Mathematical Structures in Computer Science (2022), **32**, pp. 271–299
doi:10.1017/S0960129522000354

CAMBRIDGE UNIVERSITY PRESS

PAPER

The combinator M and the Mockingbird lattice

Samuele Giraudo

LIGM, Université Gustave Eiffel, CNRS, ESIEE Paris, F-77454 Marne-la-Vallée, France
Email: samuele.giraudo@univ-eiffel.fr

(Received 8 April 2022; revised 11 September 2022; accepted 20 September 2022; first published online 13 October 2022)

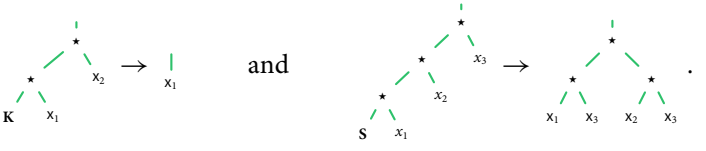
Abstract

We study combinatorial and order theoretic structures arising from the fragment of combinatory logic spanned by the basic combinator **M**. This basic combinator, named as the Mockingbird by Smullyan, is defined by the rewrite rule $\mathbf{M}x_1 \to x_1x_1$. We prove that the reflexive and transitive closure of this rewrite relation is a partial order on terms on **M** and that all connected components of its rewrite graph are Hasse diagram of lattices. This last result is based on the introduction of new lattices on duplicative forests, which are sorts of treelike structures. These lattices are not graded, not self-dual, and not semi-distributive. We present some enumerative properties of these lattices like the enumeration of their elements, of the edges of their Hasse diagrams, and of their intervals. These results are derived from formal power series on terms and on duplicative forests endowed with particular operations.

Keywords: Partial orders; lattices; combinatory logic; rewrite systems; treelike structures; formal power series

1. Introduction

Combinatory logic is a model of computation introduced by Schönfinkel (1924) and developed by Curry (1930) with the objective to abstain from the need of bound variables specific to the λ-calculus. Its combinatorial heart is formed by terms, which are binary trees with labeled leaves, and rules to compute a result from a term, which are rewrite relations on trees (Bezem et al. 2003; Baader and Nipkow 1998) (clear and complete modern references about combinatory logic are Barendregt 1981; Bimbo 2011; Hindley and Seldin 2008; Wolfram 2021). An important instance of combinatory logic is the system containing the basic combinators **K** and **S** together with the two rewrite rules



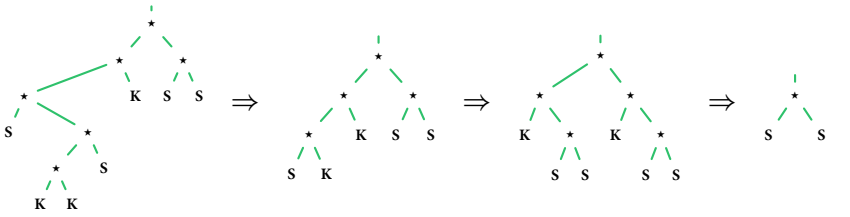
(1.0.1)

In this system, we have for instance the sequence of computation

This research has been partially supported by the projects CARPLO (ANR-20-CE40-0007) and LambdaComb (ANR-21-CE48-0017) of the Agence nationale de la recherche.

© The Author(s), 2022. Published by Cambridge University Press

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press



(1.0.2)

obtained by performing at each step one rewrite according to the previous rules. This system is important because it is combinatorially complete: each λ-term can be translated, by so-called bracket abstraction algorithms (Curry and Feys 1958; Schönfinkel 1924), into a term over **K** and **S** emulating it.

A lot of other basic combinators with their own rewrite rules have been introduced by Smullyan in Smullyan (1985) after – now widely used – bird names, forming the enchanted forest of combinator birds. For instance, **K** is the Kestrel and **S** is the Starling. Usual computer science-oriented questions consist in considering a fragment of combinatory logic, that is a finite set of combinators with their rewrite rules which is not necessarily combinatorially complete, and ask for the following questions:

- **(a)** Given two terms t and t', can we decide if t and t' can be rewritten eventually in a same term? This is known as the word problem (Baader and Nipkow 1998; Statman 2000). This question admits a positive answer for some basic combinators like among others the Lark (Statman 1989; Sprenger and Wymann-Böni 1993) and the Warbler (Sprenger and Wymann-Böni 1993) but it still open for the Starling (Barendregt et al. 2017);
- **(b)** Given a term t, can we decide if all rewrite sequences starting from t are finite? This is known as the strong normalization problem. This question admits a positive answer for, among others, the Starling (Waldmann 2000) and the Jay (Probst and Studer 2001). Related to this, see also David et al. (2013), Bendkowski et al. (2017) for a probabilistic and asymptotic study of strong normalizing terms.

Here, we decide to pursue this study in a different direction by asking questions from a combinatorial point of view, including the study of order theoretic structures and adopting an enumerative approach. In particular, by denoting by $\preccurlyeq$ (resp. by $\equiv$) the reflexive and transitive (resp. reflexive, symmetric, and transitive) closure of the rewrite relation, we ask the following questions:

- **(a′)** Determine if $\preccurlyeq$ is a partial order relation;
- **(b′)** In this case, determine if each interval of this poset is a lattice;
- **(c′)** Enumerate the $\equiv$-equivalence classes of terms w.r.t. the minimal degrees of their terms.

This work fits in this general project consisting in mixing combinatory logic with combinatorics. These enumerative questions, including the enumeration w.r.t. some size notions (like the height or the number of basic combinators of the terms) of some particular terms of a system (like its normal forms or the minimal or maximal elements of its partial order), lead to density properties of the counted terms among the terms of the same size. This gives information about the probability for a term of a given size generated uniformly at random to satisfy some property. Besides, some of the mentioned questions can be addressed by working with rewrite graphs, that are graphs labeled by terms and where there is a directed edge between two terms if the first can be obtained from the second by a rewriting step. Such graphs appear naturally in the study of term rewrite systems (Berry and é 1979). Some different graphs have been considered in the literature, where vertices are no longer terms but rather sequences of rewriting steps (see for instance Berry and é

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

1979; Venturini Zilli 1984 and Barendregt 1981, Exercise 12.4.2.). These graphs, as well as some of their semilattice properties (Berry and é 1979, Theorem 2.3.1.), have no direct links with the present work.

We choose here to start this project by studying the system made of a single combinator, the combinator **M**, known as the Mockingbird (Smullyan 1985) or as the little omega. This combinator is a very simple and an important one (see for instance Statman 2011, 2017). By drawing the portions of the rewrite graph starting from terms on **M**, the first properties that stand out are that the graph does not contain any nontrivial loops and that its connected components are finite and have exactly one minimal and one maximal element. At this stage, driven by computer exploration, we conjecture that the relation $\preccurlyeq$ on the terms on **M** is a partial order relation and that each $\equiv$-equivalence class is a lattice w.r.t. this partial order relation. This lattice property is for us a good clue for the fact that this system contains rather rich combinatorial properties.

To prove this last property, we introduce new lattices on duplicative forests, that are kinds of treelike structures, and show that each maximal interval from any term on **M** is isomorphic as a poset to a maximal interval of a lattice of duplicative forests (Theorem 3.3.4). We define the Mockingbird lattice of order $d \geqslant 0$ as the lattice $\mathrm{M}(d)$ consisting in the closed terms on **M** equal as or greater than the right comb closed term on **M** of degree d. We prove that each interval of the poset of duplicative forests is contained as an interval in $\mathrm{M}(d)$ for a certain $d \geqslant 0$ (Theorem 3.3.5). Since any closed term on **M** can be seen as a binary tree, this provides a new lattice structure on these objects. A lot of similar lattices have been studied on binary trees such as, among others the very famous Tamari lattice (Tamari 1962), the Kreweras lattice (Kreweras 1972), the Stanley lattice (Stanley 1975), the phagocyte lattice (Baril and Pallo 2006), and the pruning-grafting lattice (Baril and Pallo 2008). However, unlike these lattices having for each order $d \geqslant 0$ a cardinality equal to the d-th Catalan number, the lattices $\mathrm{M}(d)$ are enumerated by a different integer sequence. To obtain enumerative results about the Mockingbird lattices and all the posets of terms on **M** in general, we use formal power series on terms and on duplicative forests, and several products on these. In this way, we enumerate the minimal and maximal elements of the infinite poset of the closed terms on **M** (Propositions 4.2.1 and 4.2.2), the lengths of the shortest and longest saturated chains of $\mathrm{M}(d)$ (Proposition 4.3.1), the cardinality of $\mathrm{M}(d)$ (Proposition 4.4.3), the number of edges of the Hasse diagram of $\mathrm{M}(d)$ (Proposition 4.4.6), and the number of intervals of $\mathrm{M}(d)$ (Proposition 4.4.10). We also provide general results for systems of combinatory logic: we give in particular a necessary condition on the rewrite rules in order to have only finite connected components in the rewrite graph (Proposition 2.2.3). This relies on a combinatorial property on basic combinators, called the hierarchical property. We also discuss some consequences of this fact in order to construct models for systems having this property. These are in fact the algebras over a certain abstract clone (Taylor 1993) defined from the system. When the relation $\preccurlyeq$ is a partial order relation, such models have a potential nice algorithmic computational complexity.

This paper is organized as follows. Section 2 contains the preliminary notions and definitions about terms, rewrite relations, and combinatory logic systems. We show here some general properties of these systems. In Section 3, we study the combinatory logic system on **M** and the Mockingbird lattices. Finally, Section 4 contains our enumerative results about Mockingbird lattices. This text ends with the presentation of some open questions raised by this work.

This paper is an extended version of Giraudo (2022), announcing the main results without any proofs. The present version contains all proofs of the stated results, more examples, and presents new results as discussions about some models of the combinatory logic system on **M** and the enumeration of isolated elements of the poset of the closed terms on **M**. This version uses also the formalism of abstract clones in order to work with terms and rewrite systems.

