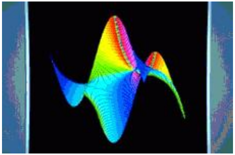
Automatic Analysis of Combinatorial Structures, Old and New

Bruno Salvy
Bruno.Salvy@inria.fr

Algorithms Project, Inria



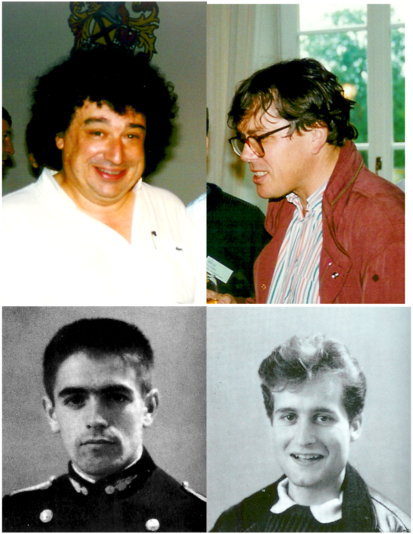
December 2, 2008
New: recent work and work in progress
with Carine Pivoteau and Michèle Soria

I Introduction

Old: ΛγΩ

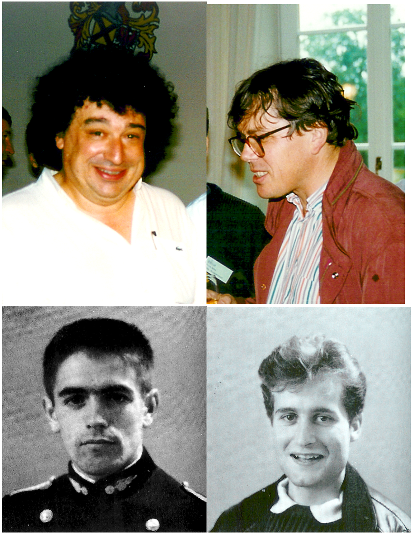
Origins: FlajoletSteyaert1981-82.
A complexity calculus for recursive tree algorithms

Old: ΛΥΩ



Origins: FlajoletSteyaert1981-82.
A complexity calculus for recursive tree algorithms

Old: ΛΥΩ



Origins: FlajoletSteyaert1981-82. *A complexity calculus for recursive tree algorithms*

The ΛΥΩ system:

Input

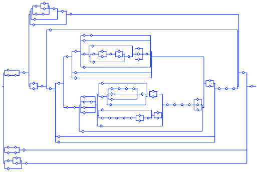
```
type bintree = epsilon | product(o, bintree, bintree);
     o = atom (1);

procedure pathlength (t : bintree);
begin
  case t of
    ()          : nil;
    o(u,v)      : begin size(u); size(v);
                        pathlength(u); pathlength(v);
                  end;
  end;
end;
```

Output

```
                               1/2   3/2
 av_tau_pathlength_n := Pi     n     + O(n)
```

General Framework

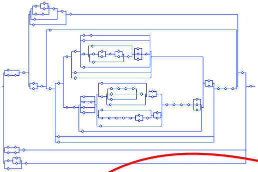
Combinatorial Objects ⟶ Equations over Generating Functions



$$Y = Y_1 + Y_2$$
$$Y_1 = (z + Y_2)^2/(1 - z - Y_2)$$
$$Y_2 = \exp(z + Y_1) - 1 - z - Y_1.$$

General Framework

Combinatorial Objects ⟶ Equations over Generating Functions

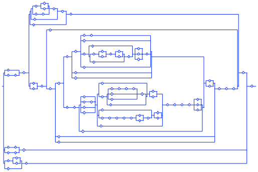


$$Y = Y_1 + Y_2$$
$$Y_1 = (z + Y_2)^2/(1 - z - Y_2)$$
$$Y_2 = \exp(z + Y_1) - 1 - z - Y_1.$$

Enumeration
3, 19, 195, 2791,. . .
→ Random generation by the recursive method [FlZiVCu94]

General Framework

Combinatorial Objects $\longrightarrow$ Equations over Generating Functions



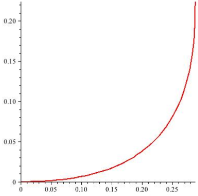
$$
\begin{aligned}
Y &= Y_1 + Y_2\\
Y_1 &= (z + Y_2)^2/(1 - z - Y_2)\\
Y_2 &= \exp(z + Y_1) - 1 - z - Y_1.
\end{aligned}
$$

Enumeration

3, 19, 195, 2791,...

