Extended boxed product and application to synchronized trees

Olivier Bodini\textsuperscript{1}, Antoine Genitrini\textsuperscript{2} et Nicolas Rolin\textsuperscript{1}

\textsuperscript{1}Université Paris 13 – LIPN
\textsuperscript{2}UPMC Paris – LIP6

GASCom 2016
1 Introduction

2 Adaptating boxed product

3 Specification of synchronized trees

4 Uniform sampling
Increasing tree

Plane tree labeled with $n$ nodes:
- Each label between 1 and $n$ is taken
- Each path from the root to a leaf is a strictly increasing path

Shape of the tree

Increasing tree
Synchronized tree

Plane tree labeled with \( n + 1 \) nodes

- Each label between 1 and \( n \) is taken
- Each path from the root to a leaf is a strictly increasing path

⇒ two nodes are sharing a label.

Shape of the tree

Synchronized tree
A combinatorial class $C$ is a set of objects, with a size function, denoted by $|·| : C \to \mathbb{N}$ and such that for every integer $n$, the subset $C_n$ of objects of size $n$, is finite with cardinality $C_n$.

We define the exponential generating function of a combinatorial class $C$ to be:

$$C(z) = \sum_{n \geq 0} C_n \frac{z^n}{n!}$$
A *combinatorial class* \( C \) is a set of objects, with a size function, denoted by \(| \cdot | : C \to \mathbb{N}\) and such that for every integer \( n \), the subset \( C_n \) of objects of size \( n \), is finite with cardinality \( C_n \).

We define the *exponential generating function* of a combinatorial class \( C \) to be:

\[
C(z) = \sum_{n \geq 0} C_n \frac{z^n}{n!}
\]

**Small dictionary**

\[
\begin{align*}
C = A + B &\quad \to \quad C(z) = A(z) + B(z) \\
C = A \star B &\quad \to \quad C(z) = A(z) \cdot B(z)
\end{align*}
\]
The boxed product on the labeled classes:

\[ C = (\mathcal{A} \boxtimes \mathcal{B}), \]

Means that \( C \) is the product of \( \mathcal{A} \) and \( \mathcal{B} \), and that the element with the smallest label comes from the \( \mathcal{A} \) component. Its translation to equation is:

\[ C = \mathcal{A} \boxtimes \mathcal{B} \rightarrow C(z) = \int_{v=0}^{v=z} \frac{dA}{dv}(v)B(v)dv. \]
The increasing trees class $\mathcal{T}$ verifies the following specification:

$$\mathcal{T} = \mathcal{Z} \Box \ast \text{Seq}(\mathcal{T})$$

\[ T(z) = \frac{z}{1 - T(z)} \text{ or } T'(z) = 1 - T(z) \text{ and } T(0) = 0. \]
Specification of increasing trees

The increasing trees class $\mathcal{T}$ verifies the following specification:

$$\mathcal{T} = \mathcal{Z} \boxdot \star \text{Seq}(\mathcal{T})$$

Hence its generating function verifies:

$$T(z) = \int_{v=0}^{v=z} \frac{dv}{1 - T(v)} \quad \text{or} \quad T'(z) = \frac{1}{1 - T(z)} \quad \text{and} \quad T(0) = 0.$$ 

So:

$$T(z) = 1 - \sqrt{1 - 2z}$$
Adapting boxed product

Increasing trees with a synchronization

We can regroup the two identical nodes:
We can regroup the two identical nodes:
Let $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ combinatorial classes with no element of size 0.

\[
\mathcal{E} = \begin{array}{c}
\mathcal{C} \\
\mathcal{A} \\
\mathcal{B}
\end{array}
\]

can be expressed with boxed operator:

- $\mathcal{E} = \mathcal{C} \underset{\square}{\star} (\mathcal{A} \star \mathcal{B})$
- $E(z) = \int_{t=0}^{t=z} C'(t)A(t)B(t)dt$. 
Let $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ combinatorial classes with no element of size 0.

