Counting λ-terms

CLA 2008 – Kraków

René David, Christophe Raffalli and Guillaume Theyssier

LAMA (CNRS, Université de Savoie)

June 7, 2008

Outline of the talk

Introduction, Problem Statement and First Results

Polynomials

Bounding

Counting specific terms

Counting variables

In the Turing world...

Research directions

Outline of the talk

The main polynomial

Number of λ-terms with at most X variables

$P_1(X) = X$

$P_{n+1}(X) = P_n(X+1) + \sum_{l+r=n} P_l(X) P_r(X)$

This gives:

$P_2(X) = 1 + X$

$P_3(X) = 2 + X + X^2$

$P_4(X) = 4 + 5X + 3X^2$

$P_5(X) = 13 + 17X + 6X^2 + 2X^3$

$P_6(X) = 42 + 49X + 26X^2 + 10X^3$

$P_7(X) = 139 + \ldots + 5X^4$

$P_8(X) = 506 + \ldots + 35X^4$

What are the coefficients of P ?

$P_n(X) = \sum_{i=0}^{\lceil \frac{n}{2} \rceil} c_{n,i} X^i$

The coefficients of the main polynomial

What are the coefficients of P ?

$P_n(X) = \sum_{i=0}^{\lceil \frac{n}{2} \rceil} c_{n,i}\, X^i$

Number of λ-terms with 1 variable occurring i-times

$c_{1,1} = 1, c_{1,k} = 0 (k \neq 1)$

$c_{n+1,k} = \sum_{i=0}^{+\infty} \binom{k+i}{i} c_{n,k+i} + \sum_{\substack{l+r=n \\ i+j=k}} c_{l,i} c_{r,j}$

The derivatives of the main polynomial

Derivation

$\frac{1}{k!}P_n^{(k)}(X) = \frac{1}{k!}\sum_{i=k}^{\lceil \frac{n}{2} \rceil - k} c_{n,i}\, \frac{i!}{(i-k)!}X^{i-k} = \sum_{i=k}^{\lceil \frac{n}{2} \rceil - k} c_{n,i}\, \binom{i}{k}X^{i-k}$

Therefore

$\frac{1}{k!}P_n^{(k)}(X)$:

Number of λ-terms with at most X variables + 1 variable occurring k-times

$P_n(X+1) = \sum_{i=0}^{+\infty} \frac{1}{k!}P_n^{(k)}(X)$

Term with exactly X variables

Formula

$$Q_n(0) = P_n(0),\ Q_n(X+1) = P_n(X+1) - P_n(X)$$
$$Q_1(1) = 1,\ Q_1(X) = 0 (X \neq 1)$$
$$Q_{n+1}(X) = Q_n(X) + Q_n(X+1) + \sum_{\substack{l+r=n \\ X_0+X_1+X_2=X}} \frac{X!}{X_0!X_1!X_2!} Q_l(X_0+X_1) Q_r(X_0+X_2)$$

Useful to count some classes of terms

Example: affine terms

$$A_1(1) = 1,\ A_1(X) = 0 (X \neq 1)$$
$$A_{n+1}(X) = A_n(X) + A_n(X+1) + \sum_{\substack{l+r=n \\ X_1+X_2=X}} \frac{X!}{X_1!X_2!} A_l(X_1) A_r(X_2)$$

Outline of the talk

Bounding

First result

$\forall \alpha < 1$ we have f, g exponential, such that:

$$f(n - n^{\alpha})(n^{\alpha} + X - 1)^{\frac{n-n^{\alpha}}{2}} \leq P_n(X) \leq g(n)(n+X)^{\lceil \frac{n}{2} \rceil}$$

The asymptotic development of $P_n(0)$

$$P_n(0) \sim n^{\lceil \frac{n}{2} \rceil} \ldots$$

Better upper bound

Size zero for variables

$P_0(X) = X$ instead of $P_1(X) = X$

$$P_n(X) \leq C(n)(n+X)^{n+1}$$

The dominant coefficient is exact.