General notations and conventions

If S is a finite set $\#S$ is the cardinality of S. For any integers i and j, $[i, j]$ denotes the set $\{i, i+1, \ldots, j\}$. For any integer i, $[i]$ denotes the set $[1, i]$ and $[\![i]$ denotes the set $[0, i]$. For any set A, A^*

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

is the set of all words on A. For any $w \in A^*$, $\ell(w)$ is the length of w, for any $i \in [\ell(w)]$, $w(i)$ is the i-th letter of w, and $|w|_a$ is the number of occurrences of the letter $a \in A$ in w. The only word of length 0 is the empty word ϵ. For any statement P, the Iverson bracket $[\![P]\!]$ takes 1 as value if P is true and 0 otherwise.

2. Terms, Rewrite Relations, and Combinatory Logic Systems

In this preliminary part, we set the main definitions and notations used in the sequel. We introduce also the central notion of combinatory logic systems, which are defined from special kinds of rewrite relations. We show also some general properties of these systems.

2.1 Terms, compositions, and rewrite relations

Let us start by presenting here the notions of terms, composition, and rewrite relations.

2.1.1 Terms

An *alphabet* is a finite set $\mathfrak{G}$. Its elements are called *constants* or *basic combinators*. Any element of the set $\mathbb{X} := \bigcup_{n \geqslant 1} \mathbb{X}_n$, where $\mathbb{X}_n := \{\mathsf{x}_1, \ldots, \mathsf{x}_n\}$, is a *variable*. The set $\mathfrak{T}(\mathfrak{G})$ of *$\mathfrak{G}$-terms* (or simply *terms* when the context is clear) is so that any variable of $\mathbb{X}$ is a $\mathfrak{G}$-term, any constant of $\mathfrak{G}$ is a $\mathfrak{G}$-term, and if $\mathfrak{t}_1$ and $\mathfrak{t}_2$ are two $\mathfrak{G}$-terms, then $(\mathfrak{t}_1 \star \mathfrak{t}_2)$ is a $\mathfrak{G}$-term. For any $n \geqslant 0$, we denote by $\mathfrak{T}_n(\mathfrak{G})$ the set of $\mathfrak{G}$-terms $\mathfrak{t}$ having all variables belonging to $\mathbb{X}_n$.

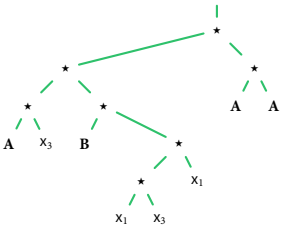
From this definition, any term is a rooted planar binary tree where leaves are decorated by variables or by constants. We shall express terms concisely by removing superfluous parentheses by considering that $\star$ associates to the left and also by removing the symbols $\star$ (see the example given in (2.1.1), (2.1.2), and (2.1.3)).

Let $\mathfrak{t}$ be a $\mathfrak{G}$-term. The *degree* $\deg(\mathfrak{t})$ of $\mathfrak{t}$ is the number of internal nodes (the nodes decorated by $\star$) of $\mathfrak{t}$ seen as a binary tree. For any $a \in \mathbb{X} \cup \mathfrak{G}$, $\deg_a(\mathfrak{t})$ is the number of nodes labeled by a in $\mathfrak{t}$. The *depth* of a node u of $\mathfrak{t}$ is the number of internal nodes in the path connecting the root of $\mathfrak{t}$ and u. The *height* $\mathrm{ht}(\mathfrak{t})$ of $\mathfrak{t}$ is the maximal depth among all the nodes of $\mathfrak{t}$. The *frontier* of $\mathfrak{t}$ is the sequence of all the variables appearing in $\mathfrak{t}$ from the left to the right. The term $\mathfrak{t}$ is *planar* if its frontier is of the form $\mathsf{x}_1\mathsf{x}_2 \ldots \mathsf{x}_n$ where $n \geqslant 0$. The term $\mathfrak{t}$ is *linear* if there are no multiple occurrences of the same variable in its frontier. A *closed term* or *combinator* is a term of $\mathfrak{T}_0(\mathfrak{G})$ having thus no occurrence of any variable.

For instance, by setting $\mathfrak{G} := \{\mathbf{A}, \mathbf{B}\}$, the $\mathfrak{G}$-term

$$\mathfrak{t} := (((\mathbf{A} \star \mathsf{x}_3) \star (\mathbf{B} \star ((\mathsf{x}_1 \star \mathsf{x}_3) \star \mathsf{x}_1))) \star (\mathbf{A} \star \mathbf{A})) \tag{2.1.1}$$

draws as the binary tree



(2.1.2)

and expresses concisely as

$$\mathfrak{t} = (\mathbf{A}\,\mathsf{x}_3)(\mathbf{B}(\mathsf{x}_1\mathsf{x}_3\mathsf{x}_1))(\mathbf{A}\,\mathbf{A}). \tag{2.1.3}$$

Its degree is 7, its height is 5, its frontier is $\mathsf{x}_3\mathsf{x}_1\mathsf{x}_3\mathsf{x}_1$ and is an element $\mathfrak{T}_n(\mathfrak{G})$ for any $n \geqslant 3$.

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

2.1.2 Compositions of terms and clones

Let $\mathfrak{t}$ and $\mathfrak{t}'_1, \ldots, \mathfrak{t}'_n$, $n \geqslant 0$ be $\mathfrak{G}$-terms. The *composition* of $\mathfrak{t}$ with $\mathfrak{t}'_1, \ldots, \mathfrak{t}'_n$ is the $\mathfrak{G}$-term $\mathfrak{t}\big[\mathfrak{t}'_1, \ldots, \mathfrak{t}'_n\big]$ obtained by simultaneously replacing for all $i \in [n]$ all occurrences of the variables $\mathfrak{x}_i$ in $\mathfrak{t}$ by $\mathfrak{t}'_i$. For instance,

$$\mathbf{A}(\mathfrak{x}_3\mathfrak{x}_1)\mathfrak{x}_3[\mathfrak{x}_1\mathfrak{x}_4, \mathfrak{x}_1\mathfrak{x}_1\mathfrak{x}_2, \mathbf{A}(\mathbf{A}\,\mathfrak{x}_1)] = \mathbf{A}(\mathbf{A}(\mathbf{A}\,\mathfrak{x}_1)(\mathfrak{x}_1\mathfrak{x}_4))(\mathbf{A}(\mathbf{A}\,\mathfrak{x}_1)). \tag{2.1.4}$$

Given two $\mathfrak{G}$-terms $\mathfrak{t}$ and $\mathfrak{s}$, $\mathfrak{s}$ is a *factor* of $\mathfrak{t}$ if

$$\mathfrak{t} = \mathfrak{t}'[\mathfrak{s}_1, \ldots, \mathfrak{s}_{i-1}, \mathfrak{s}, \mathfrak{s}_{i+1}, \ldots, \mathfrak{s}_n][\mathfrak{r}_1, \ldots, \mathfrak{r}_m] \tag{2.1.5}$$

for some integers $n, m \geqslant 0$ and $\mathfrak{G}$-terms $\mathfrak{t}', \mathfrak{s}_1, \ldots, \mathfrak{s}_{i-1}, \mathfrak{s}_{i+1}, \ldots, \mathfrak{s}_n, \mathfrak{r}_1, \ldots, \mathfrak{r}_m$, where $\mathfrak{x}_i$ appears in $\mathfrak{t}'$. When this property does not hold, $\mathfrak{t}$ *avoids* $\mathfrak{s}$. When in (2.1.5), for all $i \in [m]$ the $\mathfrak{r}_i$ are variables, $\mathfrak{s}$ is a *suffix* of $\mathfrak{t}$. When this property does not hold, $\mathfrak{t}$ *suffix-avoids* $\mathfrak{s}$. For instance, the term defined in (2.1.3) admits $\mathbf{B}(\mathfrak{x}_1\mathfrak{x}_2)$ as factor, $\mathfrak{x}_1\mathfrak{x}_2\mathfrak{x}_3$ as suffix, but suffix-avoids $\mathfrak{x}_1\mathfrak{x}_1\mathfrak{x}_1$.

Observe that we have, for any $n, m \geqslant 0$, $i \in [n]$, and $\mathfrak{t}, \mathfrak{s}_1, \ldots, \mathfrak{s}_n, \mathfrak{r}_1, \ldots, \mathfrak{r}_m \in \mathfrak{T}(\mathfrak{G})$,

$$\mathfrak{x}_i[\mathfrak{t}_1, \ldots, \mathfrak{t}_n] = \mathfrak{t}_i, \tag{2.1.6a}$$

$$\mathfrak{t}[\mathfrak{x}_1, \ldots \mathfrak{x}_n] = \mathfrak{t}, \tag{2.1.6b}$$

$$\mathfrak{t}[\mathfrak{s}_1, \ldots, \mathfrak{s}_n][\mathfrak{r}_1, \ldots, \mathfrak{r}_m] = \mathfrak{t}[\mathfrak{s}_1[\mathfrak{r}_1, \ldots, \mathfrak{r}_m], \ldots, \mathfrak{s}_n[\mathfrak{r}_1, \ldots, \mathfrak{r}_m]]. \tag{2.1.6c}$$

These three relations imply that the set $\mathfrak{T}(\mathfrak{G})$ together with the composition operation is an abstract clone (Taylor 1993). This is in fact the free abstract clone generated by $\mathfrak{G} \sqcup \{\star\}$ where $\star$ is a binary generator.