$E = \begin{array}{c} C \\ \downarrow \\ \mathcal{A} \end{array}$
$E = \begin{array}{c} C \\ \downarrow \\ \mathcal{B} \end{array}$

Can be expressed with boxed operator:

- $E = \mathcal{C} \Box * (\mathcal{A} \star \mathcal{B})$
- $E(z) = \int_{t=0}^{t=z} C'(t)A(t)B(t)dt.$

$D = \begin{array}{c} \mathcal{A} \\ \downarrow \\ \mathcal{C} \end{array}$
$D = \begin{array}{c} \mathcal{B} \\ \downarrow \\ \mathcal{C} \end{array}$

Can be expressed with boxed operator:

- $D = \mathcal{A} \Box * (\mathcal{B} \Box * \mathcal{C}) + \mathcal{B} \Box * (\mathcal{A} \Box * \mathcal{C})$
- $D(z) = \int_{x=0}^{z} \int_{y=0}^{x} A'(x)B'(y)C(y)dxdy + \int_{y=0}^{z} \int_{x=0}^{y} A'(y)B'(x)C(x)dxdy.$
We now look at the class $\mathcal{P}_{k,p}$. The previous method gives:

$$\mathcal{P}_{k,p} = A \boxtimes X_1 \boxtimes (Y_1 \boxtimes (\ldots) + X_2 \boxtimes (\ldots))$$

$$+ A \boxtimes Y_1 \boxtimes (X_1 \boxtimes (\ldots) + Y_2 \boxtimes (\ldots))$$

The two branches interlace.

$\Rightarrow$ The number of terms in the sum is $\binom{k+p}{k}$
Stanley showed that computing the number of linear extensions of a partial order reduces to compute the volume of convex polytopes.

\[
\#\{\text{linear extension of } \succ \} = 4! \int \int \int \int dz dt dx dy = 2
\]

Graph portraying the partial order $\succ$:

- $z \succ a, z \succ b, a \succ t, b \succ t$
Factorize the generating function:

\[ D(z) = \int_{x=0}^{z} \int_{y=0}^{x} A'(x) B'(y) C(y) \, dx \, dy + \int_{y=0}^{z} \int_{x=0}^{y} A'(x) B'(y) C(x) \, dx \, dy \]
Factorize the generating function:

\[
D(z) = \int_{x=0}^{z} \int_{y=0}^{x} A'(x) B'(y) C(y) \, dx \, dy + \int_{y=0}^{z} \int_{x=0}^{y} A'(x) B'(y) C(x) \, dx \, dy
\]

we swap the integration order of \( y \) and \( x \) on the right term

\[
= \int_{x=0}^{z} \int_{y=0}^{x} A'(x) B'(y) C(y) \, dx \, dy + \int_{x=0}^{z} \int_{y=x}^{z} A'(x) B'(y) C(x) \, dx \, dy
\]
Simple example: factorization

Factorize the generating function:

\[
D(z) = \int_{x=0}^{z} \int_{y=0}^{x} A'(x)B'(y)C(y)\,dx\,dy + \int_{y=0}^{z} \int_{x=0}^{y} A'(x)B'(y)C(x)\,dx\,dy
\]

we swap the integration order of \( y \) and \( x \) on the right term

\[
= \int_{x=0}^{z} \int_{y=0}^{x} A'(x)B'(y)C(y)\,dx\,dy + \int_{x=0}^{z} \int_{y=x}^{z} A'(x)B'(y)C(x)\,dx\,dy
\]

we replace \( y \) and \( x \) by \( \min(x, y) \) in \( C \)

\[
= \int_{x=0}^{z} \int_{y=0}^{x} A'(x)B'(y)C(\min(x, y))\,dx\,dy
\]
A more complicated example: factorization

\[
P_{k,p}(z) = \int_{u=0}^{z} \int_{x_1=0}^{u} \int_{y_1=0}^{x_1} \cdots \int_{x_p=0}^{y_{k-1}} \int_{y_k=0}^{\min(x_p,y_k)} \int_{t=0}^{u} A'(u) X'_1(x_1) X'_2(x_2) \cdots X'_p(x_p) Y'_1(y_1) Y'_2(y_2) \cdots Y'_k(y_k) R'(t) dudtdx_1 \cdots
\]
We change the integration order:

\[
D(z) = \int_{x=0}^{z} \int_{y=0}^{z} \int_{t=0}^{\min(x,y)} A'(x) B'(y) C'(t) \, dx \, dy \, dt
\]
Simple example: reordering