$\rightarrow$ Random generation by the recursive method [FlZiVCu94]

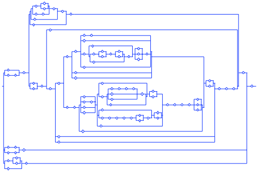
Numerical Values



$\rightarrow$ Boltzmann Sampling [DuFlLoSc04,FlFuPi07] (G. Schaeffer yesterday)

General Framework

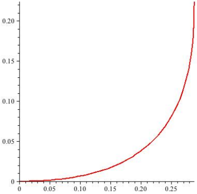
Combinatorial Objects ⟶ Equations over Generating Functions



$$Y = Y_1 + Y_2$$
$$Y_1 = (z + Y_2)^2/(1 - z - Y_2)$$
$$Y_2 = \exp(z + Y_1) - 1 - z - Y_1.$$

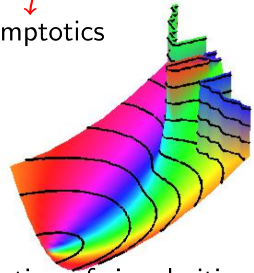
Enumeration
3, 19, 195, 2791,. . .
→ Random generation by the recursive method [FlZiVCu94]

Numerical Values



→ Boltzmann Sampling [DuFlLoSc04,FlFuPi07] (G. Schaeffer yesterday)

Asymptotics



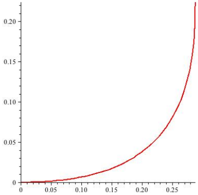
Location of singularities
Singularity analysis [FlajoletOdlyzko1990]

Computation

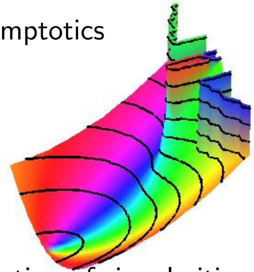
Enumeration

3, 19, 195, 2791,. . .

Numerical Values



Asymptotics



Location of singularities

Singularity analysis

Algorithms in $\Lambda\gamma\Omega$:

- Enumeration: $O(N^2)$;
- Location of singularities: for *explicit* g.f.;
- Local behaviour: *explicit* g.f., no Pólya operator.

$\longrightarrow$ Numerical values from the truncated g.f.

Computation

Algorithms in $\Lambda\Upsilon\Omega$:

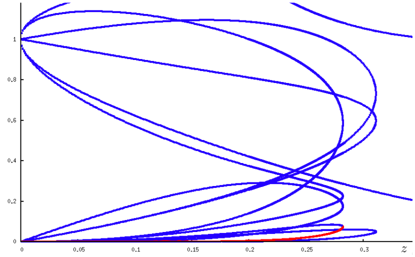
- Enumeration: $O(N^2)$;
- Location of singularities: for *explicit* g.f.;
- Local behaviour: *explicit* g.f., no Pólya operator.

$\longrightarrow$ Numerical values from the truncated g.f.

New

- Enumeration: $O(\mathrm{M}(N))$;
- Numerical evaluation: fast;
- Location of singularities: directly from a system .

$\mathrm{M}(N)$: cost of polynomial multiplication.