2.1.3 Rewrite relations

A *rewrite relation* on $\mathfrak{T}(\mathfrak{G})$ is a binary relation $\rightarrow$ on $\mathfrak{T}(\mathfrak{G})$. The *context closure* of $\rightarrow$ is the binary relation $\Rightarrow$ on $\mathfrak{T}(\mathfrak{G})$ satisfying

$$\mathfrak{t}[\mathfrak{s}_1, \ldots, \mathfrak{s}_i, \ldots, \mathfrak{s}_n][\mathfrak{r}_1, \ldots, \mathfrak{r}_m] \Rightarrow \mathfrak{t}\big[\mathfrak{s}_1, \ldots, \mathfrak{s}'_i, \ldots, \mathfrak{s}_n\big][\mathfrak{r}_1, \ldots, \mathfrak{r}_m] \tag{2.1.7}$$

where $\mathfrak{t}$ is a planar $\mathfrak{G}$-term, $\mathfrak{x}_i$ appears in $\mathfrak{t}$, and $\mathfrak{s}_1, \ldots, \mathfrak{s}_i, \mathfrak{s}'_i, \ldots, \mathfrak{s}_n$, $n \geqslant 1$, $i \in [n]$, and $\mathfrak{r}_1, \ldots, \mathfrak{r}_m$, $m \geqslant 0$, are $\mathfrak{G}$-terms, and $\mathfrak{s}_i \rightarrow \mathfrak{s}'_i$. Intuitively, $\mathfrak{t} \Rightarrow \mathfrak{t}'$ if $\mathfrak{t}'$ can be obtained from $\mathfrak{t}$ by replacing a factor $\mathfrak{s}$ of $\mathfrak{t}$ by $\mathfrak{s}'$ and by identifying variables with terms in a coherent way, whenever $\mathfrak{s} \rightarrow \mathfrak{s}'$. For instance, for $\mathfrak{G} := \{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$ and the rewrite relation $\rightarrow$ satisfying $\mathfrak{x}_1(\mathbf{A}\,\mathfrak{x}_1) \rightarrow \mathfrak{x}_1\mathfrak{x}_2$, we have

$$\mathfrak{x}_3\,\mathbf{B}(\mathbf{A}(\mathfrak{x}_3\,\mathbf{B}))(\mathfrak{x}_4\,\mathbf{A}) \Rightarrow \mathfrak{x}_3\,\mathbf{B}\,\mathbf{C}(\mathfrak{x}_4\,\mathbf{A}) \tag{2.1.8}$$

because

$$\mathfrak{x}_3\,\mathbf{B}(\mathbf{A}(\mathfrak{x}_3\,\mathbf{B}))(\mathfrak{x}_4\,\mathbf{A}) = \mathfrak{x}_1(\mathfrak{x}_2\,\mathbf{A})\;[\mathfrak{x}_1(\mathbf{A}\,\mathfrak{x}_1), \mathfrak{x}_4]\;[\mathfrak{x}_3\,\mathbf{B}, \mathbf{C}, \mathfrak{x}_3, \mathfrak{x}_4] \tag{2.1.9}$$

and

$$\mathfrak{x}_3\,\mathbf{B}\,\mathbf{C}(\mathfrak{x}_4\,\mathbf{A}) = \mathfrak{x}_1(\mathfrak{x}_2\,\mathbf{A})\;[\mathfrak{x}_1\mathfrak{x}_2, \mathfrak{x}_4]\;[\mathfrak{x}_3\,\mathbf{B}, \mathbf{C}, \mathfrak{x}_3, \mathfrak{x}_4]. \tag{2.1.10}$$

Observe in this example that the variable $\mathfrak{x}_2$ occurs in the right member of $\rightarrow$ and not in the left member. For this reason, this variable can be replaced by any term in the context closure of $\rightarrow$ (it is replaced by $\mathbf{C}$ in the example).

2.1.4 Applicative term rewrite systems

An *applicative term rewrite system* (or *ATRS* for short) is a pair $(\mathfrak{G}, \rightarrow)$ such that $\mathfrak{G}$ is an alphabet and $\rightarrow$ is a rewrite relation on $\mathfrak{T}(\mathfrak{G})$. Let $\mathcal{T} := (\mathfrak{G}, \rightarrow)$ be an ATRS. We denote by $\preccurlyeq$ the reflexive and transitive closure of $\Rightarrow$. This relation $\preccurlyeq$ is by construction a preorder. We denote by $\equiv$ the reflexive, symmetric, and transitive closure of $\Rightarrow$ and by $[\mathfrak{t}]_{\equiv}$ the $\equiv$-equivalence class

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

of $\mathfrak{t} \in \mathfrak{T}(\mathfrak{G})$. This relation $\equiv$ is by construction an equivalence relation. The *rewrite graph* $\mathrm{G}_{\mathcal{T}}$ of $\mathcal{T}$ is the digraph on the set $\mathfrak{T}(\mathfrak{G})$ of vertices where there is an arc from $\mathfrak{t} \in \mathfrak{T}(\mathfrak{G})$ to $\mathfrak{t}' \in \mathfrak{T}(\mathfrak{G})$ if $\mathfrak{t} \Rightarrow \mathfrak{t}'$. For any $\mathfrak{t} \in \mathfrak{T}(\mathfrak{G})$, we also denote by $\mathrm{G}_{\mathcal{T}}(\mathfrak{t})$ the subgraph of $\mathrm{G}_{\mathcal{T}}$ restrained on the set $\left\{\mathfrak{t}' \in \mathfrak{T}(\mathfrak{G}) : \mathfrak{t} \preccurlyeq \mathfrak{t}'\right\}$ of vertices. The ATRS $\mathcal{T}$ is *locally finite* if for any term $\mathfrak{t}$, $[\mathfrak{t}]_{\equiv}$ is finite. This is equivalent to the fact that all the connected components of $\mathrm{G}_{\mathcal{T}}$ are finite. A $\mathfrak{G}$-term $\mathfrak{t}$ is a *normal form* of $\mathcal{T}$ if there is no arc of source $\mathfrak{t}$ in $\mathrm{G}_{\mathcal{T}}$. We say that $\mathfrak{t}$ is *weakly normalizing* if there is at least one normal form in $\mathrm{G}_{\mathcal{T}}(\mathfrak{t})$ and that $\mathfrak{t}$ is *strongly normalizing* if $\mathrm{G}_{\mathcal{T}}(\mathfrak{t})$ is finite and acyclic. When all the $\mathfrak{G}$-terms are strongly normalizing, $\mathcal{T}$ is *terminating*. Besides, if for any $\mathfrak{t}, \mathfrak{s}_1, \mathfrak{s}_2 \in \mathfrak{T}(\mathfrak{G})$, $\mathfrak{t} \preccurlyeq \mathfrak{s}_1$ and $\mathfrak{t} \preccurlyeq \mathfrak{s}_2$ implies the existence of $\mathfrak{t}' \in \mathfrak{T}(\mathfrak{G})$ such that $\mathfrak{s}_1 \preccurlyeq \mathfrak{t}'$ and $\mathfrak{s}_2 \preccurlyeq \mathfrak{t}'$, then $\mathcal{T}$ is *confluent*.

2.1.5 Partial orders

When $\mathcal{T} := (\mathfrak{G}, \to)$ is an ATRS such that $\preccurlyeq$ is antisymmetric, $\preccurlyeq$ is a partial order relation. In this case, $\mathcal{T}$ has the *poset property* and we shall denote by $\mathcal{P}_{\mathcal{T}}$ the poset $(\mathfrak{T}(\mathfrak{G}), \preccurlyeq)$. For any $\mathfrak{t} \in \mathfrak{T}(\mathfrak{G})$, let $\mathcal{P}_{\mathcal{T}}(\mathfrak{t})$ be the subposet of $\mathcal{P}_{\mathcal{T}}$ having $\mathfrak{t}$ as least element. Again in this case, a $\mathfrak{G}$-term $\mathfrak{t}$ is *minimal* (resp. *maximal*) if $\mathfrak{t}$ is a minimal (resp. maximal) element of $\mathcal{P}_{\mathcal{T}}$. Observe that any normal form is maximal but the converse is false because a maximal element $\mathfrak{t}$ could satisfy $\mathfrak{t} \Rightarrow \mathfrak{t}$. A $\mathfrak{G}$-term $\mathfrak{t}$ is *isolated* if $\mathfrak{t}$ is both minimal and maximal. When $\mathcal{T}$ has the poset property, $\mathcal{T}$ is *rooted* if for any term $\mathfrak{t}$, the subposet $[\mathfrak{t}]_{\equiv}$ of $\mathcal{P}_{\mathcal{T}}$ has a unique minimal element. Finally, $\mathcal{T}$ has the *lattice property* if $\mathcal{T}$ has the poset property and for all $\mathfrak{t} \in \mathfrak{T}(\mathfrak{G})$, all posets $\mathcal{P}_{\mathcal{T}}(\mathfrak{t})$ are lattices.

2.2 Combinatory logic systems

We begin by defining combinatory logic systems as particular ATRS and present some consequence of nonerasing and hierarchical combinatory logic systems (notions defined thereafter). We explain some of the consequences of the local finiteness, the poset property, and the lattice property on the models of combinatory logic systems.