We change the integration order:

\[
D(z) = \int_{x=0}^{z} \int_{y=0}^{\min(x,z)} \int_{t=0}^{\min(x,y)} A'(x)B'(y)C'(t) \, dx \, dy \, dt
\]

\[
D(z) = \int_{x=0}^{z} \int_{t=0}^{\min(x,z)} \int_{y=0}^{t} A'(x)B'(y)C'(t) \, dx \, dy \, dt
\]
Simple example: reordering

We change the integration order:

\[ D(z) = \int_{x=0}^{z} \int_{y=0}^{z} \int_{t=0}^{\min(x,y)} A'(x)B'(y)C'(t) \, dx \, dy \, dt \]

\[ D(z) = \int_{x=0}^{z} \int_{t=0}^{\min(x,z)} \int_{y=t}^{z} A'(x)B'(y)C'(t) \, dx \, dy \, dt \]

\[ D(z) = \int_{t=0}^{z} \int_{x=t}^{z} \int_{y=t}^{z} A'(x)B'(y)C'(t) \, dx \, dy \, dt \]
We change the integration order:

\[
P_{k,p}(z) = \int_{u=0}^{z} \int_{x_1=0}^{u} \int_{x_2=0}^{t} \cdots \int_{x_p=0}^{t} \int_{y_1=0}^{t} \int_{y_2=0}^{t} \cdots \int_{y_k=0}^{t} \int_{t=0}^{\min(x_p,y_k)}
A'(u)X'_1(x_1)X'_2(x_2) \cdots X'_p(x_p)Y'_1(y_1)Y'_2(y_2) \cdots Y'_k(y_k)R'(t)dudtdx_1 \ldots
\]

\[
P(z) = \int_{u=0}^{z} \int_{t=0}^{u} \int_{x_1=t}^{u} \int_{x_2=t}^{x_1} \cdots \int_{x_p=t}^{x_1} \int_{y_1=t}^{x_p} \int_{y_2=t}^{y_1} \cdots \int_{y_k=t}^{y_{k-1}} \int_{t=0}^{y_k} \int_{t=0}^{x_1}
A'(u)X'_1(x_1)X'_2(x_2) \cdots X'_p(x_p)Y'_1(y_1)Y'_2(y_2) \cdots Y'_k(y_k)R'(t)dudtdx_1 \ldots
\]
Construction of synchronized trees

Construction:

- Shape in "pendulum"
Construction of synchronized trees

Construction:
- Shape in "pendulum"
- Increasing forests $\mathcal{F}$ grafted on it
\[ S = \mathcal{Z} \star \star (\mathcal{F} \star S \star \mathcal{F}) + \mathcal{P} \]

With \( \mathcal{F} \) the class of increasing forests.

\[ \mathcal{P} \]
Specification of synchronized trees

Construction of synchronized trees

\[ S(z) = \int_{t=0}^{z} \frac{\sqrt{1 - 2t}}{\sqrt{1 - 2z}} P'(t) \, dt \]

\[ P(z) = \int_{u=0}^{z} F^3(u) \int_{t=0}^{z} F^2(t) \left( \frac{\sqrt{1 - 2t}}{\sqrt{1 - 2z}} \right)^2 \, dt \, du \]

\[ = \frac{4 z^2 + (3 z - 1) \sqrt{-2 z + 1} - 4 z + 1}{3 (4 z^2 - 4 z + 1)} \]
Specification of synchronized trees

Application to synchronized trees

\[ S(z) = \frac{2z}{1 - 2z} - \log \left( \frac{1}{1 - 2z} \right) \]

\[ S(z) = 1 \frac{z^2}{2!} + 11 \frac{z^3}{3!} + 122 \frac{z^4}{4!} + 1518 \frac{z^5}{5!} + 21423 \frac{z^6}{6!} + \ldots \]

\[ S_n = 2^n \ n! \ \sqrt{\frac{n}{\pi}} \left( \frac{1}{2} - \frac{\log(n)}{4n} \right) \cdot \left( 1 + O \left( \frac{1}{\sqrt{n}} \right) \right). \]
1 Introduction