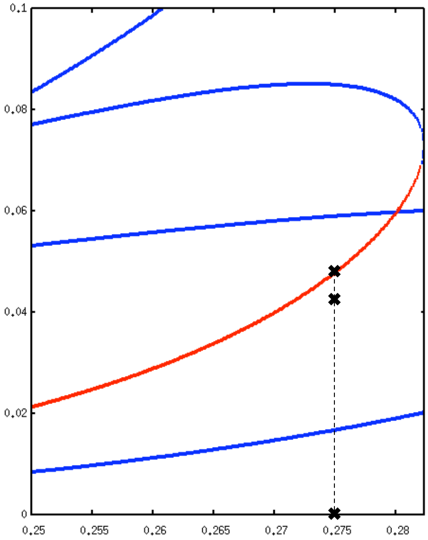
Main Result of (Pivoteau, Salvy, Soria 2008)

Theorem

For all (well-founded) systems of generating series, Newton iteration starting from 0 converges to the desired solution.

↓

Oracle for Boltzmann sampling.



Examples (I): Polynomial Systems (Darrasse 2008)

Random generation following given XML grammars.

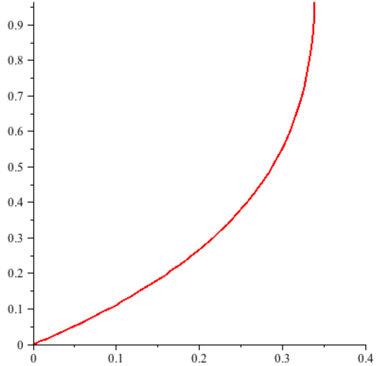
Grammar	nb eqs	max deg	nb sols	oracle (s.)	FGb (s.)
rss	10	5	2	0.02	0.03
PNML	22	4	4	0.05	0.1
xslt	40	3	10	0.4	1.5
relaxng	34	4	32	0.4	3.3
xhtml-basic	53	3	13	1.2	18
mathml2	182	2	18	3.7	882
xhtml	93	6	56	3.4	1124
xhtml-strict	80	6	32	3.0	1590
xmlschema	59	10	24	0.5	6592
SVG	117	10		5.8	>1.5Go
docbook	407	11		67.7	>1.5Go
OpenDoc	500			3.9	

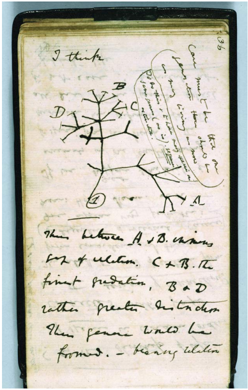
We exploit the combinatorial structure

Example (II): A Non-Polynomial “System”

Unlabelled rooted trees:

$$f(x) = x\exp\left(f(x) + \frac{1}{2}f(x^2) + \frac{1}{3}f(x^3) + \cdots\right)$$





II Combinatorics

Combinatorial Specifications

Language and Gen. Fcns (labelled)

$\mathcal{A} \cup \mathcal{B}$	$A(z) + B(z)$
$\mathcal{A} \times \mathcal{B}$	$A(z) \times B(z)$
$\mathcal{A}'$	$A'(z)$
$\mathrm{SEQ}(\mathcal{C})$	$\frac{1}{1-C(z)}$
$\mathrm{CYC}(\mathcal{C})$	$\log \frac{1}{1-C(z)}$
$\mathrm{SET}(\mathcal{C})$	$\exp(C(z))$

Examples

- Binary trees: $\mathcal{B} = \mathcal{Z} \cup \mathcal{Z} \times \mathcal{B} \times \mathcal{B}$;

$$B(z) = z + zB(z)^2$$

- General trees: $\mathcal{T} = \mathcal{Z} \times \mathrm{SET}(\mathcal{T})$;

$$T(z) = z\exp(T(z))$$

Combinatorial Specifications

Language and Gen. Fcns (labelled)

$\mathcal{A}\cup\mathcal{B}$	$A(z)+B(z)$
$\mathcal{A}\times\mathcal{B}$	$A(z)\times B(z)$
$\mathcal{A}'$	$A'(z)$
$\mathrm{SEQ}(\mathcal{C})$	$\frac{1}{1-C(z)}$
$\mathrm{CYC}(\mathcal{C})$	$\log\frac{1}{1-C(z)}$
$\mathrm{SET}(\mathcal{C})$	$\exp(C(z))$