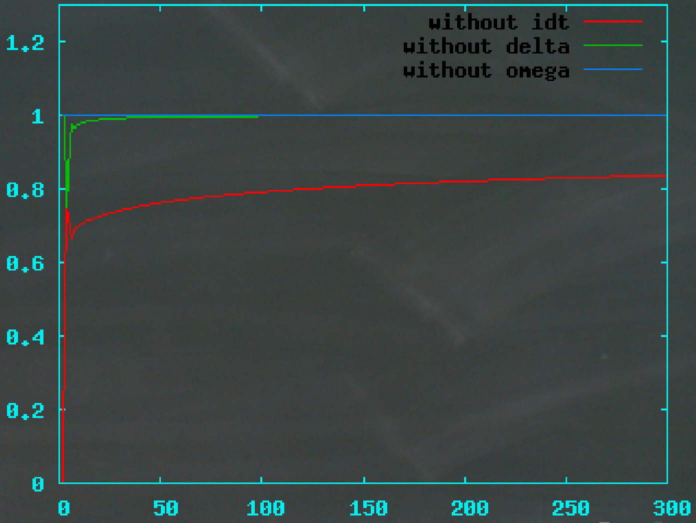
Conjecture for size one variables

$$P_{2n}(X) \leq (2n-1)C(n-1)(2n-1+X)^n$$

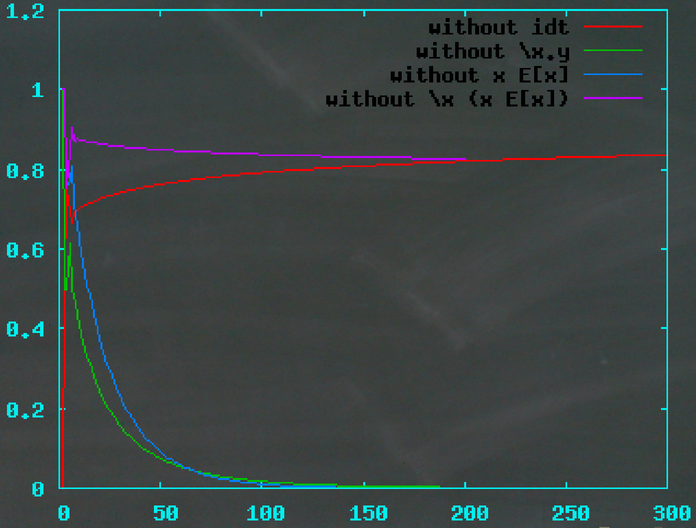
$$P_{2n+1}(X) \leq C(n)(2n+X)^n$$

Outline of the talk

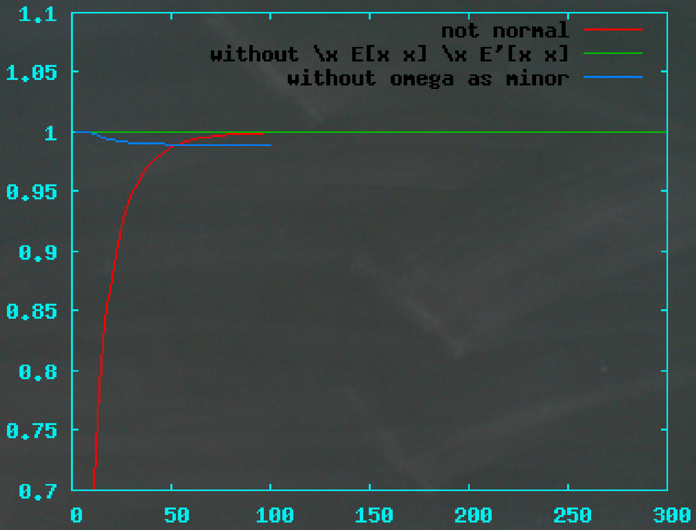
Terms avoiding a given term as subterm



Terms avoiding a given pattern as subterm

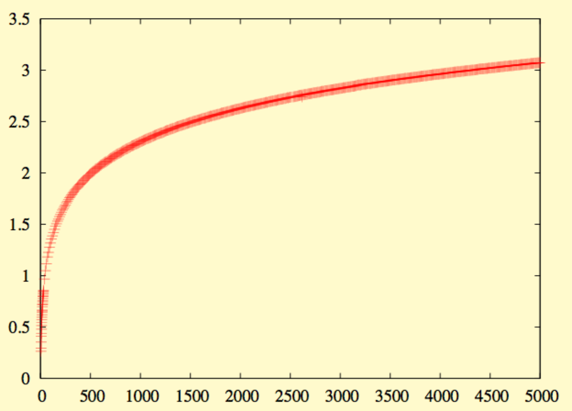


Normalisable terms



Other results by Jue Wang (Boston) 2004

- Typable terms converges toward probability 0
- Ratio of application nodes / lambda nodes



Outline of the talk

General setting

Recurrence formula

$$\begin{aligned} P_0(X) &= 0 \\ P_{n+1}(X) &= V_{n+1}(X) + P_n(X+1) + \sum_{i+j=n} P_i(X)P_j(X) \end{aligned}$$

- n : size
- X : number of available variables
- $V_n(X)$: the number of terms which are variables

General setting

Recurrence formula

$$\begin{aligned} P_0(X) &= 0 \\ P_{n+1}(X) &= V_{n+1}(X) + P_n(X+1) + \sum_{i+j=n} P_i(X)P_j(X) \end{aligned}$$

- n : size
- X : number of available variables
- $V_n(X)$: the number of terms which are variables

Until now

unit size: $V_1(X) = X$ and $V_n(X) = 0$ if $n > 1$

Is it realistic?