2 Adaptating boxed product

3 Specification of synchronized trees

4 Uniform sampling
Uniform sampling

Simple example: binary tree

\[ B(z) = z + zB(z)B(z) \]

We put a weight on the nodes of the left tree:
Uniform sampling

Simple example: binary tree

\[ B(z) = z + zB(z)B(z) \]

We put a weight on the nodes of the left tree:

\[ B(z, u) = z + zB(zu)B(z) \]

Left tree of size \( k \):

\[ P_{k,n} = \frac{[u^k z^n]B(z, u)}{[z^n]B(z, 1)} \]
Uniform sampling

More complex example: plane trees

\[ T(z) = \frac{1}{1 - T(z)} \]

We put a weight for each child of the root:
Uniform sampling

More complex example : plane trees

\[ T(z) = \frac{1}{1 - T(z)} \]

We put a weight for each child of the root:

\[ T(z, u) = \frac{1}{1 - u T(z)} \]

Root with \( k \) children:

\[ \mathbb{P}_{k, n} = \frac{[u^k z^n] T(z, u)}{[z^n] T(z, 1)} \]
More complex example: plane trees

\[ T(z) = z T(z)^5 \]
More complex example : plane trees

\[ T(z) = zT(z)^5 \]

We sample the size of each subtree:

\[ \text{Children from left to right of size respectively } p_1, p_2, p_3, p_4 \text{ and } p_5: \]

\[ \mathbb{P}_P = \left[ u_1^{p_1} u_2^{p_2} u_3^{p_3} u_4^{p_4} u_5^{p_5} z^n \right] T (z, u_1, u_2, u_3, u_4, u_5) \]

\[ \left[ z^n \right] T (z, 1, 1, \ldots, 1) \]
Synchronized trees

Add three variables: 

- $u$ for the size of the trunk 
- $b$ for the size of the left branch 
- $g$ for the size of the right branch
Add three variables:

- \( u \) for the size of the trunk
- \( b \) for the size of the left branch
- \( g \) for the size of the right branch
Trunk length

\[ S(z, u) = -\frac{((u - 3)z + 1)\sqrt{1 - 2z} (1 - 2z)^{\frac{1}{2}}u - 4z^2 + 4z - 1}{(4z^2 - 4z + 1)(u^2 - 4u + 3)(1 - 2z)^{\frac{1}{2}}u} \]

We sample a given length of the trunk \( m \) with probability

\[ \mathbb{P}_m = \frac{[u^m z^n] S(z, u)}{[z^n] S(z, 1)}. \]
Uniform sampling

Trunk length
Uniform sampling

\[
P(z, b, g) = \frac{\left( (b + g) - (b + g + 1)(1 - 2z)^{-\frac{1}{2}} + (1 - 2z)^{-\frac{b+g+1}{2}} \right)}{(b + g)(b + g + 1)}.
\]

We calculate \( S^{(m)}(z, b, g) \) using the parameter \( m \), and choose a left and right side length of respectively \( l \) and \( r \) with probability

\[
P_{l,r} = \frac{[b^l g^r z^n]S^{(m)}(z, b, g)}{[z^n]S^{(m)}(z, 1, 1)}.
\]
2(m + l + r + 1) + 1 instances of $\mathcal{F}$

To sample the shape we:

- use a variable for each instance of $\mathcal{F}$
- draw a size repartition
- sample each $\mathcal{F}$ independently
Label sampling

Uniform sampling
Uniform sampling

Label sampling

Diagram showing labeled trees with numbers 1, 2, 3, 4, and 5.
Uniform sampling

Label sampling

- mark each node whose label is greater than the synchronized label.
- \(2(l + r + m) + 1\) different variables, one for each string.
- A string of length \(n\) has for equation \(\frac{1-u_k^{n+1}}{1-u_k}z^n\)
Label sampling
Label sampling
Label sampling

Uniform sampling
Uniform sampling

Label sampling
Label sampling

Uniform sampling

---

Extended boxed product
We showed:

- a generalized boxed product that deals with any kind of partial order
- the exact generating function for synchronized trees
- a method to uniformly sample synchronized trees