Examples

- Binary trees: $\mathcal{B}=\mathcal{Z}\cup\mathcal{Z}\times\mathcal{B}\times\mathcal{B}$;

$$B(z)=z+zB(z)^2$$

- General trees: $\mathcal{T}=\mathcal{Z}\times\mathrm{SET}(\mathcal{T})$;

$$T(z)=z\exp(T(z))$$

System (Σ): $\boldsymbol{\mathcal{Y}}=\boldsymbol{\mathcal{H}}(\mathcal{Z},\boldsymbol{\mathcal{Y}})$

Definition (Well-founded)

$\boldsymbol{\mathcal{H}}(\varnothing,\boldsymbol{\varnothing})=\boldsymbol{\varnothing}$ and Jacobian $\partial\boldsymbol{\mathcal{H}}/\partial\boldsymbol{\mathcal{Y}}$ nilpotent at $(\varnothing,\boldsymbol{\varnothing})$.

Proposition

If (Σ) well-founded,

- $\boldsymbol{\mathcal{Y}}_{n+1}=\boldsymbol{\mathcal{H}}(\mathcal{Z},\boldsymbol{\mathcal{Y}}_n)$ *converges to a limit species* $\boldsymbol{\mathcal{Y}}$ *[Joyal1981]*
- $\mathbf{Y}(z)$ *is analytic in a neighborhood of 0.*

Combinatorial Specifications

Language and Gen. Fcns (unlabelled)

$\mathcal{A}\cup\mathcal{B}$	$A(z)+B(z)$
$\mathcal{A}\times\mathcal{B}$	$A(z)\times B(z)$
$\mathcal{A}'$	$A'(z)$
$\mathrm{SEQ}(\mathcal{C})$	$\frac{1}{1-C(z)}$
$\mathrm{PSET}(\mathrm{C})$	$\exp(\sum(-1)^i C(z^i)/i)$
$\mathrm{MSET}(\mathrm{C})$	$\exp(\sum C(z^i)/i)$
$\mathrm{CYC}(\mathrm{C})$	$\sum_{k\geq 1}\frac{\phi(k)}{k}\log\frac{1}{1-C(z^k)}$

Examples

- Binary trees: $\mathcal{B}=\mathcal{Z}\cup\mathcal{Z}\times\mathcal{B}\times\mathcal{B}$;

$$B(z)=z+zB(z)^2$$

- General trees: $\mathcal{T}=\mathcal{Z}\times\mathrm{SET}(\mathcal{T})$;

$$T(z)=z\exp(T(z)+T(z^2)/2+\cdots)$$

System (Σ): $\boldsymbol{\mathcal{Y}}=\boldsymbol{\mathcal{H}}(\mathcal{Z},\boldsymbol{\mathcal{Y}})$

Definition (Well-founded)

$\boldsymbol{\mathcal{H}}(\varnothing,\boldsymbol{\varnothing})=\boldsymbol{\varnothing}$ and Jacobian $\partial\boldsymbol{\mathcal{H}}/\partial\boldsymbol{\mathcal{Y}}$ nilpotent at $(\varnothing,\boldsymbol{\varnothing})$.

Proposition

If (Σ) well-founded,

- $\boldsymbol{\mathcal{Y}}_{n+1}=\boldsymbol{\mathcal{H}}(\mathcal{Z},\boldsymbol{\mathcal{Y}}_n)$ *converges to a limit species* $\boldsymbol{\mathcal{Y}}$ *[Joyal1981]*
- $\mathbf{Y}(z)$ *is analytic in a neighborhood of 0.*

$\Lambda\Upsilon\Omega$ via attribute grammars

- Attribute grammars: Knuth1968;
- Parameters of trees: DelestFedou1992;
- Decomposable structures: Mishna2003.