General setting

Recurrence formula

$$\begin{aligned} P_0(X) &= 0 \\ P_{n+1}(X) &= V_{n+1}(X) + P_n(X+1) + \sum_{i+j=n} P_i(X)P_j(X) \end{aligned}$$

- n : size
- X : number of available variables
- $V_n(X)$: the number of terms which are variables

Until now

unit size: $V_1(X) = X$ and $V_n(X) = 0$ if $n > 1$

Is it realistic? Well...

- 1 bit to represent many possible choices
- a super-exponential number of terms of a given “size”

Non-unit cost of variables

A concrete notion of size

- use De Bruijn indexes
- take into account the cost of such indexes

Non-unit cost of variables

A concrete notion of size

- use De Bruijn indexes
- take into account the cost of such indexes

Possible choices

- unary De Bruijn index:

$$V_n(X) = \begin{cases} 1 & \text{if } n \leq X \\ 0 & \text{else} \end{cases}$$

- binary De Bruijn index:

$$V_n(X) = \#\{i : i \leq X \text{ and } |bin(i)| = n\}$$

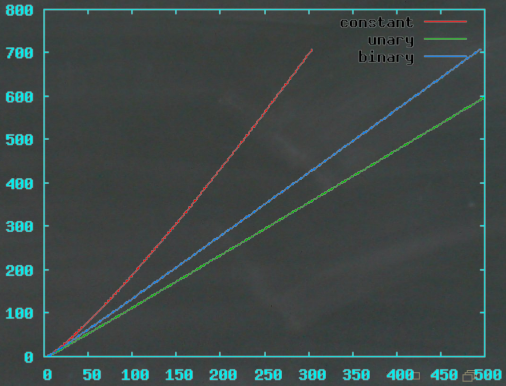
Small closed terms in the 3 models

size	constant	unary	binary
2	$\lambda x.x$	$\lambda.1$	$\lambda.0$
3	$\lambda x_1, x_2.x_i \quad (\times 2)$	$\lambda^2.1$	$\lambda^2.0$ $\lambda^2.1$
4	$\lambda x_1, x_2, x_3.x_i \quad (\times 3)$ $\lambda x.x\ x$	$\lambda^3.1$ $\lambda^2.11$ $\lambda.1\ 1$	$\lambda^3.0$ $\lambda^3.1$ $\lambda.0\ 0$
5	$\lambda x_1, x_2, x_3, x_4.x_i \quad (\times 4)$ $\lambda x_1, x_2.x_i\ x_j \quad (\times 4)$ $\lambda x_1.x_1\ (\lambda x_2.x_i) \quad (\times 4)$ $(\lambda x.x)\ (\lambda x.x)$	$\lambda^4.1$ $\lambda^3.11$ $\lambda^2.1\ 1$ $\lambda.(1\ \lambda.1)$ $\lambda.(\lambda.1\ 1)$ $(\lambda.1)\ (\lambda.1)$	$\lambda^4.(0/1) \quad (\times 2)$ $\lambda^3.11$ $\lambda^2.(0/1)\ (0/1) \quad (\times 4)$ $\lambda.0\ (\lambda.(0/1)) \quad (\times 4)$ $(\lambda.0)\ (\lambda.0)$