2.2.1 Main definitions

A *combinatory logic system* (or *CLS* for short) is an ATRS $\mathcal{C} := (\mathfrak{G}, \to)$ such that for each constant $\mathbf{X}$ of $\mathfrak{G}$, there is exactly one rewrite rule where $\mathbf{X}$ appears, and this rule is of the form

$$\mathbf{X}\,\mathrm{x}_1 \ldots \mathrm{x}_n \to \mathfrak{t}_{\mathbf{X}} \tag{2.2.1}$$

where $n \geqslant 1$ and $\mathfrak{t}_{\mathbf{X}}$ is a term having no constants and having all variables in $\mathbb{X}_n$. Moreover, all rewrite rules of $\mathcal{C}$ must be of the form (2.2.1). The integer n is the *order* of $\mathbf{X}$ in $\mathcal{C}$. Some well-known terms $\mathfrak{t}_{\mathbf{X}}$ appearing among others in Smullyan (1985) are,

- with order 1, $\mathfrak{t}_{\mathbf{I}} := \mathrm{x}_1$ (*Identity bird*), $\mathfrak{t}_{\mathbf{M}} := \mathrm{x}_1\mathrm{x}_1$ (*Mockingbird*);
- with order 2, $\mathfrak{t}_{\mathbf{K}} := \mathrm{x}_1$ (*Kestrel*), $\mathfrak{t}_{\mathbf{T}} := \mathrm{x}_2\mathrm{x}_1$ (*Thrush*), $\mathfrak{t}_{\mathbf{M_1}} := \mathrm{x}_1\mathrm{x}_1\mathrm{x}_2$ (*Mockingbird 1*), $\mathfrak{t}_{\mathbf{W}} := \mathrm{x}_1\mathrm{x}_1\mathrm{x}_2$ (*Warbler*), $\mathfrak{t}_{\mathbf{L}} := \mathrm{x}_1(\mathrm{x}_2\mathrm{x}_2)$ (*Lark*), $\mathfrak{t}_{\mathbf{O}} := \mathrm{x}_2(\mathrm{x}_1\mathrm{x}_2)$ (*Owl*), $\mathfrak{t}_{\mathbf{U}} := \mathrm{x}_2(\mathrm{x}_1\mathrm{x}_1\mathrm{x}_2)$ (*Turing Bird*);
- with order 3, $\mathfrak{t}_{\mathbf{C}} := \mathrm{x}_1\mathrm{x}_3\mathrm{x}_2$ (*Cardinal*), $\mathfrak{t}_{\mathbf{V}} := \mathrm{x}_3\mathrm{x}_1\mathrm{x}_2$ (*Vireo*), $\mathfrak{t}_{\mathbf{B}} := \mathrm{x}_1(\mathrm{x}_2\mathrm{x}_3)$ (*Bluebird*), $\mathfrak{t}_{\mathbf{S}} := \mathrm{x}_1\mathrm{x}_3(\mathrm{x}_2\mathrm{x}_3)$ (*Starling*);
- with order 4, $\mathfrak{t}_{\mathbf{J}} := \mathrm{x}_1\mathrm{x}_2(\mathrm{x}_1\mathrm{x}_4\mathrm{x}_3)$ (*Jay*).

The constant $\mathbf{X}$ is *nonerasing* if $\mathfrak{t}_{\mathbf{X}}$ contains at least one occurrence of each variable x_i for any $i \in [n]$. The constant $\mathbf{X}$ is *hierarchical* if for any $i \in [n]$, x_i appears in $\mathfrak{t}_{\mathbf{X}}$ at depth $n + 1 - i$. For instance, the terms $\mathfrak{t}_{\mathbf{X}}$ such that $\mathbf{X}$ are hierarchical and of order 3 or less are

$$\mathrm{x}_1\mathrm{x}_1,\ \ \mathrm{x}_1\mathrm{x}_1\mathrm{x}_2,\ \ \mathrm{x}_2(\mathrm{x}_1\mathrm{x}_1),\ \ \mathrm{x}_1\mathrm{x}_1\mathrm{x}_2\mathrm{x}_3,\ \ \mathrm{x}_2(\mathrm{x}_1\mathrm{x}_1)\mathrm{x}_3,\ \ \mathrm{x}_3(\mathrm{x}_1\mathrm{x}_1\mathrm{x}_2),\ \ \mathrm{x}_3(\mathrm{x}_2(\mathrm{x}_1\mathrm{x}_1)). \tag{2.2.2}$$

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

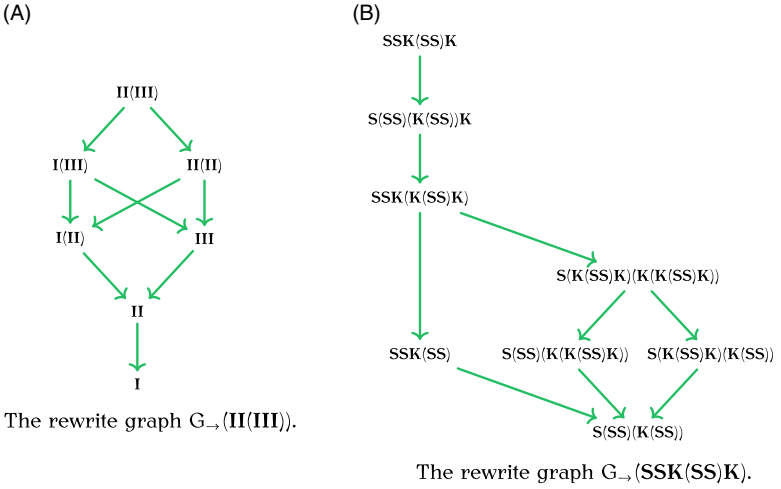


Figure 1. Some subgraphs of rewrite graphs of some CLS.

In particular, if **X** is hierarchical, then **X** is nonerasing. We say that $\mathcal{C}$ is *nonerasing* (resp. *hierarchical*) if all constants of $\mathfrak{G}$ are nonerasing (resp. hierarchical). Other interesting properties of basic combinators are introduced in Bimbo (2011) but they do not intervene directly in this work.

Consider for instance the CLS $\mathcal{C} := (\mathfrak{G}, \to)$ where $\mathfrak{G}$ contains only the constant **I** where $\mathfrak{t}_{\mathbf{I}}$ is defined above. It is straightforward to show that $\mathcal{C}$ has the poset property. Nevertheless, $\mathcal{C}$ has not the lattice property, as suggested by the Hasse diagram shown in Figure 1A. On the other side, the CLS $\mathcal{C} := (\mathfrak{G}, \to)$ where $\mathfrak{G}$ contains only the constants **K** and **S** where $\mathfrak{t}_{\mathbf{K}}$ and $\mathfrak{t}_{\mathbf{S}}$ are defined above does not have the poset property. Indeed, $\preccurlyeq$ is not antisymmetric because by setting $\mathfrak{t} := \mathbf{S}(\mathbf{S}(\mathbf{KS})(\mathbf{K}(\mathbf{SKK})))(\mathbf{K}(\mathbf{SKK}))$, we have $\mathfrak{s} \preccurlyeq \mathfrak{s}'$ and $\mathfrak{s}' \preccurlyeq \mathfrak{s}$ where $\mathfrak{s} := \mathfrak{tt}(\mathfrak{ttt})$ and $\mathfrak{s}' := \mathfrak{ttt}(\mathfrak{ttt})$. This can be seen by noticing that $\mathfrak{t}\,x_1x_2 \preccurlyeq x_2x_2$. Figure 1B shows a part of the rewrite graph of $\mathcal{C}$.

2.2.2 Properties of CLS

Let us state some properties of CLS related to their properties of confluence, termination, and local finiteness.

Proposition 2.2.1. *Any CLS is confluent.*

Proof. By definition, the underlying ATRS of any CLS is orthogonal (Baader and Nipkow 1998, Chapter 6). The statement follows from the fact that all orthogonal rewrite systems are confluent (Rosen 1973). □

Lemma 2.2.2. *Let $\mathcal{C} := (\mathfrak{G}, \to)$ be a CLS. If $\mathcal{C}$ is nonerasing, then for any $n \geqslant 0$ and $\mathfrak{t} \in \mathfrak{T}_n(\mathfrak{G})$, $[\mathfrak{t}]_{\equiv} \subseteq \mathfrak{T}_n(\mathfrak{G})$.*

Proof. Assume that $\mathfrak{t}$ and $\mathfrak{t}'$ are two $\mathfrak{G}$-terms such that $\mathfrak{t} \Rightarrow \mathfrak{t}'$. From the definition of $\Rightarrow$ from $\to$ provided by (2.1.7), the fact that all constants of $\mathfrak{G}$ are nonerasing implies that for any $x_i \in \mathbb{X}$, $\deg_{x_i}(\mathfrak{t}) \geqslant 1$ if and only if $\deg_{x_i}(\mathfrak{t}') \geqslant 1$. Therefore, both $\mathfrak{t}$ and $\mathfrak{t}'$ belong to $\mathfrak{T}_n(\mathfrak{G})$ where n is an integer nonsmaller than the greatest index of the variables appearing in $\mathfrak{t}$ and $\mathfrak{t}'$. Since $\equiv$ is the reflexive, symmetric, and transitive closure of $\Rightarrow$, the result follows. □

Proposition 2.2.3. *Let $\mathcal{C} := (\mathfrak{G}, \to)$ be a CLS. If $\mathcal{C}$ is hierarchical, then $\mathcal{C}$ is locally finite and all the $\mathfrak{G}$-terms of a same connected component of $\mathrm{G}_{\mathcal{C}}$ have the same height.*

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

Proof. Assume that t and t' are two $\mathfrak{G}$-terms such that $t \Rightarrow t'$. From the definition of $\Rightarrow$ from $\rightarrow$ provided by (2.1.7), the fact that all constants of $\mathfrak{G}$ are hierarchical implies that $\mathrm{ht}(t) = \mathrm{ht}(t')$. Moreover, since $\mathcal{C}$ is also nonerasing, by Lemma 2.2.2, for any $x_i \in \mathbb{X}$, $\deg_{x_i}(t) \geqslant 1$ if and only if $\deg_{x_i}(t') \geqslant 1$. Observe moreover that for any $n \geqslant 0$, the number of terms of the same height in $\mathfrak{T}_n(\mathfrak{G})$ is finite. Since $\equiv$ is the reflexive, symmetric, and transitive closure of $\Rightarrow$, and since two terms t and t' are $\equiv$-equivalent if and only if t and t' belong to the same connected component of $\mathrm{G}(\mathcal{C})$, the result follows. □

Let us use these results to state some properties of a CLS $\mathcal{C}$. By Proposition 2.2.1 and Bezem et al. (2003, Theorem 1.2.2), each $\equiv$-equivalence class of terms of $\mathcal{C}$ admits at most one normal form. If $\mathcal{C}$ is hierarchical and has the poset property, it follows by Proposition 2.2.3 that for each term t of $\mathcal{C}$, the subposet on $[t]_{\equiv}$ of $\mathcal{P}_{\mathcal{C}}$ admits exactly one maximal element. If additionally $\mathcal{C}$ is rooted, then for each term of $\mathcal{C}$, the subposet on $[t]_{\equiv}$ of $\mathcal{P}_{\mathcal{C}}$ admits exactly one minimal element. In this case, there is a one-to-one correspondence between the set of the minimal terms and the set of the maximal terms of $\mathcal{C}$ and this correspondence preserves the $\equiv$-equivalence classes.