Specification + attribute grammar → Multivariate g.f.

```
tree:={B=Union(Epsilon,Prod(Z,B,B))}:
path_length:={pl(B)=Union(0,Prod(0,size(B)+pl(B),size(B)+pl(B)))}:
combstruct[agfeqns](tree,path_length,unlabelled,z,[[u,pl]]);
```

$$\left[B(z,u)=1+z\,B(z\,u,u)^2,\ Z(z,u)=z\,u\right]$$

- Generalizes $\Lambda\Upsilon\Omega$;
- All moments available by differentiating wrt u;
- Opens the way to the automatic computation of limit laws (e.g., with [DrmotaGittenbergerKlausner2005]).

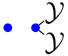
Combinatorial Newton (Décoste, Labelle, Leroux 1982)

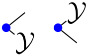
For one equation: $\mathcal{Y} = \mathcal{H}(\mathcal{Z}, \mathcal{Y}) \rightarrow \mathcal{Y}_{n+1} = \mathcal{N}_{\mathcal{H}}(\mathcal{Y}_n)$

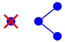
$$\mathcal{N}_{\mathcal{H}}(\mathcal{Y}) = \mathcal{Y} \cup \operatorname{SEQ}\left(\frac{\partial \mathcal{H}}{\partial \mathcal{Y}}(\mathcal{Z}, \mathcal{Y})\right) \times (\mathcal{H}(\mathcal{Z}, \mathcal{Y}) - \mathcal{Y}).$$

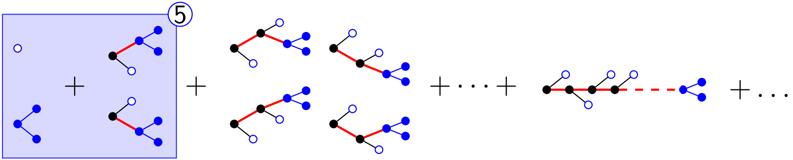
Contact doubles at each step.

Binary trees:

$\mathcal{H}(\mathcal{Z}, \mathcal{Y}) = \mathcal{Z} \cup \mathcal{Z} \times \mathcal{Y} \times \mathcal{Y} =$ 

$\frac{\partial \mathcal{H}}{\partial \mathcal{Y}}(\mathcal{Z}, \mathcal{Y}) =$ 

$\mathcal{Y}_0 = \varnothing$ $\quad \mathcal{Y}_1 = \bullet$ $\quad \mathcal{H}(\mathcal{Y}_1) - \mathcal{Y}_1 = \mathcal{Z} + \mathcal{Z}\mathcal{Y}_1^2 - \mathcal{Y}_1 =$ 

$\mathcal{Y}_2 =$ 

Combinatorial Newton (Décoste, Labelle, Leroux 1982)

For one equation: $\mathcal{Y} = \mathcal{H}(\mathcal{Z}, \mathcal{Y}) \rightarrow \mathcal{Y}_{n+1} = \mathcal{N}_{\mathcal{H}}(\mathcal{Y}_n)$

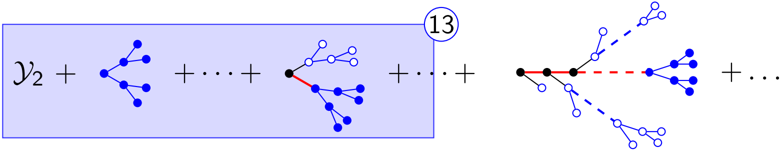
$$\mathcal{N}_{\mathcal{H}}(\mathcal{Y}) = \mathcal{Y} \cup \textsc{Seq}\left(\frac{\partial \mathcal{H}}{\partial \mathcal{Y}}(\mathcal{Z}, \mathcal{Y})\right) \times (\mathcal{H}(\mathcal{Z}, \mathcal{Y}) - \mathcal{Y}).$$

Binary trees:

$\mathcal{H}(\mathcal{Z}, \mathcal{Y}) = \mathcal{Z} \cup \mathcal{Z} \times \mathcal{Y} \times \mathcal{Y} =$ $\qquad$ $\frac{\partial \mathcal{H}}{\partial \mathcal{Y}}(\mathcal{Z}, \mathcal{Y}) =$

