
The Ordered and Colored Products in Analytic Combinatorics: Application to the Quantitative Study of Synchronizations in Concurrent Processes*

Olivier Bodini[†], Matthieu Dien[‡], Antoine Genitrini[‡] and Frédéric Peschanski[‡].

Olivier.Bodini@lipn.univ-paris13.fr and

{Matthieu.Dien, Antoine.Genitrini, Frederic.Peschanski}@lip6.fr.

July 12, 2017

In this paper, we study two operators for composing combinatorial classes: the *ordered product* and its dual, the *colored product*. These operators have a natural interpretation in terms of *Analytic Combinatorics*, in relation with combinations of Borel and Laplace transforms. Based on these new constructions, we exhibit a set of *transfer theorems* and closure properties. We also illustrate the use of these operators to specify increasingly labeled structures tightly related to Series-Parallel constructions and concurrent processes. In particular, we provide a quantitative analysis of Fork/Join (FJ) parallel processes, a particularly expressive example of such a class.

Keywords: Analytic Combinatorics, Concurrency Theory, Increasing Graph-Structures, Boxed Product.

1. Introduction

Our study is concerned with the increasing labeling of combinatorial structures, which tightly relate to the notion of concurrent processes. Our goal is to conduct this study using the tools of Analytic Combinatorics. In [4], we studied the restricted subclass of pure parallel processes, corresponding to the well-known increasing trees. One measure of particular importance is the number of possible executions of a process, which corresponds to the number of increasing labelings of the underlying combinatorial structure. More recently, we began the exploration of *interacting* processes, in particular through *synchronization*. The idea is to allow not only the *fork* of independent processes, as in the pure parallel case, but also their *join*, i.e. their ultimate synchronization. This class of *Fork-Join (FJ) processes* is a much more involved setting since labels can now be shared among distinct processes. This means we now face a class of increasing DAGs (directed acyclic graphs). As a starting point, we studied in [3], the restricted class of *diamond* processes. A diamond is a fork of two process that must be later joined, and their subprocesses are diamonds also. The main restriction is that diamonds may not be composed in series, which is a very strong constraint.

In terms of Analytic Combinatorics, the study of the pure parallel processes and the diamonds heavily relied on the *boxed product* operator (cf. [10] and [8, pp.139–142]). Unfortunately, the

*This research is partially supported by the ANR project *MetACOnc*, ANR-15-CE40-0014.

[†]Laboratoire d'Informatique de Paris-Nord, CNRS UMR 7030 - Institut Galilée - Université Paris-Nord, 99, avenue Jean-Baptiste Clément, 93430 Villetaneuse, France.

[‡]Laboratoire d'Informatique de Paris 6, CNRS UMR 7606 and Université Pierre et Marie Curie, 4 place Jussieu, 75005 Paris, France.

operator only singles out one node wrt. a sub-structure in the increasing labeling process. In order to put e.g. two diamonds in series, we have to combine two sub-structures A and B and ensure that all the labels of A are strictly lower than those of B . In the literature, the *ordered product* studied in the *species theory* [2] appears to match our requirement (it is also used as a technical tool in the context of Boltzmann sampling in [6]). However, as far as we know, the operator has not been studied in the realms of Analytic Combinatorics, a situation we intend to remedy with the present paper. The second operator we introduce is a form of a dual that we call the *colored product*, which has an interesting combinatorial interpretation, and to our knowledge is studied for the first time.

The outline of this paper is as follows. In Section 2, we define the ordered and colored product and provide the analogues of the standard iterated operators (SEQ, SET and CYC) based on the ordered product. We also provide transfer theorems that give the asymptotic behavior of the number of structures just by reading the way they are specified. Then we exhibit several combinatorial properties and examples, especially some classical structures whose specification can be rephrased through these new products. Next, in Section 3, we detail the construction of Fork-Join processes and their increasing labelings. We recall a result of Möhring and derive a formula to compute efficiently the number of runs of a given \mathcal{FJ} process. This corresponds to a so-called *hook-length formula*. We exhibit some technical difficulties related to \mathcal{FJ} processes through interesting subclasses, like the series combinations of the diamonds. Finally, some proof details are provided in dedicated Appendices.

2. The ordered and the colored operator

We briefly introduce some notations we will use through the paper. A combinatorial class is denoted by \mathcal{A} . Its number of objects of size n is given by A_n . Thus if \mathcal{A} is a non-labeled class, its ordinary generating function is

$$A(z) = \sum_{n \geq 0} A_n z^n.$$

And if \mathcal{A} is a labeled class, its exponential generating function is

$$A(z) = \sum_{n \geq 0} A_n \frac{z^n}{n!}.$$

2.1. Context and definitions

When studying tree-structured processes [4] we used the boxed product operator introduced by Greene [10] and later developed in the context of the symbolic method [8]. It is used to encode a global increasing labeling constraint for combinatorial classes. However the boxed product cannot be adapted to our context. In fact, the class $\mathcal{A}^\square \star \mathcal{B}$ given by the boxed product, contains well labeled objects from the product $\mathcal{A} \star \mathcal{B}$ with the constraint that the smallest label belongs to the component issued from \mathcal{A} .

In the case of our study of increasingly labeled structures, such a constraint is not sufficient. In particular, we need to build the combinatorial class of objects of $\mathcal{A} \star \mathcal{B}$ such that *all* the smallest labels belong to the component issued from \mathcal{A} . We will call the latter product the *ordered product*. We will introduce a second product, that may appear as less important in the context of labeled combinatorial class construction, but in fact, it is central in some quantitative analysis (e.g. Section 3). Both products are somehow dual.

In the literature an operator analogous to the ordered product, appears in the context of combinatorial species (cf. [2, Chapter 5]), named the *ordinal product*. However it has not been thoroughly studied, especially from a quantitative point of view. The name *ordered product* is introduced in the paper [6], where it is used as a technical tool in the context of Boltzmann sampling.

Definition 1. Let \mathcal{A} and \mathcal{B} be two labeled combinatorial classes and α and β be two structures respectively in \mathcal{A} and in \mathcal{B} . We define the class of labeled structures induced by α and β :

$$\alpha \boxtimes \beta = \{(\alpha, f_{|\alpha|}(\beta)) \mid f_{|\alpha|}(\cdot) \text{ shifts the labels} \\ \text{from } \beta \text{ by } |\alpha|\},$$

such that the function $f_{|\alpha|}$ is a relabeling function (by shifting by $+|\alpha|$ the previous labels). We extend the ordered product to combinatorial classes

$$\mathcal{A} \boxtimes \mathcal{B} = \bigcup_{\alpha \in \mathcal{A}, \beta \in \mathcal{B}} \alpha \boxtimes \beta.$$

In fact, the ordered product of $\mathcal{A} \boxtimes \mathcal{B}$ contains objects from the product $\mathcal{A} \star \mathcal{B}$ such that all the labels of component of \mathcal{A} are smaller than the ones of the component of \mathcal{B} .

Remark. In the case where one of the operands of the ordered product is reduced to an atom, then the product corresponds to the classical boxed product (e.g. [8, p.139]).

Definition 2. Let \mathcal{A} and \mathcal{B} be two unlabeled combinatorial classes and α and β be two structures respectively in \mathcal{A} and in \mathcal{B} . We define the class of unlabeled structures induced by α and β

$$\alpha \odot \beta = \{(\tilde{\alpha}, \tilde{\beta}) \text{ satisfying the conditions (C)}\},$$

(C) $\tilde{\alpha}$ and $\tilde{\beta}$ have respectively the structure of α and β and their atoms are 2-colored, with the constraint that among the $|\alpha| + |\beta|$ atoms exactly $|\alpha|$ atoms are colored with the first color (and thus $|\beta|$ atoms are colored with the second color).

We extend the colored product to combinatorial classes

$$\mathcal{A} \odot \mathcal{B} = \bigcup_{\alpha \in \mathcal{A}, \beta \in \mathcal{B}} \alpha \odot \beta.$$

Remark. In the case where one of the operands of the colored product is the atomic class, then the product corresponds to the classical pointing operator (e.g. [8, p.86]), in the following sense:

$$\mathcal{Z} \odot \mathcal{A} = \Theta(\mathcal{Z} \times \mathcal{A})$$

Let us now introduce the definitions of both products on the generating functions from two distinct points of view: a formal definition and an integral definition.

We first recall the classical integral transforms: the combinatorial Laplace and the Borel transforms¹. From a combinatorial point of view, they define a bridge between exponential generating functions and ordinary generating functions. More precisely, we have respectively

$$\mathcal{L}_c \left(\sum_{n \geq 0} a_n \frac{z^n}{n!} \right) = \sum_{n \geq 0} a_n z^n; \quad \mathcal{B}_c \left(\sum_{n \geq 0} a_n z^n \right) = \sum_{n \geq 0} a_n \frac{z^n}{n!}.$$

From a functional point of view, the combinatorial Laplace and the Borel transforms correspond respectively to

$$\mathcal{L}_c(f) = \int_0^\infty \exp(-t) f(zt) dt;$$

$$\mathcal{B}_c(f) = \frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{\exp(zt)}{t} f\left(\frac{1}{t}\right) dt,$$

where the real constant c is greater than the real part of all singularities of $f(1/t)/t$.