2.2.3 Models

Let $\mathcal{C} := (\mathfrak{G}, \rightarrow)$ be a CLS. A *model* of $\mathcal{C}$ is an algebra over the abstract clone defined as the quotient of $\mathfrak{T}(\mathfrak{G})$ by the clone congruence $\equiv$ (see for instance Taylor 1993 for a general description of algebras over abstract clones). In more concrete terms, for this particular case of abstract clone, this is the data of a magma $(\mathcal{M}, \star)$ such that $\mathfrak{G} \subseteq \mathcal{M}$ and, for each rule $\mathbf{X}\, x_1 \ldots x_n \rightarrow t_{\mathbf{X}}$, the axiom $(\ldots((\mathbf{X} \star x_1) \star x_2) \ldots) \star x_n = e_{\mathbf{X}}$ holds, where $e_{\mathbf{X}}$ is the expression having $t_{\mathbf{X}}$ as syntax tree. For instance, a model of the CLS on the constants $\mathbf{K}$ and $\mathbf{S}$ is a set $\mathcal{M}$ containing $\mathbf{K}$ and $\mathbf{S}$, and endowed with an operation $\star : \mathcal{M}^2 \rightarrow \mathcal{M}$ satisfying $(\mathbf{K} \star x_1) \star x_2 = x_1$ and $((\mathbf{S} \star x_1) \star x_2) \star x_3 = (x_1 \star x_3) \star (x_2 \star x_3)$. Such structures are known as *combinatory algebras* (Hindley and Seldin 2008).

The *trivial model*) of $\mathcal{C}$ is the magma $(\mathcal{M}, \star)$ where $\mathcal{M}$ is the set $\mathfrak{T}(\mathfrak{G})/_{\equiv}$ and, for any $[t_1]_{\equiv}, [t_2]_{\equiv} \in \mathcal{M}$, $[t_1]_{\equiv} \star [t_2]_{\equiv}$ is the $\equiv$-equivalence class of $t_1 t_2$ where t_1 is any element of $[t_1]_{\equiv}$ and t_2 is any element of $[t_2]_{\equiv}$. Moreover, when $\mathcal{C}$ is nonerasing, for any $n \geqslant 0$, the *n-trivial model* of $\mathcal{C}$ is the trivial model restrained to $\equiv$-equivalence classes of terms of $\mathfrak{T}_n(\mathfrak{G})$. Thanks to Lemma 2.2.2, the n-trivial model is well-defined.

When $\mathcal{C}$ is hierarchical and has the poset property, as already noticed, each $\equiv$-equivalence class contains exactly one maximal element. Therefore, the set $\mathcal{M}^{\max}$ of the maximal terms of $\mathfrak{T}(\mathfrak{G})$ is a set of representatives of the set of the $\equiv$-equivalence classes and forms a model of $\mathcal{C}$ isomorphic to the trivial model. When additionally $\mathcal{C}$ is rooted, as already noticed, each $\equiv$-equivalence class contains exactly one minimal element. For this reason, the set $\mathcal{M}^{\min}$ of the minimal terms of $\mathfrak{T}(\mathfrak{G})$ is a set of representatives of the set of the $\equiv$-equivalence classes and forms a model of $\mathcal{C}$ isomorphic to the previous ones. Here also, for any $n \geqslant 0$, we define the model $\mathcal{M}_n^{\max}$ (resp. $\mathcal{M}_n^{\min}$) as the restriction of $\mathcal{M}^{\max}$ (resp. $\mathcal{M}^{\min}$) on $\mathfrak{T}_n(\mathfrak{G})$. The interest of these two models $\mathcal{M}^{\max}$ and $\mathcal{M}^{\min}$ relies on considerations about algorithmic complexity for the computation of the product $t_1 \star t_2$ where t_1 and t_2 are two terms of the models.

3. The Mockingbird Lattice

This central section of this work concerns the study of the poset associated with the CLS containing the basic combinator $\mathbf{M}$. We shall prove that this CLS has the poset property, is rooted, and has also the lattice property by introducing new lattices on some kind of treelike structures, called duplicative forests.

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

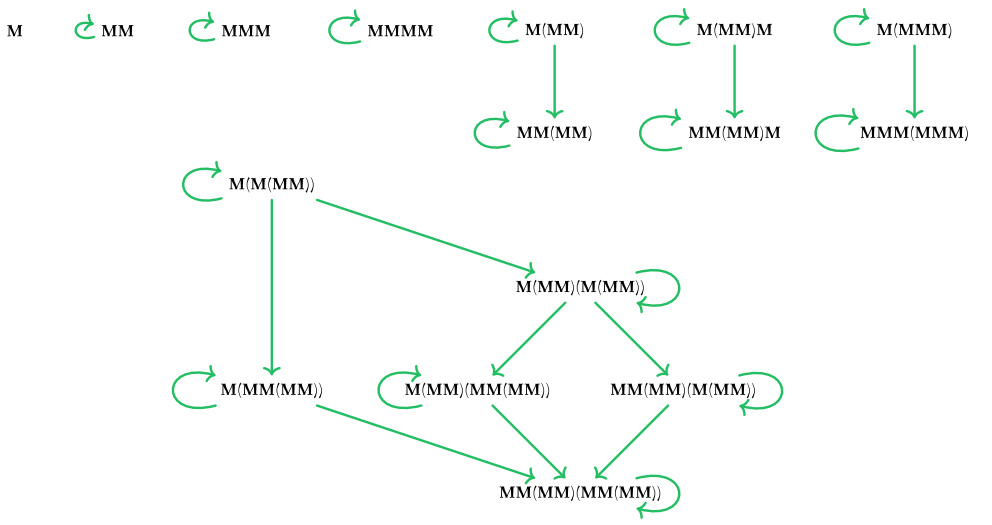


Figure 2. The fragment of the rewrite graph of $\mathcal{C}$ restrained to the terms reachable from closed terms of degrees 3 or less.

3.1 The combinator **M** *and its poset*

Let the CLS $\mathcal{C} := (\mathfrak{G}, \to)$ such that $\mathfrak{G} := \{\mathbf{M}\}$ and $t_{\mathbf{M}} = x_1x_1$. Since **M** is the Mockingbird basic combinator, we call $\mathcal{C}$ the *Mockingbird CLS*. Observe that **M** is hierarchical so that $\mathcal{C}$ satisfies the properties stated by Proposition 2.2.3 and is in particular locally finite. From now, we shall simply write G instead of $G_{\mathcal{C}}$. Figure 2 provides an example of a fragment of G.

3.1.1 First properties

Let us present some first properties of $\mathcal{C}$. We begin by providing a recursive description of the rewrite relation $\Rightarrow$.

Lemma 3.1.1. *For any* $t, t' \in \mathfrak{T}(\mathfrak{G})$*, we have* $t \Rightarrow t'$ *if and only if* $t = t_1t_2$, $t' = t'_1t'_2$ *with* $t_1, t_2, t'_1, t'_2 \in \mathfrak{T}(\mathfrak{G})$*, and at least one of the following assertions holds:*

(i) $t_1 \Rightarrow t'_1$ *and* $t_2 = t'_2$*;*
(ii) $t_2 \Rightarrow t'_2$ *and* $t_1 = t'_1$*;*
(iii) $t_1 = \mathbf{M}$ *and* $t_2 = t'_1 = t'_2$*.*

Proof. This is a direct consequence of the definitions of the context closure $\Rightarrow$ from the rewrite relation $\to$ and of the term $t_{\mathbf{M}}$. □

Proposition 3.1.2. *The CLS* $\mathcal{C}$

(i) is locally finite;
(ii) has the poset property;
(iii) is rooted.

Proof. First, since **M** is hierarchical, by Proposition 2.2.3, $\mathcal{C}$ is locally finite. Therefore, (i) holds.

Let us prove that $\mathcal{C}$ has the poset property. For this, let $\theta : \mathfrak{T}(\mathfrak{G}) \to \mathbb{Z}^2$ be the map defined for any $t \in \mathfrak{T}(\mathfrak{G})$ by $\theta(t) := \big(\deg(t), -\deg_{\mathbf{M}}(t)\big)$. We denote by $\leqslant$ the lexicographic order on $\mathbb{Z}^2$ and by $\dot{+}$ the pointwise addition on $\mathbb{Z}^2$. Let us prove by structural induction on $\mathfrak{T}(\mathfrak{G})$ that for

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

any $\mathfrak{G}$-terms t and t', $t \neq t'$ and $t \Rightarrow t'$ implies $\theta(t) < \theta(t')$. By Lemma 3.1.1, we have $t = t_1 t_2$ and $t' = t'_1 t'_2$ for some $\mathfrak{G}$-terms t_1, t_2, t'_1, and t'_2, and we have the three following cases. First, if $t_1 \Rightarrow t'_1$ and $t_2 = t'_2$, then $\theta(t) = \theta(t_1) \dot{+} \theta(t_2) \dot{+} (1,0)$ and $\theta(t') = \theta(t'_1) \dot{+} \theta(t_2) \dot{+} (1,0)$. Since, by induction hypothesis, $\theta(t_1) < \theta(t'_1)$, we have $\theta(t) < \theta(t')$. Second, if $t_2 \Rightarrow t'_2$ and $t_1 = t'_1$, the same arguments as the previous ones (by interchanging the indices 1 and 2 in the concerned terms) apply. Finally, if $t_1 = \mathbf{M}$ and $t_2 = t'_1 = t'_2$, then $\theta(t) = \theta(t_2) \dot{+} (1,-1)$ and $\theta(t') = \theta(t_2) \dot{+} \theta(t_2) \dot{+} (1,0)$. Since $t \neq t'$, we have $t_2 \neq \mathbf{M}$ so that $\theta(t) < \theta(t')$. This implies that $\preccurlyeq$ is antisymmetric, so that (ii) holds.