$\mathcal{Y}_0 = \varnothing \qquad \mathcal{Y}_1 = \bullet \qquad \mathcal{H}(\mathcal{Y}_1) - \mathcal{Y}_1 = \mathcal{Z} + \mathcal{Z}\mathcal{Y}_1^2 - \mathcal{Y}_1 =$

$\mathcal{Y}_3 =$



Combinatorial Newton (Décoste, Labelle, Leroux 1982)

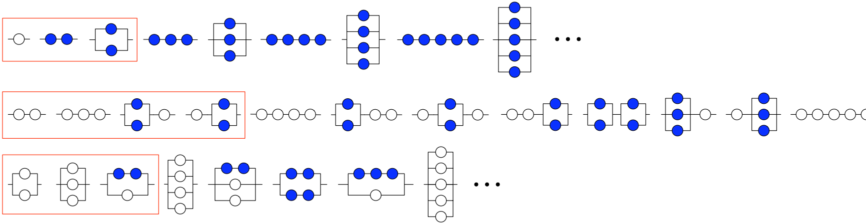
For one equation: $\mathcal{Y} = \mathcal{H}(\mathcal{Z}, \mathcal{Y}) \to \mathcal{Y}_{n+1} = \mathcal{N}_{\mathcal{H}}(\mathcal{Y}_n)$

$$\mathcal{N}_{\mathcal{H}}(\mathcal{Y}) = \mathcal{Y} \cup \textsc{Seq}\left(\frac{\partial \mathcal{H}}{\partial \mathcal{Y}}(\mathcal{Z}, \mathcal{Y})\right) \times (\mathcal{H}(\mathcal{Z}, \mathcal{Y}) - \mathcal{Y}).$$

Extension to systems (Pivoteau, Soria, Salvy 2008):

- $\textsc{Seq}(\frac{\partial \mathcal{H}}{\partial \mathcal{Y}})$ replaced by $\boldsymbol{\mathcal{U}} = \cup_{k\geq 0}(\frac{\partial \boldsymbol{\mathcal{H}}}{\partial \boldsymbol{\mathcal{Y}}})^k$ (Labelle's bloomings);

$\mathcal{Y}_1^{[2]}$, $\mathcal{Y}_2^{[2]}$, $\mathcal{Y}_3^{[2]}$



Combinatorial Newton (Décoste, Labelle, Leroux 1982)

For one equation: $\mathcal{Y} = \mathcal{H}(\mathcal{Z}, \mathcal{Y}) \rightarrow \mathcal{Y}_{n+1} = \mathcal{N}_{\mathcal{H}}(\mathcal{Y}_n)$

$$\mathcal{N}_{\mathcal{H}}(\mathcal{Y}) = \mathcal{Y} \cup \textsc{Seq}\left(\frac{\partial \mathcal{H}}{\partial \mathcal{Y}}(\mathcal{Z}, \mathcal{Y})\right) \times (\mathcal{H}(\mathcal{Z}, \mathcal{Y}) - \mathcal{Y}).$$

Extension to systems (Pivoteau, Soria, Salvy 2008):

- $\textsc{Seq}(\frac{\partial \mathcal{H}}{\partial \mathcal{Y}})$ replaced by $\boldsymbol{\mathcal{U}} = \cup_{k \geq 0} (\frac{\partial \boldsymbol{\mathcal{H}}}{\partial \boldsymbol{\mathcal{Y}}})^k$ (Labelle's bloomings);
- this union $\boldsymbol{\mathcal{U}}$ itself is computed by Newton iteration:

$$\begin{aligned} \boldsymbol{\mathcal{U}}_{n+1} &= \boldsymbol{\mathcal{U}}_n + \boldsymbol{\mathcal{U}}_n \boldsymbol{\mathcal{T}}_{n+1} \\ \boldsymbol{\mathcal{T}}_{n+1} &= \boldsymbol{\beta}_n \boldsymbol{\mathcal{U}}_n + \boldsymbol{\mathcal{T}}_n^2 \\ \boldsymbol{\beta}_n &= \frac{\partial \boldsymbol{\mathcal{H}}}{\partial \boldsymbol{\mathcal{Y}}}(\mathcal{Z}, \boldsymbol{\mathcal{Y}}_n) - \frac{\partial \boldsymbol{\mathcal{H}}}{\partial \boldsymbol{\mathcal{Y}}}(\mathcal{Z}, \boldsymbol{\mathcal{Y}}_{n-1}) \end{aligned}$$