¹cf. Appendix A in which we recall the relations between the classical Laplace and Borel transforms and their combinatorial definitions.

Analogously to the traditional Laplace transform, the product of Laplace transforms can be expressed with a convolution product:

$$z \cdot \mathcal{L}_c(f) \cdot \mathcal{L}_c(g) = \mathcal{L}_c \left(\int_0^z f(t)g(z-t)dt \right).$$

Equivalently

$$\mathcal{L}_c(f) \cdot \mathcal{L}_c(g) = \mathcal{L}_c \left(\int_0^z f(t)g'(z-t)dt + g(0)f(z) \right).$$

We denote by $f * g$ the combinatorial convolution $\int_0^z f(t)g'(z-t)dt + g(0)f(z)$.

Proposition 3. *Let \mathcal{A} and \mathcal{B} be two labeled combinatorial classes. The exponential generating function $C(z)$, associated to $\mathcal{C} = \mathcal{A} \boxtimes \mathcal{B}$, satisfies the three following equations (according to the context: formal or integrable functions)*

$$\begin{aligned} C(z) &= \mathcal{B}_c(\mathcal{L}_c A(z) \cdot \mathcal{L}_c B(z)) \\ &= \sum_{n \geq 0} \frac{\sum_{k=0}^n a_k b_{n-k}}{n!} z^n \\ &= A(z) * B(z). \end{aligned}$$

Observe that the ordered product gives a combinatorial interpretation of this adapted convolution. Note that the integral interpretation is valid when both generating function $A(z)$ and $B(z)$ are integrable in their definition domain. However, for example if $A(z) = 1/(1-z)$, although $\mathcal{L}_c A(z)$ is not analytic, the function $A(z)$ can be a component of the ordered product. The proof of the result is given in Appendix A.

Proposition 4. *Let \mathcal{A} and \mathcal{B} be two unlabeled combinatorial classes. The ordinary generating function $C(z)$, associated to $\mathcal{C} = \mathcal{A} \odot \mathcal{B}$, satisfies the following equations*

$$C(z) = \mathcal{L}_c(\mathcal{B}_c A(z) \cdot \mathcal{B}_c B(z)) = \sum_{n \geq 0} \sum_{k=0}^n \binom{n}{k} a_k b_{n-k} z^n.$$

2.2. Some properties of the products

Let us first give some algebraic properties of the products.

Proposition 5. *The ordered and the colored product are associative and commutative. Furthermore, each product is distributive with the $+$ operator.*

We next introduce the differentiation of the ordered product.

Proposition 6. *Let \mathcal{A} and \mathcal{B} be labeled combinatorial classes. Then*

$$\begin{aligned} (A(z) * B(z))' &= A(z) * B'(z) + A'(z)B(0) \\ &= A'(z) * B(z) + B'(z)A(0). \end{aligned}$$

The interpretation of the previous result is direct. The series $(A(z) * B(z))'$ corresponds to the ordered product of \mathcal{A} and \mathcal{B} in which one node is removed. Obviously, since we know to which component belongs the remaining label, we know if the node has been removed from \mathcal{A} or from \mathcal{B} . We must keep in mind the specific case when one of the component has size 0.

And finally we provide two properties of the colored product.

Proposition 7. *Let \mathcal{A} and \mathcal{B} be unlabeled combinatorial classes. Then*

$$\begin{aligned} A(z) \odot B(z) &= A(0)B(0) \\ &+ z \left(\frac{A(z) - A(0)}{z} \odot B(z) + A(z) \odot \frac{B(z) - B(0)}{z} \right). \end{aligned}$$

This proposition suggests an interesting combinatorial interpretation. The series $\frac{A(z)-A(0)}{z} \odot B(z) + A(z) \odot \frac{B(z)-B(0)}{z}$ corresponds to the class of the colored product of \mathcal{A} and \mathcal{B} with one node removed. Due to the coloration of the remaining nodes, we can guess from which component (of \mathcal{A} or \mathcal{B}) the node has been removed, and also its color. The component $A(0)B(0)$ is to handle the special cases of size 0 structures. Ultimately, we get an object from the colored product of \mathcal{A} and \mathcal{B} .

Proposition 8. *Let \mathcal{A} be an unlabeled combinatorial class. The ordinary generating function $C(z)$ associated to the class $\mathcal{A} \odot \text{SEQ } \mathcal{Z}$ is similar to the classical binomial transform [11] of $A(z)$.*

$$C(z) = \sum_n \sum_{k=0}^n \binom{n}{k} a_k = \frac{1}{1-z} A\left(\frac{z}{1-z}\right).$$

The proofs of all these properties rely on basic properties of the Laplace and Borel transform, and are thus omitted.

2.3. Operators for iterated products

In this section we introduce the ordered analogue of the SEQ, SET and CYC operators. We first define the ordered exponentiation of a combinatorial labeled class: $\mathcal{A}^{\square k} = \underbrace{\mathcal{A} \boxtimes \dots \boxtimes \mathcal{A}}_{k \text{ times}}$. We take

the convention $\mathcal{A}^{\square 0} = \mathcal{E} = \{\epsilon\}$.

Then, an object of $\mathcal{A}^{\square k}$ is a k -tuple where the atoms of each component are labeled by all numbers of an interval of $\mathbb{N} \setminus \{0\}$, all intervals are pairwise disjoint and the union of all intervals is an interval starting by one. Moreover, the increasing labeling constraint states: for α a tuple in $\mathcal{A}^{\square k}$, then for all $i < j$ the labels of the component α_i are all smaller than the ones of α_j . Thus, α is a canonical representation of the set $\{\alpha_1, \dots, \alpha_k\}$. This observation naturally leads to the following definition.

Definition 9. *The ordered set $\text{SET}^{\square}(\mathcal{A})$ of a labeled combinatorial class \mathcal{A} (without structures of size 0) is defined by:*

$$\text{SET}^{\square}(\mathcal{A}) = \bigcup_{k \geq 0} \mathcal{A}^{\square k}.$$

This specification directly translates into a functional equation satisfied by the generating function:

$$\text{SET}^{\square}(A(z)) = \sum_{n \geq 0} A(z)^{\square n} = \mathcal{B}_c\left(\frac{1}{1 - \mathcal{L}_c(A(z))}\right).$$

In the same way, we define the ordered sequence that enumerates all permutations of a tuple $\alpha \in \mathcal{A}^{\square k}$.

Definition 10. *The ordered sequence $\text{SEQ}^{\square}(\mathcal{A})$ of a labeled combinatorial class \mathcal{A} (without structures of size 0) is*

$$\text{SEQ}^{\square}(\mathcal{A}) = \bigcup_{k \geq 0} \bigcup_{\sigma \in \mathcal{S}_k} \{\sigma(\alpha) \mid \forall \alpha \in \mathcal{A}^{\square k}\},$$

where \mathcal{S}_k denotes the set of permutations of size k . As usual, this combinatorial specification translates into the functional equations ²:

$$\text{SEQ}^{\square}(A(z)) = \sum_{k \geq 0} k! \cdot A(z)^{\square k} = \mathcal{B}_c(\mathcal{L}_c[\frac{1}{1-u}](\mathcal{L}_c(A(z)))).$$

It remains to define the ordered cycle. In this case, we need to distinguish the permutations of α up to cycle. Like in the case of the classical operator CYC for the labeled classes, a cycle is a sequence up to all of its circular shifts.

² The function $\mathcal{B}_c(\mathcal{L}_c[\frac{1}{1-u}](\mathcal{L}_c(A(z))))$ corresponds to the Laplace transform of the sequence operator applied to Laplace transform of A

Definition 11. The ordered cycle $\text{CYC}^{\square}(\mathcal{A})$ of a labeled combinatorial class \mathcal{A} (without structures of size 0) is:

$$\text{CYC}^{\square}(\mathcal{A}) = \bigcup_{k \geq 0} \bigcup_{\sigma \in \mathcal{S}_k \setminus \text{shift}_k} \{\sigma(\alpha) \mid \forall \alpha \in \mathcal{A}^{\square k}\},$$

where $\mathcal{S}_k \setminus \text{shift}_k$ is the set of size k permutations quotiented by the function $\text{shift}_k(i) = \begin{cases} 0 & \text{if } i = k \\ i + 1 & \text{otherwise} \end{cases}$.

Thus,

$$\begin{aligned} \text{CYC}^{\square}(A(z)) &= \sum_{k \geq 1} (k-1)! \cdot A(z)^{\square k} \\ &= \mathcal{B}_c \left(\mathcal{L}_c(A(z)) \cdot \sum_{k \geq 0} k! \mathcal{L}_c(A(z))^k \right) \\ &= A(z) * \text{SEQ}^{\square}(A(z)). \end{aligned}$$

Some examples will illustrate the use of these operators in the Section 2.6.

Iterated variants of the colored product are also worth studying, but this is left as a future work.

2.4. Closure Properties

Beyond being formal operators, the ordered and the colored product also enjoy important closure properties. Let us recall the definition of a holonomic generating function.

