n...

$T(n,k)$ as a polynomial

We can look at $T(n,k)$ as at the polynomial in k:

$T(n, k)$ as a polynomial

We can look at $T(n, k)$ as at the polynomial in k:

$$
\begin{aligned}
T(1,k) &= k,\\
T(2,k) &= k+1,\\
T(3,k) &= k^2+k+2,\\
T(4,k) &= 3k^2+5k+4,\\
T(5,k) &= 2k^3+6k^2+17k+13,\\
T(6,k) &= 10k^3+26k^2+49k+42,\\
T(7,k) &= 5k^4+O(k^3),\\
T(8,k) &= 35k^4+O(k^3).
\end{aligned}
$$

d_n sequence

Let d_i be defined as follows: $d_0 = 0$ and d_n is the greatest exponent of k in the polynomial $T(n, k)$.

d_n sequence

Let d_i be defined as follows: $d_0 = 0$ and d_n is the greatest exponent of k in the polynomial $T(n,k)$.

Observation

For any $n \geqslant 1$ we have

$$d_n = \left\lceil \frac{n}{2} \right\rceil$$

Moreover,

- if n is even then $d_n = d_{n-1} = d_i + d_{n-i-1}$ for $i = 1, \ldots, n-2$,
- if n is odd then $d_n = d_{2i-1} + d_{n-2i}$ for $i = 1, \ldots, \frac{n-1}{2}$

w_n sequence

Let w_i be the sequence of leading coefficients in $T(i,k)$.

w_n sequence

Let w_i be the sequence of leading coefficients in $T(i,k)$.

Observation

For any $s \in \mathbb{N}$ we have

$$\begin{aligned} w_0 &= 0 \\ w_1 &= 1 \\ w_{2s+1} &= \sum_{i=1}^{s} w_{2i-1} \cdot w_{2s-2i+1} \\ w_{2s+2} &= w_{2s+1} + \sum_{i=0}^{2s} w_i \cdot w_{2s-i+1}. \end{aligned}$$

w_n sequence

Lemma

For any natural number s the following equalities hold:

$$w_{2s+1} = \frac{1}{s+1}\binom{2s}{s} = C(s),$$
$$w_{2s} = \frac{1}{2}\binom{2s}{s} = \frac{s+1}{2}C(s).$$

. . . and a bound at last

Lower bound

For any $n > 0$ the following inequalities hold:

$$T(2n,k) \geqslant \frac{(n+1)C(n)}{2}k^n$$
$$T(2n+1,k) \geqslant C(n)k^{n+1}.$$

... and a bound at last

Lower bound

For any $n > 0$ the following inequalities hold:

$$T(2n,k) \geqslant \frac{(n+1)C(n)}{2}k^n$$
$$T(2n+1,k) \geqslant C(n)k^{n+1}.$$

Upgraded lower bound

For any $n > 0$ the following inequality holds:

$$T(n,k) \geqslant M(n-1) + \sum_{i=1}^{\lceil \frac{n}{2} \rceil} M(n,i)k^i$$

Upper bound

Upper bound

For any $n > 0$ the following inequality holds:

$$T(n,k) \leqslant \sum_{i=1}^{\lceil \frac{n}{2} \rceil} M(n,i)(n-2i+k+1)^i$$

A new idea

Let us consider $F(n, m, k)$ where

- n is the number of internal nodes
- m is the number of “open” leaves
- k is the number of “closed” leaves

A new idea

Let us consider $F(n, m, k)$ where

- n is the number of internal nodes
- m is the number of "open" leaves
- k is the number of "closed" leaves

Now, $\sum_{n+k=N} F(n, 0, k)$ is the number of lambda trees of size N

Define

$$f(z, u, v) = \sum_{n,m,k \in \mathbb{N}} F(n, m, k) z^n u^m v^k$$

The equation (D. Gardy, B. Gittenberger)

Functional equation

$$f(z,u,v) = zf(z,u,v)^2 + zf(z,u+v,v) + u$$

The equation (D. Gardy, B. Gittenberger)

Functional equation

$$f(z,u,v) = zf(z,u,v)^2 + zf(z,u+v,v) + u$$

What we need is

Goal

$$f(z,0,z)$$

Partial answer

The solution is in the form of nested radicals. Its radius of convergence is equal to 0.