To prove that $\mathcal{C}$ is rooted, we consider the rewrite relation $\to'$ obtained by inverting $\to$. Thus, $\to'$ satisfies $x_1 x_1 \to' \mathbf{M} x_1$. This ATRS $(\mathfrak{T}(\mathfrak{G}), \to')$ does not admit any overlapping term (see Bezem et al. 2003, Chapter 2). Therefore, it is confluent. A consequence of (i), is that this ATRS has only finite $\equiv'$-equivalence classes. By using additionally (Bezem et al. 2003, Theorem 1.2.2), each $\equiv'$-equivalence class admits exactly one term t such that for any $t' \in [t]_{\equiv'}$, $t' \preccurlyeq' t$. This implies that in $\mathcal{C}$, each $\equiv$-equivalence class admits exactly one minimal term. Therefore, (iii) holds. □

By Proposition 3.1.2, $\mathcal{P}_{\mathcal{C}}$ is a well-defined poset, called *Mockingbird poset*. From now, we shall simply write $\mathcal{P}$ instead of $\mathcal{P}_{\mathcal{C}}$. We have the following recursive description of $\preccurlyeq$.

Lemma 3.1.3. *For any $t, t' \in \mathfrak{T}(\mathfrak{G})$, we have $t \preccurlyeq t'$ if and only if at least one of the following assertions holds:*

(i) $t = \mathbf{M} = t'$;
(ii) $t = x_i = t'$ *for an* $x_i \in \mathbb{X}$;
(iii) $t = t_1 t_2$, $t' = t'_1 t'_2$, $t_1 \preccurlyeq t'_1$, *and* $t_2 \preccurlyeq t'_2$ *where* $t_1, t'_1, t_2, t'_2 \in \mathfrak{T}(\mathfrak{G})$;
(iv) $t = \mathbf{M} t_2$, $t' = t'_1 t'_2$, $t_2 \preccurlyeq t'_1$, *and* $t_2 \preccurlyeq t'_2$ *where* $t'_1, t_2, t'_2 \in \mathfrak{T}(\mathfrak{G})$.

Proof. This is a direct consequence of Lemma 3.1.1 and of the fact that $\preccurlyeq$ is the reflexive and transitive closure of $\Rightarrow$. □

Proposition 3.1.4. *Let* $t \in \mathfrak{T}(\mathfrak{G})$.

(i) The term t is a maximal element of $\mathcal{P}$ if and only if t avoids $\mathbf{M}(x_1 x_2)$ and suffix-avoids $\mathbf{M} x_1$.
(ii) The term t is a minimal element of $\mathcal{P}$ if and only if t avoids $(x_1 x_2)(x_1 x_2)$ and suffix-avoids $x_1 x_1$.

Proof. The term t is maximal in $\mathcal{P}$ if and only if $t \Rightarrow t'$ implies that $t' = t$. By Lemma 3.1.1, this is equivalent to the fact t is maximal in $\mathcal{P}$ if and only if when t has an internal node having $\mathbf{M}$ as left subtree, this internal node admits necessarily $\mathbf{M}$ as right subtree. Therefore, (i) holds.

The term t is minimal in $\mathcal{P}$ if and only if $t' \Rightarrow t$ implies that $t' = t$. By Lemma 3.1.1, this is equivalent to the fact t is minimal in $\mathcal{P}$ if and only if when t has an internal node such that its two children are equal, these two children are necessarily equal to $\mathbf{M}$. Therefore, (ii) holds. □

3.1.2 Some models

We begin this discussion about models of $\mathcal{C}$ by a very simple observation: any idempotent monoid $(\mathcal{M}, \cdot, e)$ is a model of $\mathcal{C}$. Indeed, by identifying $\mathbf{M}$ with e (the unit of $\mathcal{M}$) and by identifying $\star$ with $\cdot$, we have for any $x \in \mathcal{M}$, $\mathbf{M} \star x = e \cdot x = x = x \cdot x = x \star x$.

Besides, since $\mathcal{C}$ is hierarchical and has, by Proposition 3.1.2, the poset property, this CLS admits the model $\mathcal{M}^{\max}$ described in Section 2.2.3. More specifically, this model is such that $\mathcal{M}^{\max}$ is the

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

set of the maximal elements of $\mathcal{P}$, and, by Proposition 3.1.4, for any $t_1, t_2 \in \mathcal{M}^{\max}$,

$$t_1 \star t_2 = \begin{cases} t_2 t_2 & \text{if } t_1 = \mathbf{M}, \\ t_1 t_2 & \text{otherwise.} \end{cases} \tag{3.1.1}$$

Since by Proposition 3.1.2, $\mathcal{C}$ is also rooted, $\mathcal{C}$ admits the model $\mathcal{M}^{\min}$ described in Section 2.2.3. This model is such that $\mathcal{M}^{\min}$ is the set of the minimal elements of $\mathcal{P}$, and, by Proposition 3.1.4, for any $t_1, t_2 \in \mathcal{M}^{\min}$,

$$t_1 \star t_2 = \begin{cases} \mathbf{M} t_2 & \text{if } t_1 = t_2, \\ t_1 t_2 & \text{otherwise.} \end{cases} \tag{3.1.2}$$

By representing $\mathfrak{G}$-terms directly has binary trees, the computations in $\mathcal{M}^{\max}$ have a better time complexity than in $\mathcal{M}^{\min}$ because to compute $t_1 \star t_2$ in $\mathcal{M}^{\min}$, we need to decide if t_1 and t_2 are equal. Nevertheless, in return, the term $t_1 \star t_2$ produced by a computation in $\mathcal{M}^{\min}$ has a number of internal nodes equal as or smaller than the analogous computation in $\mathcal{M}^{\max}$.

3.2 Duplicative lattices

We introduce here duplicative forests and a partial order relation on these objects. We show that this partial relation endow all intervals of this poset with lattice structures.

3.2.1 Duplicative forests

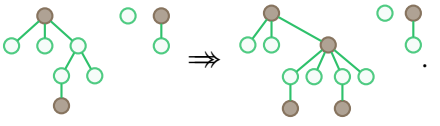
A *duplicative tree* is a planar rooted tree such that each internal node is either a *black node* ● or a *white node* ○. A *duplicative forest* is a word $\mathfrak{f} = \mathfrak{f}(1) \dots \mathfrak{f}(\ell)$, $\ell \geqslant 0$, of duplicative trees. In particular, the empty word ϵ is the *empty forest*. We denote by $\mathcal{D}$ (resp. $\mathcal{D}^*$) the set of such trees (resp. forests). The *height* $\mathrm{ht}(\mathfrak{f})$ of $\mathfrak{f}$ is the number of internal nodes in a longest path following edges connecting a node to one of its child. In the sequel, we shall write expressions using some occurrences of □: each such expression denotes the two expressions obtained by replacing simultaneously all □ either by ○ or by ●. The *grafting product* is the unary operation □ on $\mathcal{D}^*$ such that for any $\mathfrak{f} \in \mathcal{D}^*$, $\square(\mathfrak{f})$ is the duplicative tree obtained by grafting the roots of the duplicative trees of $\mathfrak{f}$ on a common root node □. The *concatenation product* is the binary operation $\cdot$ on $\mathcal{D}^*$ such that for any $\mathfrak{f}_1, \mathfrak{f}_2 \in \mathcal{D}^*$, $\mathfrak{f}_1 \cdot \mathfrak{f}_2$ is the duplicative forest made of the trees of $\mathfrak{f}_1$ and then of the trees of $\mathfrak{f}_2$.

3.2.2 A poset on duplicative forests

Let $\twoheadrightarrow$ be the binary relation on $\mathcal{D}^*$ defined recursively as follows. For any $\mathfrak{f}, \mathfrak{f}' \in \mathcal{D}^*$, we have $\mathfrak{f} \twoheadrightarrow \mathfrak{f}'$ if $\mathfrak{f}$ and $\mathfrak{f}'$ have the same length $\ell \geqslant 1$ and

- either $\ell = 1$, $\mathfrak{f} = \circ(\mathfrak{g})$ and $\mathfrak{f}' = \bullet(\mathfrak{g} \cdot \mathfrak{g})$ where $\mathfrak{g} \in \mathcal{D}^*$;
- or $\ell = 1$, $\mathfrak{f} = \square(\mathfrak{g})$, $\mathfrak{f}' = \square(\mathfrak{g}')$ where $\mathfrak{g}, \mathfrak{g}' \in \mathcal{D}^*$ and $\mathfrak{g} \twoheadrightarrow \mathfrak{g}'$;
- or $\ell \geqslant 2$ and there is a $j \in [\ell]$ such that $\mathfrak{f}(j) \twoheadrightarrow \mathfrak{f}'(j)$, and for all $i \in [\ell] \setminus \{j\}$, $\mathfrak{f}(i) = \mathfrak{f}'(i)$.