Definition 12. A generating function $F(z)$ is holonomic (or D-finite) if it satisfies a linear differential equation whose coefficients are rational polynomials in z :

$$q_0(z)F(z) + q_1(z)F^{(1)}(z) + \dots + q_r(z)F^{(r)}(z) = 0,$$

for some polynomials $q_s(X) \in \mathbb{Q}[X]$ (with q_r different from 0) and such that $F^{(i)}(z) = \frac{d^i}{dz^i} F(z)$.

Proposition 13. Let \mathcal{A} and \mathcal{B} be two labeled (resp. unlabeled) combinatorial classes whose exponential (resp. ordinary) generating functions are holonomic. Then the exponential generating function $A(z) * B(z)$ (resp. ordinary generating function $A(z) \odot B(z)$) is holonomic.

The proof of the proposition is obvious since both the Laplace and the Borel combinatorial transform are closed for the holonomy property. Let us now focus on a smaller set of functions: rational generating functions.

Definition 14. Let P and Q be two polynomials, with $Q \neq 0$, then $F = P/Q$ is a rational function.

Proposition 15. Let \mathcal{A} and \mathcal{B} be two unlabeled combinatorial classes which are associated to rational ordinary generating functions. Then the ordinary generating function $A(z) \odot B(z)$ is rational.

The proof relies on the existence of a partial fraction expansion for the rational functions and the following formula $\mathcal{L}_c\left(\frac{1}{1-az}\right) = e^{a \cdot z}$.

Remark that the analogous proposition for the ordered product of rational generating functions is not true. In fact, let $A(z) = z$ and $B(z) = 1/(1-z)$. Then we obtain $A(z) * B(z) = -\ln(1-z)$ that is not rational.

2.5. Transfer theorems

This section is dedicated to transfer theorems that give the asymptotic behavior of the number of objects built through the products and their iterations.

The following theorem, concerning the ordered product, is analogous to the transfer theorem proved in [1] in the context of rapidly growing power series.

Theorem 16. Let \mathcal{A} and \mathcal{B} be two combinatorial labeled classes. We denote by a (resp. b) the size of the smallest objects of \mathcal{A} (resp. \mathcal{B}). Let \mathcal{C} be the class $\mathcal{A} \boxtimes \mathcal{B}$.

If both exponential generating functions $A(z)$ and $B(z)$ have their dominant singularities in $[0, +\infty[$ and if there exists a positive integer r such that $\sum_{k=r}^{n-r} A_k B_{n-k} = \mathcal{O}(A_{n-r} + B_{n-r})$, then

$$C_n \underset{n \rightarrow \infty}{\sim} B_b A_{n-b} + A_a B_{n-a}.$$

Remark. In many cases $B_b A_{n-b}$ and $A_a B_{n-a}$ are not of the same order (e.g. both dominant singularities are distinct, or the sub-exponential factors in the asymptotic behaviors of A_n and B_n are distinct, or the constants a and b are distinct).

The proof of this result is given in Appendix B.

Note that if the exponential generating functions $A(z)$ and $B(z)$ (with dominant singularities ρ_A and ρ_B) are such that there exist two constants α and β such that for large enough n we get $A_n \leq n^\alpha \rho_A^{-n} n!$ and $B_n \leq n^\beta \rho_B^{-n} n!$, then the assumption $\sum_{k=r}^{n-r} A_k B_{n-k} = \mathcal{O}(A_{n-r} + B_{n-r})$ is satisfied.

Theorem 17. Let $A(z)$ be an exponential generating function (with $A_0 = 0$). We define $S(z) = \text{SET}^{\boxtimes}(A(z))$, $L(z) = \mathcal{L}_c(A(z))$ and ρ the radius of convergence of $L(z)$. Then, the asymptotic behavior of $[z^n]S(z)$, when n tends to infinity, is

1. $[z^n]S(z) \underset{n \rightarrow \infty}{\sim} [z^n]A(z)$, if $\rho = 0$ and $\frac{A_n}{A_{n-1}} = \Omega(n^\alpha)$, where α is a constant greater than 0.
2. $[z^n]S(z) = \frac{1}{n! \sigma L'(\sigma)} \cdot \sigma^n (1 + C^n)$ if $\rho \in]0, +\infty[$ and $L(\rho) > 1$, with C being a real constant such that $0 < C < 1$ and σ being the root of $L(\sigma) = 1$.

Proof. 1. In this case, we proceed by induction. Let define $S_k = 1 + A * S_{k-1}$ (and $S_1 = 1 + A$), which is the generating series of the ordered set containing at most k \mathcal{A} -structures. Thus, we have $[z^n]S_1 = [z^n]A$. Now, let us prove $[z^n]S_{k+1} \underset{n \rightarrow \infty}{\sim} [z^n]A$: We start from $S_{k+1} = 1 + A * S_k$, then $[z^n]S_{k+1} = [z^n]A * S_k$. Since $\frac{A_n}{A_{n-1}} = \Omega(n^\alpha)$, we use the Theorem 16 and get $[z^n]S_{k+1} \underset{n \rightarrow \infty}{\sim} A_a [z^{n-a}]S_k + [z^n]A$ where $a > 0$ is the valuation of A . So, using the induction hypothesis, we obtain $[z^n]S_{k+1} \underset{n \rightarrow \infty}{\sim} A_a [z^{n-a}]A + [z^n]A \underset{n \rightarrow \infty}{\sim} [z^n]A$. We thus get the conclusion by induction.

2. This case is a direct application of the theorem for *asymptotics of supercritical sequence* proved by Flajolet and Soria [9]. □

Remark. The first case of this theorem states that if the number of structures of a given class \mathcal{A} grows fast enough, then, asymptotically, an ordered set of structures of \mathcal{A} contains, almost surely, a single big structure.

2.6. Examples

We conclude this section by revisiting classical combinatorial classes using the ordered and colored products but specified with our products. Some new interesting structures are also presented.

Example 18. *Binary search trees.*

Let $B(z)$ be the exponential generating function related to the binary search trees (BSTs) containing distinct keys from 1 to the size of the tree. Using the ordered product, we can easily specify BSTs

$$\mathcal{B} = \mathcal{E} + (\mathcal{B} \boxtimes \mathcal{Z}) \boxtimes \mathcal{B}.$$

By the symbolic method, we get

$$B(z) = 1 + (B(z) * z) * B(z).$$

There is one tree containing zero key and otherwise the left sub-tree contains keys that are smaller than the root-key, and all of them are smaller than the keys from the right sub-tree. By using the definition of the ordered product, we prove that the number of BSTs built on the n distinct keys is equal to the n -th Catalan number. More precisely, we get back the classical recursive equation satisfied by the Catalan numbers.

Example 19. *Integer compositions.*

An integer composition of an integer n is a sequence of strictly positive integers whose sum is n . Usually, the specification $\text{SEQ}(\text{SEQ}_{\geq 1}(\mathcal{Z}))$ characterizes the class, where the inner $\text{SEQ}_{\geq 1}(\mathcal{Z})$ corresponds to the class of strictly positive integers. Even if this description leads to ordinary or exponential generating functions, its way of thinking is clearly an unlabeled one: the main idea is to represent an integer by a sequence of atoms without matter about their labels which gives $n!$ objects representing the integer n ($n! [z^n]_{\frac{1}{1-z}}$).

Using the ordered set, we obtain another way to express it in a more labeled fashion. The idea is to encode an integer n by a set of labeled atoms of size n . Thus, one object represents one integer ($n! [z^n] e^z = 1$). Then, we represent the integer compositions by the ordered sets of labeled integers: $\text{SET}^{\square}(\text{SET}_{\geq 1}(\mathcal{Z}))$. Each integer (i.e. $\text{SET}_{\geq 1}(\mathcal{Z})$) in the ordered set represents a part of the composition and the interval labeling this integer represents its position in the composition. For example, the set $\{\{1, 2, 3\}, \{6, 5\}, \{7, 8, 9\}, \{4\}\}$ represents the composition $3 + 1 + 2 + 3$ of 9.

Then, applying the symbolic method, we obtain the exponential generating function

$$\mathcal{B}_c \left(\frac{1}{1 - \mathcal{L}_c(e^z - 1)} \right) = \mathcal{B}_c \left(\frac{1 - z}{1 - 2z} \right)$$

Although we can compute the general term (which is 2^{n-1}) of the series directly, note that this generating function corresponds to the second case of Theorem 17. Thus, the expected result (2^{n-1}) is obtained directly.

Example 20. *Permutations with interval cycles.*

It is well known that a permutation can be decomposed as a set of cycles, which gives the following equation in term of the symbolic method: $\text{SET}(\text{CYC}(\mathcal{Z}))$.

Now, let us define the permutations subclass in which every cycles contain labeled atoms (i.e. integers) where the labels are all of the same integer interval. For example (231489567) is a such permutation, it can be decomposed as $\{(\overline{123})(\overline{4})(\overline{58697})\}$ (where $c = \overline{a_0 \dots a_{n-1}}$ is such that $c(a_i) = a_{i+1 \pmod n}$). So, these permutations with interval cycles (we will denote their combinatorial class by \mathcal{P}) can be encoded with the symbolic method using the SET^{\square} operator:

$$\mathcal{P} = \text{SET}^{\square}(\text{CYC}(\mathcal{Z}))$$

Then, we obtain the generating functions P given by $P(z) = \mathcal{B}_c \left(\frac{1}{1 - \mathcal{L}_c \left(\log \left(\frac{1}{1-z} \right) \right)} \right)$ or by elementary

combinatorial argument $P(z) = \sum_{n \geq 0} \left(\sum_{c \in C_n} \binom{n}{c} \right) \frac{z^n}{n!}$ where C_n is the set of integer compositions of n (see as tuples) and $\binom{n}{c}$ is a multinomial coefficient.