Partial answer

The solution is in the form of nested radicals. Its radius of convergence is equal to 0.

$$\frac{z}{2}-\frac{z}{2}\sqrt{1-2z-4z^2}$$

$$\frac{z}{2}-\frac{z}{2}\sqrt{1-2z-4z^2+2z\sqrt{1-2z-8z^2}}$$

$$\frac{z}{2}-\frac{z}{2}\sqrt{1-2z-4z^2+2z\sqrt{1-2z-8z^2+2z\sqrt{1-2z-12z^2}}}$$

$$\ldots$$

A guess

$$T_n \sim \frac{3^n n}{9} \Gamma\left(\frac{n}{3}\right)$$

where $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$

λI-terms

λI-terms

A lambda term is called a λI-term if in its every subterm of the form $\lambda x.M$, x is a free variable in M.

λI-terms

λI-terms

A lambda term is called a λI-term if in its every subterm of the form $\lambda x.M$, x is a free variable in M.

$I(n,k)$

Let $I(n,0)$ denote the number of closed λI-terms of length n and $I(n,k)$ for $k \geqslant 1$ — the number of such λI-terms M that are of length n and $FV(M)$ is of cardinality k.

λI-terms

λI-terms

A lambda term is called a λI-term if in its every subterm of the form $\lambda x.M$, x is a free variable in M.

$I(n,k)$

Let $I(n,0)$ denote the number of closed λI-terms of length n and $I(n,k)$ for $k \geqslant 1$ — the number of such λI-terms M that are of length n and $FV(M)$ is of cardinality k.

First values of $I(n,0)$

$$0, 0, 1, 0, 1, 5, 2, 26, 65, 141, \ldots$$

λI-terms

. . . and the recurrence

$$
\begin{aligned}
I(0,k) =& I(1,0) = 0\\
I(1,1) =& 1\\
I(n,k) =& I(n-1,k+1) + \sum_{i=0}^{n-1} \sum_{k_1,k_2 \geqslant 0} \binom{k}{k_1+k_2-k} \cdot \\
& \cdot \binom{2k-k_1-k_2}{k-k_2} I(i,k_1) I(n-i-1,k_2).
\end{aligned}
$$

λBCI-terms

λBCI-terms

A lambda term is called a λBCI-term if it is a λI-term and in its every subterm of the form $\lambda x.M$, x occurs free in M exactly once.

λBCI-terms

λBCI-terms

A lambda term is called a λBCI-term if it is a λI-term and in its every subterm of the form $\lambda x.M$, x occurs free in M exactly once.

$B(n, k)$

Let $B(n, k)$ denote the number of λBCI-terms of length n and exactly k free variables.

λBCI-terms

λBCI-terms

A lambda term is called a λBCI-term if it is a λI-term and in its every subterm of the form $\lambda x.M$, x occurs free in M exactly once.

$B(n,k)$

Let $B(n,k)$ denote the number of λBCI-terms of length n and exactly k free variables.

First values of $B(n,0)$

$$0, 0, 1, 0, 0, 5, 0, 0, 60, 0, 0, 1105, \ldots$$

λBCI-terms

. . . and the recurrence

$$\begin{aligned}
B(0,k) &= B(1,0) = 0 \\
B(1,1) &= 1 \\
B(n,k) &= B(n-1,k+1) + \\
&\quad + \sum_{i=0}^{n-1} \sum_{j=0}^{k} \binom{k}{j} B(i,j) B(n-i-1,k-j).
\end{aligned}$$

λBCI-terms

Let us define

$$b(x,y) = \sum_{k,n \geqslant 0} \frac{B(n,k)}{k!} x^n y^k$$

λBCI-terms

Let us define

$$b(x,y) = \sum_{k,n\geqslant 0} \frac{B(n,k)}{k!} x^n y^k$$

Differential equation

Function $b(x,y)$ satisfies the following equation

$$b(x,y) = xy + xb^2(x,y) + x\frac{\mathrm{d}b(x,y)}{\mathrm{d}y}$$

Application vs. abstraction

Problem

What is the density of lambda terms being in form of abstraction?

Application vs. abstraction

Problem

What is the density of lambda terms being in form of abstraction?

Conjecture

A random closed lambda term is in the form of abstraction.

The end.