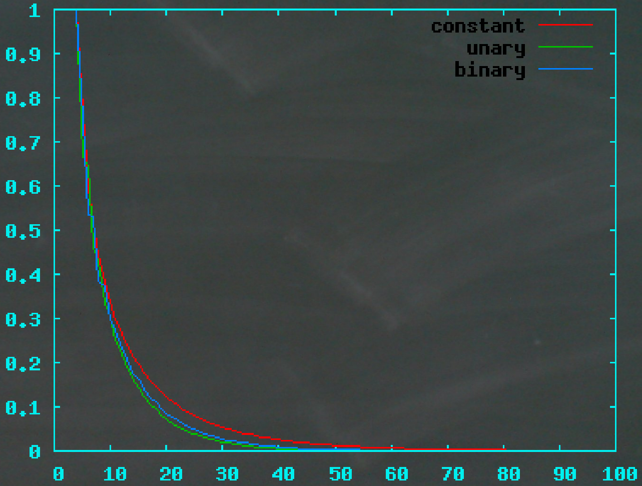
Differences between the 3 models

For non-unit costs

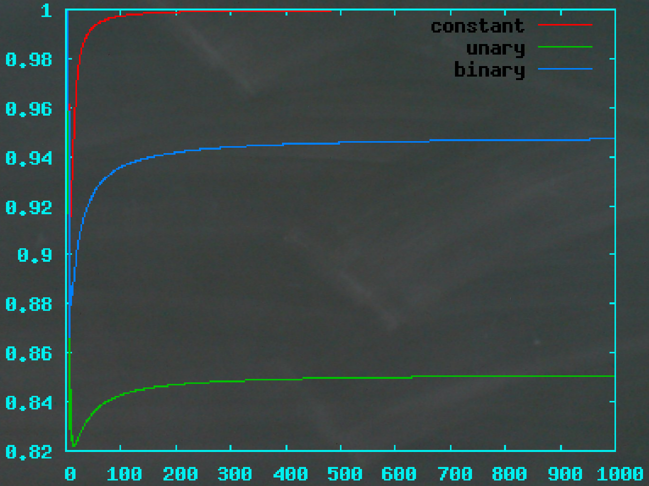
- the growth rate is only exponential
- $P_n(X)$ is not a polynomial in X



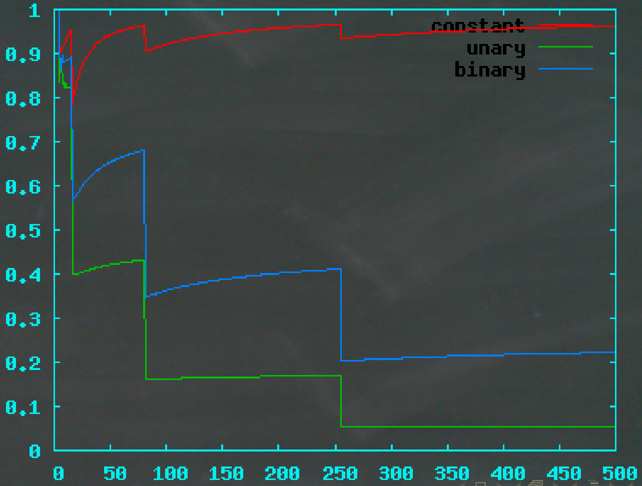
Proportion of terms in normal form



Proportion of terms starting with a lambda

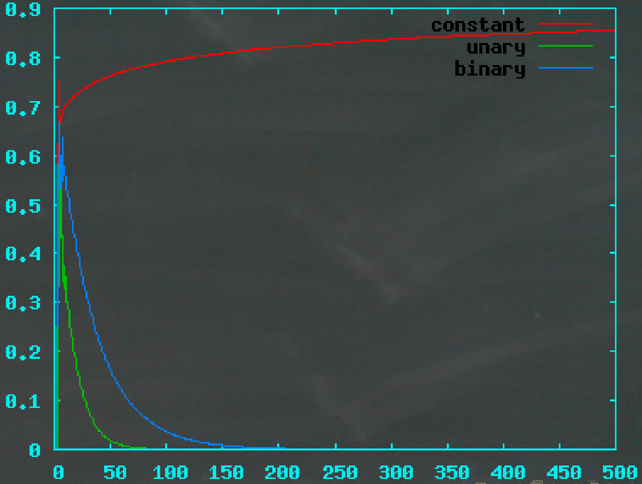


Proportion of terms starting with $\left\lceil \sqrt{\sqrt{n}} \right\rceil$ lambdas



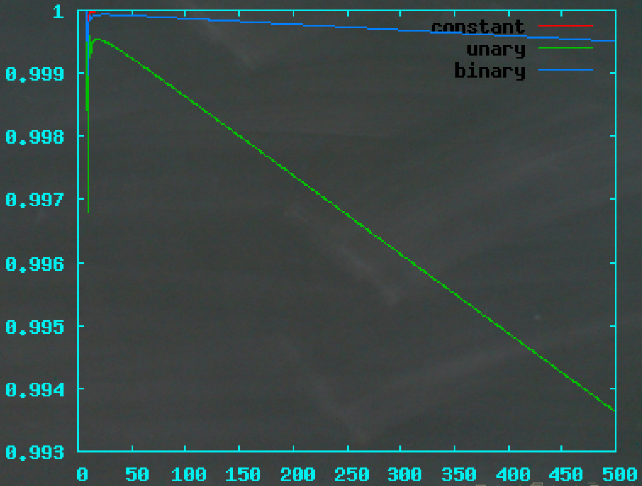
Proportion of terms not containing identity