Then, noticing that $\mathcal{L}_c \left(\log \frac{1}{1-z} \right)$ has a zero convergence radius, we apply our transfer Theorem 17 and directly obtain $[z^n]P(z) \sim (n-1)!$.

Example 21. *Bell numbers.*

The Bell numbers, denoted B_n , are known to count the number of partitions of a set. They verify the recurrence equation $B_n = \sum_{k=0}^{n-1} \binom{n-1}{k} B_k$ (with $B_0 = 1$) and the labeled combinatorial class of partitions of sets is specified by $\mathcal{B} = \text{SET}(\text{SET}_{\geq 1} \mathcal{Z})$. A partition of a set is a set of sets such that the sum of the sizes of the inner sets is equal to the size of the sliced set. From this specification, by applying the symbolic method, we derive the formula of their exponential generating function: $B(z) = \sum_{n \geq 0} B_n \frac{z^n}{n!} = e^{e^z - 1}$.

We now propose to revisit the Bell numbers but in the unlabeled context. This provides an interesting use case for the colored product.

In the unlabeled world, a partition of a set can be seen as follows: let n be aligned points on a line (in the plane), then a partition of a set of n elements in k parts is isomorphic to the k -coloring of the n points. So, a partition of a set of n elements is a part of size p containing the first point (a coloring of p points with the color 1), and an other partition of a set of size $n - p$ (a coloring of the remaining points with $k - 1$ colors).

Translating this decomposition, using the symbolic method and the colored product, we obtain the following unlabeled specification:

$$\mathcal{B} = \mathcal{E} + \mathcal{Z} \times (\text{SEQ}(\mathcal{Z}) \odot \mathcal{B}).$$

The colored product is here used to make the choice of which atoms belong to the first part of the partitions and which belong to the sub-partition.

Thus, we obtain the following functional equation for the ordinary generating series of Bell numbers:

$$B(z) = 1 + z \cdot \mathcal{L}_c \left(\mathcal{B}_c(B(z)) \cdot \mathcal{B}_c \left(\frac{1}{1-z} \right) \right).$$

Applying the Borel transformation on this equation, and after using one of the identities given in Appendix A, we obtain:

$$\mathcal{B}_c(B(z))' = \mathcal{B}_c(B(z)) e^z.$$

Solving this differential equation leads to the expected result:

$$\mathcal{B}_c(B(z)) = e^{e^z - 1}.$$

3. Application: fork-join processes

We now illustrate the use of the ordered and colored products for enumerating the increasing labelings of series parallel structures. The motivation comes from concurrency theory where the control graph of concurrent processes can be seen as directed acyclic graphs (DAG). Each run (or execution) of a given concurrent program then corresponds to an increasing labeling of the DAG.

Definition 22. An increasing labeling of a DAG D is a labeling of the n nodes of D where the labels are distinct, between 1 and n , and such that the sequence of labels induced by a path of D is strictly increasing.

In the figures, the direction of the edges of the DAGs are assumed from top to bottom. See Figure 3 for an example of an increasing DAG.

We do not know how to deal with the increasing labellings of arbitrary DAGs. The counting problem already is known to be \sharp P-complete [5]. Our objective is thus to find and study interesting and more tractable subclasses. In [4] we provide a complete analytic study of the asymptotic average number of increasing labelings for trees as well as a linear time counting algorithm. In [3] we study the *diamond* processes. A diamond is a fork of two processes that must be later joined, and the subprocesses are diamonds also. The main restriction is that diamonds may not be composed in series, which is a very strong constraint.

In the present paper, we study the more expressive class of fork-join processes, defined as follows.

Definition 23. The DAG associated to a \mathcal{FJ} process is structured according to the specification depicted in Figure 1. Thus an unambiguous specification is

$$\mathcal{FJ} = \bullet \mid \bullet \times \mathcal{FJ} \mid \bullet \times (\mathcal{FJ} \times \mathcal{FJ}) \times \mathcal{FJ}.$$

This class is named after the way Unix processes work: A parent process can *fork* a child sub-process, and wait for the termination of the child, which is named a *join* operation.

There is an elegant isomorphism between the unlabeled structure of \mathcal{FJ} processes and a simple variety of trees with nodes of arity 0, 1 or 3. An example is such a correspondance is given in Figure 2. In the ternary nodes, the rightmost sub-tree is the sub-process rooted at the *join node* of

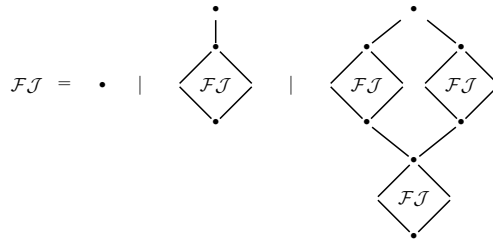


Figure 1: The \mathcal{FJ} processes.

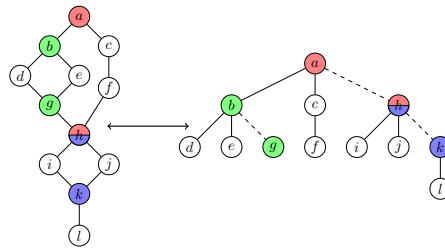


Figure 2: \mathcal{FJ} processes as ternary trees.

the DAG, and the two leftmost sub-trees correspond to the sub-processes in parallel. This encoding provides a natural induction principle for \mathcal{FJ} processes.

Similarly to the *hook-length formula* for trees (cf. [11, p. 67]), we can get an efficient mean of counting increasing labelings in \mathcal{FJ} processes. The formula is based on the ternary tree correspondence of a \mathcal{FJ} process P . For each node ν in P , we define P_ν to be the fringe subtree rooted in ν (and containing all descendants of ν).

For each ternary node ν , whose children are the nodes ν_1, ν_2 and ν_3 (from left to right), we get the three fringe subtrees P_{ν_1}, P_{ν_2} and P_{ν_3} . For example, for the node h in Figure 2, the fringe subtree P_{h_1} is reduced to i and the fringe subtree P_{h_3} contains k and l .

Theorem 24. (Hook-length formula for \mathcal{FJ} processes). *The number of runs of a \mathcal{FJ} process P (i.e. the number of increasing labelings of the DAG of P) is*

$$|\mathcal{L}^{\mathcal{E}}(P)| = \prod_{\substack{\nu \text{ ternary} \\ \text{node}}} \frac{|P_{\nu_1} \cup P_{\nu_2}|!}{|P_{\nu_1}|! \cdot |P_{\nu_2}|!}.$$

Furthermore, by denoting by n the size of P , then by memoizing all factorial values of the integers from 1 to n , the number $|\mathcal{L}^{\mathcal{E}}(P)|$ is computed in $\Theta(n)$ arithmetic operations in the worst case and with a single traversal of P .

An application of the formula for our example in Figure 2 gives $(2! 6! 2!) / (1! 1! 4! 2! 1! 1!) = 2880/48 = 60$. Thus there are 60 different runs induced by the example process.

For a process P , if its tree correspondence contains no ternary node, i.e. P is a chain, thus we put $|\mathcal{L}^{\mathcal{E}}(P)| = 1$. Theorem 24 is derived directly from Möhring's formula for series-parallel Posets [12] (the connection between FJ processes and SP-posets is obvious).

A first step to understanding what a typical \mathcal{FJ} processes looks like is computing the number of processes of a given size. Let $F(z)$ be the ordinary generating function enumerating \mathcal{FJ} processes. We denote by F_n the number of \mathcal{FJ} processes of size n . Thus we get $F(z) = \sum_n F_n z^n$. The first coefficients $(F_n)_{n>0}$ are

$$1, 1, 1, 2, 5, 11, 24, 57, 141, 349, 871, 2212, 5688, 14730, \\ 38403, 100829, 266333, 706997, 1885165, 5047522, \dots$$

This sequence is stored in OEIS³ A071879 and enumerates the subclass of $(0, 1, 3)$ -trees according to the number of edges instead of nodes. The bijection is direct.

Proposition 25. *Asymptotically when n tends to infinity, the number of \mathcal{FJ} processes of size n is*

$$F_n \underset{n \rightarrow \infty}{\sim} \frac{1}{2} \sqrt{\frac{1}{3} + \frac{1}{2^{2/3}}} \frac{\rho^{-n}}{\sqrt{\pi n^3}},$$

with the dominant singularity $\rho = (1 + 3 \cdot 2^{-2/3})^{-1}$.

The latter Proposition 25 is directly derived from the enumeration of simple varieties of trees based on the symbolic method. This method is well introduced in the book of Flajolet and Sedgewick [8]. However we give a complete proof in Appendix A.

Once this first measure computed, our aim is to compute the average number of runs of a \mathcal{FJ} process. We would like to follow the strategy from [4], however a major barrier appears immediately. The classical operators defined in the symbolic method are not sufficiently expressive to apply to \mathcal{FJ} processes.