At iteration $\boldsymbol{\mathcal{Y}}_n$, perform a single step of the calculation of $\boldsymbol{\mathcal{U}}$.

III Enumeration Sequences

Combinatorial Iteration → Iteration for Generating Series

- Same iteration in labelled & unlabelled universes:
  Example: General trees

$$\mathcal{Y} = \mathcal{H}(\mathcal{Y}) = \mathcal{Z} \times \textsc{Set}(\mathcal{Y}) \to \mathcal{N}_H(Y) = Y + \frac{H(Y) - Y}{1 - H(Y)}.$$

$$Y(z) = z \exp(Y(z) + Y(z^2)/2 + \cdots)$$

$$Y_0(z) = 0$$

$$Y_1(z) = \mathbf{z}$$

$$Y_2(z) = \mathbf{z} + \mathbf{z^2} + \mathbf{2z^3} + 3z^4 + 5z^5 + 7z^6 + 11z^7 + 15z^8 + 22z^9 + \cdots$$

$$Y_3(z) = \mathbf{z} + \mathbf{z^2} + \mathbf{2z^3} + \mathbf{4z^4} + \mathbf{9z^5} + \mathbf{20z^6} + \mathbf{48z^7} + 114z^8 + 280z^9$$

$$Y_4(z) = \mathbf{z} + \mathbf{z^2} + \mathbf{2z^3} + \mathbf{4z^4} + \mathbf{9z^5} + \mathbf{20z^6} + \mathbf{48z^7} + \mathbf{115z^8} + \mathbf{286z^9}$$

Combinatorial Iteration → Iteration for Generating Series

- Same iteration in labelled & unlabelled universes:
  Example: General trees

$$\mathcal{Y} = \mathcal{H}(\mathcal{Y}) = \mathcal{Z} \times \textsc{Set}(\mathcal{Y}) \to \mathcal{N}_H(Y) = Y + \frac{H(Y) - Y}{1 - H(Y)}.$$

- Small arithmetic complexity:
  first N coefficients in $O(\mathrm{M}(\mathrm{N}))$ operations
  (M(N): cost of multiplying two series).
  → fast enumeration,
  → fast random generation by the recursive method.

Combinatorial Iteration → Iteration for Generating Series

- Same iteration in labelled & unlabelled universes:
  Example: General trees

$$\mathcal{Y} = \mathcal{H}(\mathcal{Y}) = \mathcal{Z} \times \textsc{Set}(\mathcal{Y}) \to \mathcal{N}_H(Y) = Y + \frac{H(Y) - Y}{1 - H(Y)}.$$

- Small arithmetic complexity:
  first N coefficients in $O(\mathrm{M(N)})$ operations
  (M(N): cost of multiplying two series).
  → fast enumeration,
  → fast random generation by the recursive method.
- Small bit complexity:
  exact enumeration in quasi-optimal time.