We observe that $\mathfrak{f} \twoheadrightarrow \mathfrak{f}'$ if and only if $\mathfrak{f}'$ can be obtained from $\mathfrak{f}$ by selecting a white node of $\mathfrak{f}$, by turning it into black, and by duplicating its sequence of descendants. For instance, we have



. (3.2.1)

Observe also that in this case, there are more black nodes in $\mathfrak{f}'$ than in $\mathfrak{f}$. Hence, the reflexive and transitive closure $\ll$ of $\twoheadrightarrow$ is antisymmetric so that $(\mathcal{D}^*, \ll)$ is a poset. We call this poset the

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

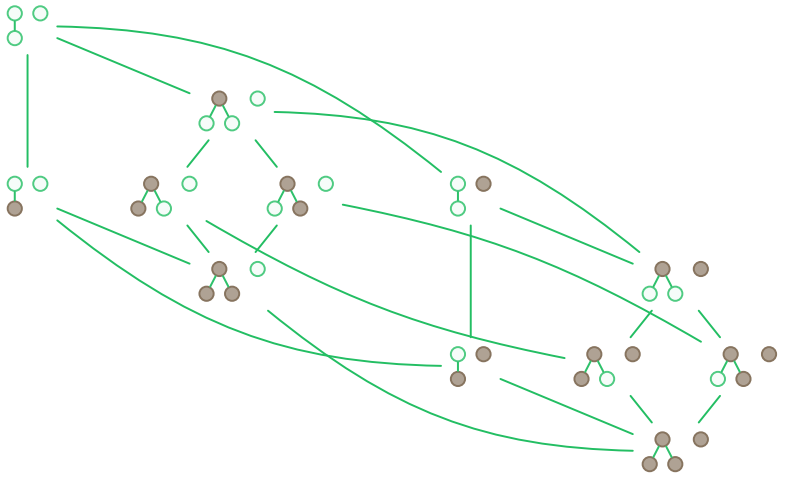


Figure 3. The Hasse diagram of a maximal interval of the duplicative forest poset.

duplicative forest poset. For any $\mathfrak{f} \in \mathcal{D}^*$, we denote by $\mathcal{D}^*(\mathfrak{f})$ the subposet of the duplicative forest poset on the set $\{\mathfrak{f}' \in \mathcal{D}^* : \mathfrak{f} \ll \mathfrak{f}'\}$. Figure 3 shows the Hasse diagram of the poset $\mathcal{D}^*(\mathfrak{f})$ where $\mathfrak{f}$ is a certain duplicative forest. According to this Hasse diagram, the duplicative forest poset is not graded. We have the following recursive description of $\ll$.

Lemma 3.2.1. *For any two duplicative forests $\mathfrak{f}$ and $\mathfrak{f}'$, we have $\mathfrak{f} \ll \mathfrak{f}'$ if and only if $\mathfrak{f}$ and $\mathfrak{f}'$ have the same length $\ell \geqslant 0$ and one of the following assertions holds:*

(i) $\ell = 0$;
(ii) $\ell = 1$, $\mathfrak{f} = \circ(\mathfrak{g})$ *and* $\mathfrak{f}' = \bullet(\mathfrak{g}' \centerdot \mathfrak{g}'')$ *where* $\mathfrak{g}, \mathfrak{g}', \mathfrak{g}'' \in \mathcal{D}^*$, $\mathfrak{g} \ll \mathfrak{g}'$, *and* $\mathfrak{g} \ll \mathfrak{g}''$;
(iii) $\ell = 1$, $\mathfrak{f} = \square(\mathfrak{g})$, $\mathfrak{f}' = \square(\mathfrak{g}')$ *where* $\mathfrak{g}, \mathfrak{g}' \in \mathcal{D}^*$ *and* $\mathfrak{g} \ll \mathfrak{g}'$;
(iv) $\ell \geqslant 2$ *and* $\mathfrak{f}(i) \ll \mathfrak{f}'(i)$ *for all* $i \in [\ell]$.

Proof. This is a direct consequence of the definitions of $\twoheadrightarrow$ and of $\ll$. □

A first consequence of Lemma 3.2.1 is that for any $\mathfrak{f}, \mathfrak{f}' \in \mathcal{D}^*$, $\mathfrak{f} \ll \mathfrak{f}'$ implies $\mathrm{ht}(\mathfrak{f}) = \mathrm{ht}(\mathfrak{f}')$.

3.2.3 Lattices on duplicative forests

We use in the sequel the usual notions and notations about lattices (see for instance Stanley 2011). Let $\wedge$ and $\vee$ be the two binary, commutative, and associative partial operations on $\mathcal{D}^*$ defined recursively, for any $\ell \geqslant 0$, $\mathfrak{f}_1, \ldots, \mathfrak{f}_\ell \in \mathcal{D}$, $\mathfrak{f}'_1, \ldots, \mathfrak{f}'_\ell \in \mathcal{D}$, and $\mathfrak{f}, \mathfrak{f}', \mathfrak{f}'' \in \mathcal{D}^*$, by

$$\mathfrak{f}_1 \centerdot \cdots \centerdot \mathfrak{f}_\ell \wedge \mathfrak{f}'_1 \centerdot \cdots \centerdot \mathfrak{f}'_\ell := (\mathfrak{f}_1 \wedge \mathfrak{f}'_1) \centerdot \cdots \centerdot (\mathfrak{f}_\ell \wedge \mathfrak{f}'_\ell), \tag{3.2.2a}$$

$$\square(\mathfrak{f}) \wedge \square(\mathfrak{f}') := \square(\mathfrak{f} \wedge \mathfrak{f}'), \tag{3.2.2b}$$

$$\circ(\mathfrak{f}) \wedge \bullet(\mathfrak{f}' \centerdot \mathfrak{f}'') := \circ(\mathfrak{f} \wedge \mathfrak{f}' \wedge \mathfrak{f}'') \tag{3.2.2c}$$

and

$$\mathfrak{f}_1 \centerdot \cdots \centerdot \mathfrak{f}_\ell \vee \mathfrak{f}'_1 \centerdot \cdots \centerdot \mathfrak{f}'_\ell := (\mathfrak{f}_1 \vee \mathfrak{f}'_1) \centerdot \cdots \centerdot (\mathfrak{f}_\ell \vee \mathfrak{f}'_\ell), \tag{3.2.3a}$$

$$\square(\mathfrak{f}) \vee \square(\mathfrak{f}') := \square(\mathfrak{f} \vee \mathfrak{f}'), \tag{3.2.3b}$$

$$\circ(\mathfrak{f}) \vee \bullet(\mathfrak{f}' \centerdot \mathfrak{f}'') := \bullet((\mathfrak{f} \vee \mathfrak{f}') \centerdot (\mathfrak{f} \vee \mathfrak{f}'')). \tag{3.2.3c}$$

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

Proposition 3.2.2. *Given a duplicative forest* $\mathfrak{f}$*, the poset* $\mathcal{D}^*(\mathfrak{f})$ *is a lattice for the operations* $\wedge$ *and* $\vee$*.*

Proof. This follows by structural induction on $\mathfrak{f}$ by establishing the fact that for any $\mathfrak{f}', \mathfrak{f}'' \in \mathcal{D}^*(\mathfrak{f})$, the duplicative forest $\mathfrak{f}' \wedge \mathfrak{f}''$ (resp. $\mathfrak{f}' \vee \mathfrak{f}''$) is well-defined and is the meet (resp. the join) of $\mathfrak{f}'$ and $\mathfrak{f}''$. This uses Lemma 3.2.1 in order to describe $\mathfrak{f}'$ and $\mathfrak{f}''$ from $\mathfrak{f}$ knowing that $\mathfrak{f} \ll \mathfrak{f}'$ and $\mathfrak{f} \ll \mathfrak{f}''$. □

By Proposition 3.2.2, for any $\mathfrak{f} \in \mathcal{D}^*$, each poset $\mathcal{D}^*(\mathfrak{f})$ is a lattice. We call it the *duplicative forest lattice* of $\mathfrak{f}$.

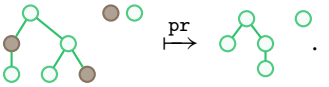
Let $\mathtt{pr} : \mathcal{D}^* \to \mathcal{D}^*$ be the map defined recursively, for any $\ell \geqslant 0$, $\mathfrak{f}_1, \ldots, \mathfrak{f}_\ell \in \mathcal{D}^*$, and $\mathfrak{f} \in \mathcal{D}^*$, by

$$\mathtt{pr}(\mathfrak{f}_1 \bullet \cdots \bullet \mathfrak{f}_\ell) := \mathtt{pr}(\mathfrak{f}_1) \bullet \cdots \bullet \mathtt{pr}(\mathfrak{f}_\ell), \tag{3.2.4a}$$

$$\mathtt{pr}(\circ(\mathfrak{f})) := \circ(\mathtt{pr}(\mathfrak{f})), \tag{3.2.4b}$$

$$\mathtt{pr}(\bullet(\mathfrak{f})) := \mathtt{pr}(\mathfrak{f}). \tag{3.2.4c}$$

The duplicative forest $\mathtt{pr}(\mathfrak{f})$ is the *pruning* of $\mathfrak{f}$. This forest has by construction no occurrence of any $\bullet$. For instance,



(3.2.5)

Lemma 3.2.3. *For any* $\mathfrak{f} \in \mathcal{D}^*$*, the posets* $\mathcal{D}^*(\mathfrak{f})$ *and* $\mathcal{D}^*(\mathtt{pr}(\mathfrak{f}))$ *are isomorphic.*

Proof. This follows by structural induction on $\mathfrak{f}$ by establishing the fact that $\mathtt{pr}$ is a one-to-one correspondence between the sets $\mathcal{D}^*(\mathfrak{f})$ and $\mathcal{D}^*(\mathtt{pr}(\mathfrak{f}))$ and that $\mathtt{pr}$ is an order isomorphism. This uses Lemma 3.2.1 in order to have a recursive description of the order relation $\ll$. □

For any $d \geqslant 0$, the *d-ladder* is the duplicative forest $\mathfrak{l}_d$ defined recursively by $\mathfrak{l}_0 := \epsilon$ and, for any $d \geqslant 1$, by $\mathfrak{l}_d := \circ(\mathfrak{l}_{d-1})$. Let also

$$\mathcal{L} := \bigcup_{d \geqslant 0} \mathcal{D}^*(\mathfrak{l}_d). \tag{3.2.6}$$

By definition, $\mathcal{L}$ is the set of the duplicative forests that are equal as or greater than a d-ladder for a $d \geqslant 0$.