That is the point, where the ordered product appears to express the constraints of the increasing labelings of \mathcal{FJ} processes DAG.

From the structure of \mathcal{FJ} processes (specified in Definition 23) and the definition of the ordered product, we deduce immediately the specification of \mathcal{FJ} processes that are increasingly labeled.

Definition 26. *Let \mathcal{P} be the labeled class of increasing \mathcal{FJ} processes. Then*

$$\mathcal{P} = \mathcal{Z} + \mathcal{Z} \square \star \mathcal{P} + \mathcal{Z} \square \star ((\mathcal{P} \star \mathcal{P}) \boxtimes \mathcal{P}).$$

First, note that the classical boxed product $\square \star$ could be replaced by the ordered product \boxtimes . Furthermore, although we have proposed several tools to deal with the ordered product, the quantitative analysis of the class \mathcal{P} seems very difficult: It relies on a second order integral equation

$$P(z) = z + \int_0^z P(t) dt + \int_0^z \int_0^t P^2(u) P'(t-u) du dt.$$

Our approach to circumvent the problem is to introduce subclasses that get closer and closer to \mathcal{P} .

A first way is to bound the fork depth, that is the number of fork nodes on each path of the DAG. An alternative way is to bound the number of subprocesses that can be put in series. In fact, the limit case where no series operation was allowed corresponds to the diamonds [3].

3.1. Parallel-constrained processes

The specification of parallel-constrained \mathcal{FJ} processes is as follows

$$\begin{aligned} \mathcal{W}_0 &= \text{SEQ}_{\geq 1} \mathcal{Z}, \\ \mathcal{W}_\ell &= \mathcal{Z} + \mathcal{Z} \square \star \mathcal{W}_\ell \\ &\quad + \mathcal{Z} \square \star ((2 \cdot \mathcal{W}_{\ell-1} \star \text{SEQ} \mathcal{Z} - \mathcal{W}_0 \star \mathcal{W}_0) \boxtimes \mathcal{W}_\ell). \end{aligned}$$

These subclasses of \mathcal{FJ} processes are an increasing iteration of subclasses, in the sense that $\mathcal{W}_0 \subset \mathcal{W}_1 \subset \mathcal{W}_2 \subset \dots$. There are two major restrictions in front of \mathcal{FJ} processes, both of them are related to the fork nodes. The first one, in \mathcal{W}_ℓ , is such that we allow only ℓ nested fork nodes. The second one is such that two processes can be put in parallel only when at least one of them is a series of atoms (a wire of atoms), eventually empty.

Let us focus on the functional equation associated to \mathcal{W}_ℓ .

Proposition 27. *Let $\ell \in \mathbb{N}$. The class \mathcal{W}_ℓ is such that $\mathcal{L}_c(\mathcal{W}_\ell(z))$ is a rational function, i.e. $\mathcal{W}_\ell(z)$ is satisfying a linear differential equation.*

³Online Encyclopedia of Integer Sequences

Proof. First differentiate it according to z , then take the Laplace transform of the whole equation, and recognize a colored product. We get

$$\begin{aligned}\mathcal{L}_c(W_\ell(z)') &= 1 + \mathcal{L}_c(W_\ell(z)) \\ &+ \left(2\mathcal{L}_c(W_{\ell-1}(z)) \odot \mathcal{L}_c(\exp(z)) - \mathcal{L}_c(W_0(z)^2)\right) \cdot \mathcal{L}_c(W_\ell(z)).\end{aligned}$$

Substitute $\mathcal{L}_c(W'_\ell(z))$ by $z^{-1}\mathcal{L}_c(W_\ell(z))$ (cf. Appendix A), and finally isolate $\mathcal{L}_c(W_\ell(z))$:

$$\mathcal{L}_c(W_\ell(z)) = \frac{z}{1 - z - z(2\mathcal{L}_c(W_{\ell-1}(z)) \odot \mathcal{L}_c(\exp(z)) - \mathcal{L}_c(W_0(z)^2))}.$$

By induction we prove that $\mathcal{L}_c(W_\ell(z))$ is a rational function, since the colored product of two rational functions is rational (cf. Proposition 15). \square \square

Easily we compute

$$\mathcal{L}_c(W_0(z)) = \frac{z}{1-z}; \quad \mathcal{L}_c(W_1(z)) = \frac{z(1-2z)}{1-3z}.$$

Let us denote by $\frac{P_\ell}{Q_\ell}(z)$ the rational function $\mathcal{L}_c(W_\ell(z))$, and by d_ℓ the degree of $Q_\ell(z)$.

Proposition 28. *Let $\ell \geq 2$. The class \mathcal{W}_ℓ is such that, for*

$$\begin{aligned}P_\ell(z) &= z(1-z)(1-2z) \cdot (1-z)^{d_{\ell-1}} Q_{\ell-1}\left(\frac{z}{1-z}\right) \\ Q_\ell(z) &= (1-4z+5z^2) \cdot (1-z)^{d_{\ell-1}} Q_{\ell-1}\left(\frac{z}{1-z}\right) \\ &\quad - 2z(1-2z) \cdot (1-z)^{d_{\ell-1}} P_{\ell-1}\left(\frac{z}{1-z}\right),\end{aligned}$$

we get $\mathcal{L}_c(W_\ell(z)) = P_\ell(z)/Q_\ell(z)$. Furthermore the polynomial P_ℓ and Q_ℓ are coprime and the degree of Q_ℓ is $d_\ell = 3\ell - 2$.

Remark that both products $(1-z)^{d_{\ell-1}} Q_{\ell-1}\left(\frac{z}{1-z}\right)$ and $(1-z)^{d_{\ell-1}} P_{\ell-1}\left(\frac{z}{1-z}\right)$ are polynomials. The proposition is proved by recurrence, in particular P_ℓ and Q_ℓ cannot have a common root, otherwise $P_{\ell-1}$ and $Q_{\ell-1}$ would have one. Furthermore, for $\ell \geq 1$, we get that the differential equation satisfied by \mathcal{W}_ℓ is of order d_ℓ since the polynomial $Q_\ell(z)$ corresponds to the characteristic equation associated to the differential equation of \mathcal{W}_ℓ . Thus, then smallest root of Q_ℓ is the exponential order in the asymptotic behavior of $[z^n]\mathcal{W}_\ell(z)$. The sequence of dominant singularities, from \mathcal{W}_1 to \mathcal{W}_{20} is decreasing (due to the growth of the families \mathcal{W}_ℓ). It starts at $1/3$ and ends, for $\ell = 20$ around 0.044743 .

Proposition 29. *The average number of runs in \mathcal{W}_{20} processes, of size n , is asymptotically $\Theta(r^n)$, with $r \approx 5.4314$.*

A simple improvement of the previous specification avoids the empty wire as a parallel sub-process. Thus, the latter specification is replaced by

$$\begin{aligned}\tilde{\mathcal{W}}_0 &= \text{SEQ}_{\geq 1} \mathcal{Z}, \\ \tilde{\mathcal{W}}_\ell &= \mathcal{Z} + \mathcal{Z}^\square \star \tilde{\mathcal{W}}_\ell \\ &\quad + \mathcal{Z}^\square \star ((2 \cdot \tilde{\mathcal{W}}_{\ell-1} \star \tilde{\mathcal{W}}_0 - \tilde{\mathcal{W}}_0 \star \tilde{\mathcal{W}}_0) \boxtimes \tilde{\mathcal{W}}_\ell).\end{aligned}$$

This obvious modification in the specification induces many difficulties in the analysis. In fact, the binomial transform given by the term $\text{SEQ}_{\geq 1} \mathcal{Z}$ is now perturbed and, although the class $\tilde{\mathcal{W}}_\ell$ is still satisfying a linear differential equation, its order is $d_\ell = 2 \cdot d_{\ell-1} + 1 + (\ell + 1 \bmod 2)$, with $d_1 = 4$ (thus $d_2 = 10, d_3 = 21, d_4 = 44 \dots$). While in the first subclasses, the order was growing linearly with ℓ ; here it grows quadratically, and consequently, from a calculation point of view, we are quickly unable to compute numerically the main statistics.

The last subclasses with some parallel constraint, we are interested in are given by

$$\begin{aligned}\mathcal{N}_0 &= \text{SEQ}_{\geq 1} \mathcal{Z}, \\ \mathcal{N}_\ell &= \mathcal{Z} + \mathcal{Z}^\square \star \mathcal{N}_\ell + \mathcal{Z}^\square \star ((\mathcal{N}_{\ell-1} \star \mathcal{N}_{\ell-1}) \boxtimes \mathcal{N}_\ell).\end{aligned}$$

The single constraint distinguishing \mathcal{N}_ℓ from the whole class \mathcal{FJ} is such that the number of nested fork nodes in a process of \mathcal{N}_ℓ is bounded by ℓ (i.e. at most 2^ℓ processes can be run in parallel). With the same proof as before, we prove that each class \mathcal{N}_ℓ is satisfying a linear differential equation. We get the recurrence

$$\begin{aligned}\mathcal{L}_c(N_0(z)) &= \frac{z}{1-z}, \\ \mathcal{L}_c(N_\ell(z)) &= \frac{z}{1-z-z\mathcal{L}_c(N_{\ell-1}(z))\odot\mathcal{L}_c(N_{\ell-1}(z))}.\end{aligned}$$

Because of the colored product whose both operands depend on ℓ , the degrees of the numerator and denominator polynomials are growing very rapidly. The sequence starts by 3, 10, 66, 2278, ... We prove that the degrees are satisfying the following recursive equation

$$d_1 = 3; \quad d_\ell = \frac{(d_{\ell-1} + 1)(d_{\ell-1} + 2)}{2}.$$

Consequently, the class \mathcal{N}_4 seems to be the limit to obtain quantitative results. We are thus far from the whole class of \mathcal{FJ} processes, especially if we assume that, as we have shown for trees [4], most of the runs are given by the widest structures.