IV Numerical values for Boltzmann Samplers

Numerical Values

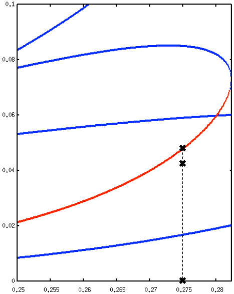
Combinatorics

- $\mathcal{Y} = \mathcal{H}(\mathcal{Z}, \mathcal{Y})$
- $\mathcal{Y}_{n+1} = \mathcal{N}_{\mathcal{H}}(\mathcal{Z}, \mathcal{Y}_n)$
- $\varnothing = \mathcal{Y}_0 \subset \mathcal{Y}_1 \subset \cdots \subset \mathcal{Y}$

$$\Downarrow$$

Generating Series

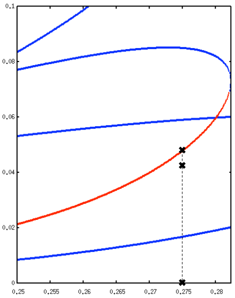
- $\mathbf{Y}(z) = \mathbf{H}(z, \mathbf{Y}(z))$
  (analytic in $|z| < r$ for some $r > 0$)
- $\mathbf{Y}_{n+1}(z) = \mathcal{N}_{\mathbf{H}}(z, \mathbf{Y}_n(z))$
- coeffs ≥ 0 and increasing $\Rightarrow$
  1. $\mathbf{Y}_n$ analytic in $|z| < r$;
  2. $\mathbf{Y}_n(x) \leq \mathbf{Y}_{n+1}(x)$ for $0 \leq x < r$;
  3. $\mathbf{Y}_n(z) \to \mathbf{Y}(z)$ for $|z| < r$.



Numerical Values

Generating Series

- $\mathbf{Y}(z) = \mathbf{H}(z, \mathbf{Y}(z))$
  (analytic in $|z| < r$ for some $r > 0$)
- $\mathbf{Y}_{n+1}(z) = \mathcal{N}_{\mathbf{H}}(z, \mathbf{Y}_n(z))$
- coeffs ≥ 0 and increasing $\Rightarrow$
  1. $\mathbf{Y}_n$ analytic in $|z| < r$;
  2. $\mathbf{Y}_n(x) \leq \mathbf{Y}_{n+1}(x)$ for $0 \leq x < r$;
  3. $\mathbf{Y}_n(z) \rightarrow \mathbf{Y}(z)$ for $|z| < r$.



$\Downarrow$

Proposition

For $0 \leq x < r$, *define*
$\alpha_0 := 0$ *and* $\boldsymbol{\alpha}_{n+1} := \mathcal{N}_{\mathbf{H}}(x, \boldsymbol{\alpha}_n)$.
Then, $\boldsymbol{\alpha}_n = \mathbf{Y}_n(x) \rightarrow \mathbf{Y}(x)$.

(Technical point: prove that **H** is analytic there.)

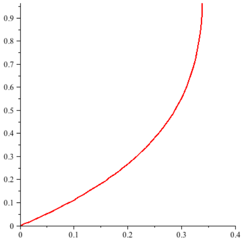
Example (II): Unlabelled Rooted Trees

- Combinatorial Newton iteration:

$$\mathcal{Y} = \mathcal{H}(\mathcal{Y}) = \mathcal{Z} \times \textsc{Set}(\mathcal{Y})$$

$$\rightarrow \mathcal{N}_H(Y) = Y + \frac{H(Y) - Y}{1 - H(Y)}.$$

$$H(Y) = z\exp(Y(z) + Y(z^2)/2 + \cdots).$$



- Numerical evaluation at $z = 0.3$:

n	$Y_n(0.3)$	$Y_n(0.3^2)$	$Y_n(0.3^3)$
0	0	0	0
1	.43021322639	**0.99**370806338e-1	**0.277**59817516e-1
2	**.5**4875612912	**0.998871**32154e-1	**0.277706291**87e-1
3	**.557**09557053	**0.998871471 9**7e-1	**0.27770629189e-1**
4	**.5571390**7945	**0.99887147198e-1**	**0.27770629189e-1**
5	**.55713908064**	**0.99887147198e-1**	**0.27770629189e-1**

V Conclusion

Conclusions

- More info at Carine Pivoteau's defense tomorrow at LIP6;
- Next: extend to $\mathcal{E}$'s (use multivariate g.f.);
- Next: use Schröder iteration at $x = r$ (exploit Drmota-Lalley-Woods);
- Next: singularity analysis of these systems (see also Bousquet-Mélou & Fusy).