Remark: id has the same size for the 3 models



Proportion of terms not containing Ω

Remark: Ω has the same size for the 3 models



Outline of the talk

Analogous questions in the Turing world

- **syntactical objects:** fully specified Turing machines

$$\delta : \{0,1\} \times Q_n \to \{0,1\} \times (Q_n \cup \{halt\}) \times \{L, R\}$$

- **studied property:** the halting problem (H)

Analogous questions in the Turing world

- **syntactical objects:** fully specified Turing machines

$$\delta : \{0,1\} \times Q_n \to \{0,1\} \times (Q_n \cup \{halt\}) \times \{L, R\}$$

- **studied property:** the halting problem (H)

Simple facts

1. there are $(4n+4)^{2n}$ machines with n states
2. $d(H) < 1 - 1/e^2$ (if it exists)

- determining $d(H)$ in the general setting is still open, but...

Result of another kind...

Theorem (J.D. Hamkins, A. Miasnikov, 2005)

There exists a set B of machines such that:

1. *B is decidable (in polynomial time);*
2. *B has asymptotic density 1;*
3. *the halting problem is decidable (in polynomial time) on B.*

Result of another kind...

Theorem (J.D. Hamkins, A. Miasnikov, 2005)

There exists a set B of machines such that:

1. *B is decidable (in polynomial time);*
2. *B has asymptotic density* 1*;*
3. *the halting problem is decidable (in polynomial time) on B.*

model sensitive result!

- it uses semi-infinite tape Turing-machines!
- if falling-off the tape is not halting then $d(H) = 0$

Result of another kind...

Theorem (J.D. Hamkins, A. Miasnikov, 2005)

There exists a set B of machines such that:

1. *B is decidable (in polynomial time);*
2. *B has asymptotic density* 1*;*
3. *the halting problem is decidable (in polynomial time) on B.*

model sensitive result!

- it uses semi-infinite tape Turing-machines!
- if falling-off the tape is not halting then $d(H) = 0$

Proof idea

Uses a trick to convert the first steps of a random Turing machine into a random walk of the head.

Yet another kind of result...

Definition

A set is **strongly generic** if its complement converges exponentially fast to 0 in density.

Yet another kind of result...

Definition

A set is **strongly generic** if its complement converges exponentially fast to 0 in density.

Theorem (A. Rybalov, 2005)

The halting problem is strongly undecidable, i.e. it is undecidable on any strongly generic set of Turing machines.

Yet another kind of result...

Definition

A set is **strongly generic** if its complement converges exponentially fast to 0 in density.

Theorem (A. Rybalov, 2005)

The halting problem is strongly undecidable, i.e. it is undecidable on any strongly generic set of Turing machines.

Proof idea

The set of machines computing a given (computable) function is not strongly negligible (one can add garbage code at some places without affecting the computation).

Outline of the talk

Research directions

- "combinatorial" proof that for any fixed tree T_0, a large random tree always contains T_0 (Kolmogorov incompressibility method?)

Research directions

- "combinatorial" proof that for any fixed tree T_0, a large random tree always contains T_0 (Kolmogorov incompressibility method?)
- typical tree shape of random λ-terms (e.g. relation between height and size)

Research directions

- "combinatorial" proof that for any fixed tree T_0, a large random tree always contains T_0 (Kolmogorov incompressibility method?)
- typical tree shape of random λ-terms (e.g. relation between height and size)
- other experiments: random generation of big λ-terms to have an idea of their shape

Research directions

- possible properties to look at:
  - *about shape:* normality, beginning by a λ,
  - *depending on the reduction rule:* SN, WN,
  - *other:* typability

Research directions

- possible properties to look at:
  - *about shape:* normality, beginning by a λ,
  - *depending on the reduction rule:* SN, WN,
  - *other:* typability
- finding sets of terms with density 1 for which SN or WN becomes decidable

Research directions

- possible properties to look at:
  - *about shape:* normality, beginning by a λ,
  - *depending on the reduction rule:* SN, WN,
  - *other:* typability
- finding sets of terms with density 1 for which SN or WN becomes decidable
- consider other distributions than the uniform one (e.g., similar to the BST model)

Research directions

- possible properties to look at:
  - *about shape:* normality, beginning by a λ,
  - *depending on the reduction rule:* SN, WN,
  - *other:* typability
- finding sets of terms with density 1 for which SN or WN becomes decidable
- consider other distributions than the uniform one (e.g., similar to the BST model)
- any non-trivial proved result about random λ-terms with any representation of variables is welcome!