3.2. Series-constrained processes

In the latter models we bounded the number of nested forks, i.e. the *fork depth*, trying to play with this parameter with the hope to guess the shape of the asymptotic number of increasing labellings of \mathcal{FJ} processes. Unfortunately, those models, while interesting by their technical properties are not as a nice approximation of \mathcal{P} as expected.

Now, we focus on another parameter: the *series depth*. We will no longer constrain the branching of the processes, within the meaning of constraining their parallelism. We will constrain the processes on their possibility to execute sequences of subprocesses. The processes will branch without restrictions, then execute only sequences of simple actions before joining all their parallel subprocesses. Once this global synchronization is done, the processes are allowed to begin a new forking step, etc.

In order to model these series-constrained processes, let us introduce the model of increasing binary diamonds which we studied in details in [3]. They are the basic blocks of this model.

Definition 30. Let \mathcal{T} (resp. \mathcal{D}) the unlabeled (resp labeled) class of (resp. increasing) binary diamonds. They are specified⁴ as follows:

$$\begin{aligned}\mathcal{T} &= \mathcal{Z} + \mathcal{Z} \times (1 + \mathcal{T} + \mathcal{T}^2) \times \mathcal{Z}, \\ \mathcal{D} &= \mathcal{Z} + \mathcal{Z}^\square \star (\mathcal{E} + \mathcal{D} + \mathcal{D} \star \mathcal{D}) \boxtimes \mathcal{Z}.\end{aligned}$$

Thus, a binary increasing diamond is similar to an increasing \mathcal{FJ} process where the join DAG is reduced to a single node (see Figure 3).

The generating functions for \mathcal{T} and \mathcal{D} satisfy $T(z) = z + z^2(1 + T(z) + T^2(z))$ and $D''(z) = 1 + D + D^2$ with $D(0) = 0$ and $D'(0) = 1$.

As seen in [3], the solution of this differential equation can be expressed as a Weierstrass \wp function from which we easily extract the asymptotic of $[z^n]D$. The radius of convergence ρ of D is equal to

$$\int_0^\infty \frac{dt}{\sqrt{\frac{2}{3}t^3 + t^2 + 2t}} \approx 3.1721709321\dots,$$

and thus

$$[z^n]D(z) \underset{n \rightarrow \infty}{\sim} 6 \frac{(n+1)!}{\rho^{n+2}}.$$

⁴The specification of \mathcal{D} can be described only with the two Greene's boxed products detailed in [8, pp.139–142]

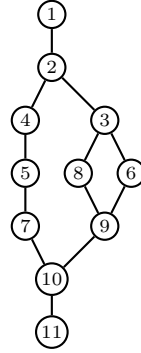


Figure 3: A binary increasing diamond

\mathcal{D} is a subclass of the \mathcal{FJ} processes but we can increase the expressivity of this model by putting several increasing binary diamonds in an ordered set. This construction models programs which can continue their execution after the last join action of the diamond.

Theorem 31. Let $\mathcal{S} = \text{SET}^{\square}(\mathcal{D})$ be the class of ordered sets of increasing binary diamonds and $T(z)$ the ordinary generating function of unlabeled binary diamonds.

- The asymptotic number of ordered sets of increasing binary diamonds is equivalent to the asymptotic number of increasing binary diamonds

$$[z^n]S \underset{n \rightarrow \infty}{\sim} [z^n]D.$$

- The asymptotic average number of increasing labelings of sequences of binary diamonds of size n is

$$\frac{[z^n] S(z)}{[z^n] (1 - T(z))^{-1}} \underset{n \rightarrow \infty}{\sim} 6 T'(\sigma) \frac{\sigma^{n+1} \cdot (n+1)!}{\rho^{n+2}}$$

$\sigma = \frac{1}{6}(\sqrt{13} - 1)$ is a solution of $T(z) = 1$, and thus $T'(\sigma) = \sqrt{13}/\sigma$. Approximations of the constant are given in the following form

$$\frac{[z^n] S(z)}{[z^n] (1 - T(z))^{-1}} \underset{n \rightarrow \infty}{\sim} \alpha \cdot \beta^{n+1} \cdot (n+1)!$$

with $\alpha \approx 15.7042 \dots$ and $\beta \approx 0.136896 \dots$

Proof. The first asymptotic behavior is obtained by a direct application of the Theorem 17. The second one is equal to the asymptotic number of ordered set of increasing binary diamonds, divided by the asymptotic number of sequences of unlabeled binary diamonds. $[z^n] \text{SEQ}(T)$ where T is the ordinary generating function of unlabeled binary diamonds: in bijection with a simple variety of trees. This is achieved using the *asymptotics of supercritical sequence* theorem of [9].

Note that the specification \mathcal{S} is a bit ambiguous: the sequences of atoms are counted several times. The length k sequence is counted exactly 2^{k-1} times, the number of integer compositions of k . But the asymptotic behavior is still relevant because the correction is asymptotically small. \square

3.3. Increasing \mathcal{FJ} processes

Finally, let us go back to the specification of increasing \mathcal{FJ} processes (cf. Definition 26). It can be seen as the limit of the increasing sequence (\mathcal{N}_ℓ) , however, the interesting property of holonomicity seems to be lost. The dominant singularity of \mathcal{P} satisfies $\eta \approx 2.31198062902106 \dots$. Let us first give an approximation of the asymptotic behavior of P_n .

Theorem 32. Let $r = 2.31197$ and $\epsilon = 2 \cdot 10^{-5}$. Then $r < \eta < r + \epsilon$. The average number on runs in all \mathcal{FJ} processes of size $n \geq 127$, satisfies

$$\gamma_0 n^{3/2} \rho_0^n (n+1)! \leq \frac{C_n}{F_n} \leq \gamma_1 n^{3/2} \rho_1^n (n+1)!,$$

with $\gamma_0 = 12\sqrt{\pi}(r+\epsilon)^2/\sqrt{1/3+1/2^{2/3}}$ and $\gamma_1 = 12\sqrt{\pi}r^2/\sqrt{1/3+1/2^{2/3}}$. Recall ρ is the dominant singularity of unlabeled \mathcal{FJ} processes, thus $\rho_0 = \rho/(r+\epsilon)$ and $\rho_1 = \rho/r$.

Obviously both bounds are exponentially far from the asymptotic behavior, however we have a detailed shape of it. The proof is given in Appendix B.

Let us remark that the shape of the asymptotic behavior favors a single large structure rather than a sequence of several \mathcal{FJ} processes. Finally, let us conclude the paper with the following result, based on an important technical assumption about the nature of the generating function $P(z)$ around its dominant singularity. In order to prove the result without assumption, we are trying to prove that the generating function $P(z)$ is a combination of elliptic functions, thus more complicated structures than diamonds.

Theorem 33. If $P(z)$ is Δ -analytic around η . Then the average number of runs of \mathcal{FJ} processes of size n is asymptotically

$$\frac{P_n}{F_n} \underset{n \rightarrow \infty}{\sim} \frac{12\sqrt{\pi} n^{3/2}}{\rho^2 \sqrt{\frac{1}{3} + \frac{1}{2^{2/3}}}} \left(\frac{\eta}{\rho}\right)^n (n+1)!,$$

with $\frac{\eta}{\rho} \approx 0.149670346096653\dots$

The average number above is to be compared to the average number of runs in tree-structures processes [4] of size n that is $0.5^n \cdot n!$. We note thus that the synchronization constrain drastically the number of runs.

4. Conclusion and future work

We defined two operators for building combinatorial classes. First, the ordered product impose a global increasing labeling constraint that can be found in many combinatorial structures such as search trees, integer compositions and of course concurrent processes. The colored product allows to specify the partition of structures. We only provided the example of Bell numbers but we intend to explore the range of possibilities concerning this operator.

Both products support the symbolic method with associated generating functions. They enjoy interesting algebraic properties (e.g. the preservation of holonomy) and we also demonstrated useful transfer theorems. As our development of fork-join processes suggests, it remains very difficult to work with the generating functions obtained from recursive specifications containing ordered and colored products. There are probably useful analytic tools that could be used to obtain interesting quantitative results (e.g. Borel resummation theory).

References

- [1] E. A. Bender. An asymptotic expansion for the coefficients of some formal power series. *Journal of the London Mathematical Society*, s2-9(3):451–458, 1975.
- [2] F. Bergeron, G. Labelle, and P. Leroux. *Combinatorial Species and Tree-like Structures*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1998.
- [3] O. Bodini, M. Dien, X. Fontaine A. Genitrini, and H.-K. Hwang. Increasing Diamonds. In *LATIN 2016: Theoretical Informatics - 12th Latin American Symposium*, pages 207–219, 2016.

- [4] O. Bodini, A. Genitrini, and F. Peschanski. A Quantitative Study of Pure Parallel Processes. *Electronic Journal of Combinatorics*, 23(1):P1.11, 39 pages, 2016.
- [5] G. Brightwell and P. Winkler. Counting linear extensions is $\sharp P$ -Complete. In C. Koutsougeras and J. S. Vitter, editors, *STOC*, pages 175–181. ACM, 1991.
- [6] A. Darrasse, K. Panagiotou, O. Roussel, and M. Soria. Biased Boltzmann samplers and generation of extended linear languages with shuffle. In *23rd International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, (AofA)*, pages 125–140, Montreal, Canada, June 2012.
- [7] P. Flajolet and A. M. Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3(2):216–240, 1990.
- [8] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.
- [9] P. Flajolet and M. Soria. General combinatorial schemas: Gaussian limit distributions and exponential tails. *Discrete Mathematics*, 114(1-3):159–180, 1993.
- [10] D. H. Greene. *Labelled Formal Languages and Their Uses*. PhD thesis, Stanford, CA, USA, 1983. AAI8320712.
- [11] D. E. Knuth. *The art of computer programming, volume 3: (2nd ed.) sorting and searching*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 1998.
- [12] R. H. Möhring. *Computationally Tractable Classes of Ordered Sets*. Institut für Ökonometrie und Operations Research: Report. 1987.

A. Appendix: ordered and colored products

A.1. Reminders on Borel and Laplace transforms

Let us recall here classical relations between combinatorial Laplace transform and the traditional Laplace transform. By definition, the traditional Laplace transform is defined by $\mathcal{L}f = \int_0^\infty \exp(-zt)f(t)dt$ instead of $\mathcal{L}_c f = \int_0^\infty \exp(-t)f(zt)dt$.

This operator is clearly linear. By a simple change of variable, we get that $\mathcal{L}f(z) = \frac{1}{z} (\mathcal{L}_c f) \left(\frac{1}{z}\right)$ or equivalently $\mathcal{L}_c f(z) = \frac{1}{z} (\mathcal{L}f) \left(\frac{1}{z}\right)$ (Notice the perfect involution !)

Laplace transforms admit a functional inverse called Borel transforms. This transform also has an integral representation: for traditional Laplace transforms, the Borel transform is $\mathcal{B}(f) = \frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \exp(zt)f(t)dt$ where c is greater than the real part of all singularities of $f(t)$.

By analogy, the combinatorial Borel transform is $\mathcal{B}_c(f) = \frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{\exp(zt)}{t} f(1/t)dt$ where c is greater than the real part of all singularities of $f(1/t)/t$. The link with traditional Borel transforms is $\mathcal{B}_c(f) = \mathcal{B}(1/zf(1/z))$ or equivalently $\mathcal{B}(f) = \mathcal{B}_c(f(1/z))' = \mathcal{B}_c(1/zf(1/z))$

Now, let us essentially concentrate our attention on combinatorial transforms. Combinatorial Laplace transforms create a bridge between exponential generating functions ($\sum_{n \geq 0} a_n \frac{z^n}{n!}$) and ordinary generating functions ($\sum_{n \geq 0} a_n z^n$). Precisely, we have:

$$\mathcal{L}_c \left(\sum_{n \geq 0} a_n \frac{z^n}{n!} \right) = \sum_{n \geq 0} a_n z^n$$

Reciprocally, we have

$$\mathcal{B}_c \left(\sum_{n \geq 0} a_n z^n \right) = \sum_{n \geq 0} a_n \frac{z^n}{n!}$$

From those formulas on formal series, one can easily derive the following identities:

- $\mathcal{L}_c f' = \frac{1}{z} (\mathcal{L}_c f - f_0)$
- $\mathcal{L}_c (\int f) = z \mathcal{L}_c f$
- $\mathcal{B}_c(zf) = \int \mathcal{B}_c f$
- $\mathcal{B}_c \left(\frac{f-f_0}{z} \right) = (\mathcal{B}_c f)'$

As for traditional Laplace transforms, the product of Laplace transform can be express using convolution product. We have :

$$z \mathcal{L}_c f \times \mathcal{L}_c g = \mathcal{L}_c \left(\int_0^z f(t)g(z-t)dt \right).$$

Or equivalently,

$$\mathcal{L}_c f \times \mathcal{L}_c g = \mathcal{L}_c \left(\int_0^z f(t)g'(z-t)dt + g_0 f(z) \right).$$

Observe that the ordered product, in fact, gives a combinatorial interpretation of this adapted convolution. We denote by $f * g$ the combinatorial convolution $\int_0^z f(t)g'(z-t)dt + g_0 f(z)$.

The product of combinatorial Borel transforms can also be expressed with convolution in the complex plane as follow: using the traditional

$$\mathcal{B}f \times \mathcal{B}g = \mathcal{B} \left(\frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} f(t)g(z-t)dt \right),$$

and compose it with the latter identities leads to the following formula

$$\mathcal{B}_c f \times \mathcal{B}_c g = \mathcal{B}_c \left(\frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{1}{(1-zt)t} f(1/t)g(z/(1-zt))dt \right).$$

A.2. Proofs of Section 2

Proof. Proof of Proposition 3

Using Definition 1, we note that an object from \mathcal{C} is given by an object from \mathcal{A} and one from \mathcal{B} only by shifting the labels of the second one. Thus the number of objects of size n in \mathcal{C} is given by $\sum_{k=1}^{n-1} A_k \cdot B_{n-k}$. The result of the composition $\mathcal{B}(\mathcal{L}(A(z)) \cdot \mathcal{L}(B(z)))$ gives directly this sum. \square

Proof. Proof of Proposition 6

Let us denote $F[A, B](z) = \int_0^z A(t)B'(z-t)dt + A(z)B(0)$ and differentiate it.

$$F'[A, B](z) = A(z)B'(0) + \int_0^z A(t)B''(z-t)dt + A'(z)B(0) = F[A, B'](z) + A'(z)B(0).$$

Since we use exponential generating functions, the differentiating operation can be seen as the removal of the atom with the smallest label. Thus we get an isomorphism between objects enumerated by $F'[A, B]$ and by C' . In conclusion, both associated generating functions are equal. \square

Proof. Proof of Theorem 16

Let us give the proof of the result, when both dominant singularities are equal and belong to $\rho \in]0, +\infty[$. The other cases are simpler and their proof is a slight adaptation of this case.

For all $\epsilon > 0$, there exists an integer n_0 , such that for all $n \geq n_0$ we get

$$\frac{1-\epsilon}{\rho} \leq \frac{A_n}{nA_{n-1}} \leq \frac{1+\epsilon}{\rho} \quad \text{and} \quad \frac{1-\epsilon}{\rho} \leq \frac{B_n}{nB_{n-1}} \leq \frac{1+\epsilon}{\rho}.$$

Thus, let us denote by $R = \max\{n_0, r\}$, then we partition the number C_n :

$$\begin{aligned} C_n &= \sum_{k=a}^{n-b} A_k B_{n-k} = A_a B_{n-a} + A_{n-b} B_b + \sum_{k=a+1}^{n-b-1} A_k B_{n-k} \\ &= A_a B_{n-a} + A_{n-b} B_b + \left(\sum_{k=a+1}^{R-1} A_k B_{n-k} + \sum_{k=n-R+1}^{n-b-1} A_k B_{n-k} + \sum_{k=R}^{n-R} A_k B_{n-k} \right). \end{aligned}$$

Let us focus on the first sum, and prove that it is negligible in front on B_{n-a} .

$$\begin{aligned} \frac{B_{n-(R-1)}}{B_{n-a}} \cdot \sum_{k=a+1}^{R-1} \frac{A_k B_{n-k}}{B_{n-(R-1)}} &= \frac{B_{n-(R-1)}}{B_{n-a}} \cdot \sum_{k=a+1}^{R-1} A_k \frac{B_{n-k}}{B_{n-(k+1)}} \frac{B_{n-(k+1)}}{B_{n-(k+2)}} \dots \frac{B_{n-(R-2)}}{B_{n-(R-1)}} \\ &\leq \frac{B_{n-(R-1)}}{B_{n-a}} \cdot \sum_{k=a+1}^{R-1} A_k \left(\frac{(n-(a+1)(1+\epsilon))^{R-1-k}}{\rho} \right). \end{aligned}$$

For n sufficiently large, there exists a constant γ such that

$$\begin{aligned} \frac{B_{n-(R-1)}}{B_{n-a}} \cdot \sum_{k=a+1}^{R-1} \frac{A_k B_{n-k}}{B_{n-(R-1)}} &\leq \frac{B_{n-(R-1)}}{B_{n-a}} \cdot \gamma \left(\frac{(n-(a+1)(1+\epsilon))^{R-a-2}}{\rho} \right) \\ &\leq \frac{B_{n-(R-1)} B_{n-(R-2)} \dots B_{n-(a+1)}}{B_{n-(R-2)} B_{n-(R-3)} \dots B_{n-a}} \cdot \gamma \left(\frac{(n-(a+1)(1+\epsilon))^{R-a-2}}{\rho} \right) \\ &\leq \gamma \left(\frac{(n-(a+1)(1+\epsilon))^{R-a-2}}{(n-(R-2))(1-\epsilon)} \right) \cdot \frac{\rho}{(n-(R-2))(1-\epsilon)}. \end{aligned}$$

We thus deduce that

$$\lim_{n \rightarrow \infty} \frac{B_{n-(R-1)}}{B_{n-a}} \cdot \sum_{k=a+1}^{R-1} \frac{A_k B_{n-k}}{B_{n-(R-1)}} = 0.$$

In the same way we prove

$$\lim_{n \rightarrow \infty} \frac{A_{n-(R-1)}}{A_{n-b}} \cdot \sum_{k=n-R+1}^{n-b-1} \frac{A_k B_{n-k}}{A_{n-(R-1)}} = 0.$$

By assumption, the last sum satisfies $\sum_{k=R}^{n-R} A_k B_{n-k} = \mathcal{O}(A_{n-R} + B_{n-R})$. And thus the statement (for this case) is proved. □

B. Appendix: Fork-Join processes

Proof. Proof of Proposition 25

Let $F(z)$ be the ordinary generating function enumerating \mathcal{FJ} processes. We denote by F_n the number of \mathcal{FJ} processes of size n . Thus we get $F(z) = \sum_n F_n z^n$. By using the specification of \mathcal{FJ} processes, and the symbolic method, we get directly a functional equation of $F(z)$

$$F(z) = z + zF(z) + zF(z)^3.$$

We can compute the first coefficients of the generating function $F(z) = z + z^2 + z^3 + 2z^4 + 5z^5 + 11z^6 + 24z^7 + 57z^8 + 141z^9 + 349z^{10} + \dots$. The combinatorial class of \mathcal{FJ} processes is a simple variety of trees (cf. [8, Part VII.3.]) and thus the way of analysing it is almost classical.

A first step consists in determining the closed form for the generating function. By solving the functional equation for $F(z)$ a first problem arises: in fact, the solution involves complex (and not real) coefficients and is thus not nice. Although the transformation we use in order to write $F(z)$ only with real coefficients is almost technical, our Computer Algebra Systems are not able to do it on their own. After technical works, we get

$$F(z) = -2\sqrt{\frac{1-z}{3z}} \sin\left(\frac{\pi}{6} + \frac{2}{3} \arctan\left(\frac{2\sqrt{1-3z+3z^2-\frac{31}{4}z^3}}{(3z)^{3/2}-2(1-z)^{3/2}}\right)\right).$$

Thus the dominant singularity of $F(z)$ is the solution of $1-3z+3z^2-\frac{31}{4}z^3=0$. The singularity is thus $\rho = \frac{1}{1+3 \cdot 2^{-2/3}}$ and is of square-root type. Using these details, we can exhibit the Puiseux expansion (cf. [8, Part VII.7.]) of $F(z)$ in order to analyze the asymptotic behavior of its coefficients.

$$F(z) \underset{z \rightarrow \rho}{=} 2^{-1/3} - \sqrt{\frac{3+2^{2/3}}{3 \cdot 2^{1/3}}} \cdot \sqrt{1-\frac{z}{\rho}} + \mathcal{O}\left(\left(1-\frac{z}{\rho}\right)^{3/2}\right).$$

From the Puiseux development around the dominant singularity, we deduce

$$F_n \underset{n \rightarrow \infty}{\sim} \frac{1}{2} \sqrt{\frac{1}{3} + \frac{1}{2^{2/3}}} \frac{\rho^{-n}}{\sqrt{\pi n^3}}.$$

□

Proof. Proof of Theorem 32

Let us introduce a proof by recurrence on n to obtain the upper bound for C_n . Let $r = 2.31197$ and $\epsilon = 2 \cdot 10^{-5}$. Then $r < \rho < r + \epsilon$.

We will prove by recurrence the upper bound. First, by calculating the first terms we get, for all $k \in \{3, \dots, 20\}$ we get $C_k \leq 6k\gamma_0 k! r^{-k-2}$, with $\gamma_0 = 1.08$. Second, for all $k \in \{21, \dots, 126\}$ we get $C_k \leq 6k\gamma_1 k! r^{-k-2}$, with $\gamma_1 = 1.004$. By calculating the terms for $k \in \{127, \dots, 951\}$ we get $C_k \leq 6k k! r^{-k-2}$,

Let n be an integer and suppose that for all $k \in \{127, \dots, n-1\}$, we get $C_k \leq 6k k! r^{-k-2}$.

We take $C_0 = 0, C_1 = 1$ and $C_2 = 1$, and Γ_0 and γ_1 are fixed.

The recurrence for the coefficients (C_n) is

$$C_n = C_{n-1} + \sum_{\ell=3}^{n-1} \sum_{k=1}^{\ell-2} \binom{\ell-1}{k} C_k C_{\ell-1-k} C_{n-\ell}.$$

Thus, by dividing both sides by $6n \cdot n! \cdot r^{-n-2}$, we get (for $n > 990$):

$$\begin{aligned} \frac{C_n}{6n \cdot n! \cdot r^{-n-2}} &= \frac{C_{n-1}}{6n \cdot n! \cdot r^{-n-2}} \\ &+ \frac{C_1}{6n \cdot n! \cdot r^{-n-2}} \left(2(n-2)C_1 C_{n-3} + (n-2)(n-3)C_2 C_{n-4} + \sum_{k=3}^{n-5} \binom{n-2}{k} C_k C_{n-2-k} \right) \\ &+ \sum_{\ell=3}^{n-2} \frac{C_{n-\ell}}{6n \cdot n! \cdot r^{-n-2}} \left(2(\ell-1)C_1 C_{\ell-2} + \sum_{k=2}^{\ell-3} \binom{\ell-1}{k} C_k C_{\ell-1-k} \right) \end{aligned}$$

In order to obtain a straight upper bound, some sum must be decompose according to small terms smaller than 21 and then the terms smaller than 126 and the other terms.

$$\begin{aligned} \frac{C_n}{6n \cdot n! \cdot r^{-n-2}} &\leq \frac{(n-1)r}{n^2} + \frac{2(n-3)r^3 + (n-4)r^4}{n^2(n-1)} + \gamma_0 \sum_{k=3}^{20} \frac{6k(n-2-k)}{n^2(n-1)} + \gamma_1 \sum_{k=21}^{126} \frac{6k(n-2-k)}{n^2(n-1)} \\ &+ \sum_{k=127}^{n-127} \frac{6k(n-2-k)}{n^2(n-1)} + \gamma_1 \sum_{k=n-126}^{n-21} \frac{6k(n-2-k)}{n^2(n-1)} + \gamma_0 \sum_{k=n-20}^{n-5} \frac{6k(n-2-k)}{n^2(n-1)} \\ &+ \frac{4(n-3)^2 r^3 + 6(n-4)r^4}{n^2(n-1)(n-2)(n-3)} + \gamma_0^3 \sum_{\ell=5}^{n-3} \frac{(n-\ell)(n-\ell)! r^\ell}{n \cdot n!} \\ &\quad \cdot \left(12(\ell-2)(\ell-1)! r^{-\ell} + 6(\ell-3)(\ell-1)! r^{-\ell+1} + 36(\ell-1)! r^{-\ell-3} \sum_{k=3}^{\ell-4} k(\ell-1-k) \right) \\ &+ \frac{1}{n \cdot n!} \left(2(n-4)(n-3)! r^4 + (n-5)(n-3)! r^5 + 6\gamma_0(n-3)! r \sum_{k=3}^{20} k(n-3-k) \right. \\ &\quad + 6\gamma_1(n-3)! r \sum_{k=21}^{126} k(n-3-k) + 6(n-3)! r \sum_{k=127}^{n-127} k(n-3-k) \\ &\quad \left. + 6\gamma_1(n-3)! r \sum_{k=n-126}^{n-21} k(n-3-k) + 6\gamma_0(n-3)! r \sum_{k=n-21}^{n-6} k(n-3-k) \right) \end{aligned}$$

The less important sums (that have been multiplied by a γ factor) can be easily bound by n times their dominant term. We thus prove that

$$\begin{aligned} \frac{C_n}{6n \cdot n! \cdot r^{-n-2}} &\leq 1 + \frac{2r-5}{n} + \frac{-7r + 1008\gamma_0^3 r^{-3} + 2r^3 + 45474 * \gamma_1 + r^4 - 46512 + 1008\gamma_0}{n^2} \\ &\quad + R(n, r, \gamma_0, \gamma_1), \end{aligned}$$

where the function $R(n, r, \gamma_0, \gamma_1)$ is negative for a positive n .

By evaluating r, γ_0 and γ_1 we prove that for all $n > 990$ we have

$$C_n \leq 6n \cdot n! \cdot r^{-n-2}.$$

The lower bound is proved in a simpler way, because from $n > 7$ we have

$$C_n \geq 6n \cdot n! \cdot (r + \epsilon)^{-n-2}.$$

□

□