Lemma 3.2.4. *For any* $\mathfrak{f} \in \mathcal{D}^*$*, there exists* $\mathfrak{g} \in \mathcal{L}$ *such that* $\mathtt{pr}(\mathfrak{f}) = \mathtt{pr}(\mathfrak{g})$*.*

Proof. This follows by structural induction on $\mathfrak{f}$ by showing that there is a $d \geqslant 0$ such that $\mathfrak{l}_d \ll \mathfrak{g}$ and $\mathtt{pr}(\mathfrak{f}) = \mathtt{pr}(\mathfrak{g})$. This uses Lemma 3.2.1. □

An important consequence of Lemmas 3.2.3 and 3.2.4 is that for any $\mathfrak{f} \in \mathcal{D}^*$, there exists $d \geqslant 0$ such that $\mathcal{D}^*(\mathfrak{f})$ is isomorphic as a poset to a maximal interval of $\mathcal{D}^*(\mathfrak{l}_d)$.

3.3 Mockingbird lattices

We show here that each subposet $\mathcal{P}(\mathfrak{t})$, $\mathfrak{t} \in \mathfrak{T}(\mathfrak{G})$, of $\mathcal{P}$ is a lattice. This is based on a poset isomorphism between $\mathcal{P}(\mathfrak{t})$ and an interval of a lattice of duplicative forests. We also define for each $d \geqslant 0$ the Mockingbird lattice of order d as a particular maximal interval of $\mathcal{P}$.

3.3.1 From terms to duplicative forests

Let $\mathtt{fr} : \mathfrak{T}(\mathfrak{G}) \to \mathcal{D}^*$ be the map defined recursively, for any $\mathtt{x}_i \in \mathbb{X}$ and $\mathfrak{t}, \mathfrak{t}' \in \mathfrak{T}(\mathfrak{G})$, by

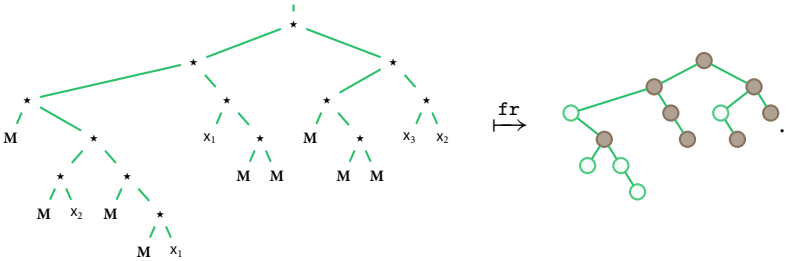
$$\mathtt{fr}(\mathtt{x}_i) := \epsilon, \tag{3.3.1a}$$

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

$$\mathtt{fr}(\mathbf{M}) := \epsilon, \tag{3.3.1b}$$

$$\mathtt{fr}(\mathfrak{t}\mathfrak{t}') := \begin{cases} \circ(\mathtt{fr}(\mathfrak{t}')) & \text{if } \mathfrak{t} = \mathbf{M} \text{ and } \mathfrak{t}' \neq \mathbf{M}, \\ \bullet(\mathtt{fr}(\mathfrak{t}) \mathbin{.} \mathtt{fr}(\mathfrak{t}')) & \text{otherwise.} \end{cases} \tag{3.3.1c}$$

For instance,



(3.3.2)

Intuitively, $\mathtt{fr}(\mathfrak{t})$ is the duplicative forest obtained from $\mathfrak{t}$ by the following process: replace each internal node of $\mathfrak{t}$ having $\mathbf{M}$ as left child and having a right child different from $\mathbf{M}$ by a $\circ$, replace each other internal node of $\mathfrak{t}$ by a $\bullet$, and remove all the leaves. Immediately from the definition, we observe that this map is not injective. Moreover, again directly from the definition, any duplicative forest $\mathfrak{f}$ in the image of $\mathtt{fr}$ is such that each white node of $\mathfrak{f}$ has no more than one child.

Lemma 3.3.1. *For any $\mathfrak{t} \in \mathfrak{T}(\mathfrak{G})$, the restriction of the map fr on the domain $\mathcal{P}(\mathfrak{t})$ is injective.*

Proof. This follows by structural induction on $\mathfrak{t}$ by establishing the fact that for any $\mathfrak{s}, \mathfrak{s}' \in \mathcal{P}(\mathfrak{t})$, $\mathtt{fr}(\mathfrak{s}) = \mathtt{fr}(\mathfrak{s}')$ implies $\mathfrak{s} = \mathfrak{s}'$. This uses in particular Lemma 3.1.3 in order to describe $\mathfrak{s}$ and $\mathfrak{s}'$ from $\mathfrak{t}$ knowing that $\mathfrak{t} \preccurlyeq \mathfrak{s}$ and $\mathfrak{t} \preccurlyeq \mathfrak{s}'$. □

Lemma 3.3.2. *Let $\mathfrak{t} \in \mathfrak{T}(\mathfrak{G})$ and $\mathfrak{s}, \mathfrak{s}' \in \mathfrak{T}(\mathfrak{t})$ such that $\mathfrak{s} \neq \mathfrak{s}'$. We have $\mathfrak{s} \Rightarrow \mathfrak{s}'$ if and only if $fr(\mathfrak{s}) \twoheadrightarrow fr(\mathfrak{s}')$.*

Proof. This follows by structural induction on $\mathfrak{t}$. This uses in particular the definition of the binary relation $\twoheadrightarrow$ on $\mathcal{D}^*$, Lemma 3.1.3 in order to describe $\mathfrak{s}$ and $\mathfrak{s}'$ from $\mathfrak{t}$ knowing that $\mathfrak{t} \preccurlyeq \mathfrak{s}$ and $\mathfrak{t} \preccurlyeq \mathfrak{s}'$, and Lemma 3.1.1 in order to describe $\mathfrak{s}'$ from $\mathfrak{s}$ knowing that $\mathfrak{s} \neq \mathfrak{s}'$ and $\mathfrak{s} \Rightarrow \mathfrak{s}'$. □

Proposition 3.3.3. *For any $\mathfrak{t} \in \mathfrak{T}(\mathfrak{G})$, the posets $\mathcal{P}(\mathfrak{t})$ and $\mathcal{D}^*(fr(\mathfrak{t}))$ are isomorphic.*

Proof. Let $D := \{\mathtt{fr}(\mathfrak{t}) : \mathfrak{t} \in \mathcal{P}(\mathfrak{t})\}$. By Lemma 3.3.1, the sets $\mathcal{P}(\mathfrak{t})$ and D are in one-to-one correspondence. Moreover, by Lemma 3.3.2, we have that $D = \mathcal{D}^*(\mathtt{fr}(\mathfrak{t}))$ and that $\mathtt{fr}$ is an order isomorphism between $\mathcal{P}(\mathfrak{t})$ and $\mathcal{D}^*(\mathtt{fr}(\mathfrak{t}))$. □

Figure 4 shows the poset $\mathcal{P}(\mathfrak{t})$ for a $\mathfrak{G}$-term $\mathfrak{t}$ and the isomorphic poset $\mathcal{D}^*(\mathtt{fr}(\mathfrak{t}))$.

Theorem 3.3.4. *For any $\mathfrak{t} \in \mathfrak{T}(\mathfrak{G})$, the poset $\mathcal{P}(\mathfrak{t})$ is a finite lattice.*

Proof. By Proposition 3.3.3, $\mathcal{P}(\mathfrak{t})$ is isomorphic as a poset to $\mathcal{D}^*(\mathtt{fr}(\mathfrak{t}))$. Hence, and since by Proposition 3.2.2, $\mathcal{D}^*(\mathtt{fr}(\mathfrak{t}))$ is a lattice, $\mathcal{P}(\mathfrak{t})$ also is. The finiteness of $\mathcal{P}(\mathfrak{t})$ is a consequence of the fact that by Proposition 3.1.2, $\mathcal{C}$ is locally finite. □

The *Mockingbird lattice* of order $d \geqslant 0$ is the lattice $\mathrm{M}(d) := \mathcal{P}(\mathfrak{r}_d)$ where $\mathfrak{r}_d$ is the term recursively defined by $\mathfrak{r}_0 := \mathbf{M}$ and, for any $d \geqslant 1$, by $\mathfrak{r}_d := \mathbf{M}\mathfrak{r}_{d-1}$. Figure 5 shows the Hasse diagrams of the first Mockingbird lattices.

Theorem 3.3.5. *For any $\mathfrak{f} \in \mathcal{D}^*$, the poset $\mathcal{D}^*(\mathfrak{f})$ is isomorphic to a maximal interval of a Mockingbird lattice.*

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

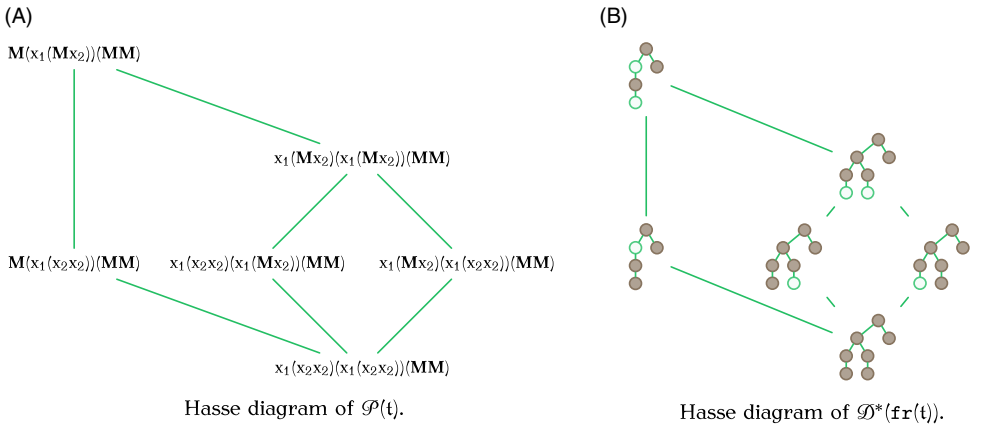


Figure 4. A maximal interval of the Mockingbird poset from the term $\mathfrak{t} := \mathbf{M}(\mathsf{x}_1(\mathbf{M}\mathsf{x}_2))(\mathbf{MM})$ and its realization as a maximal interval in the duplicative forest poset.

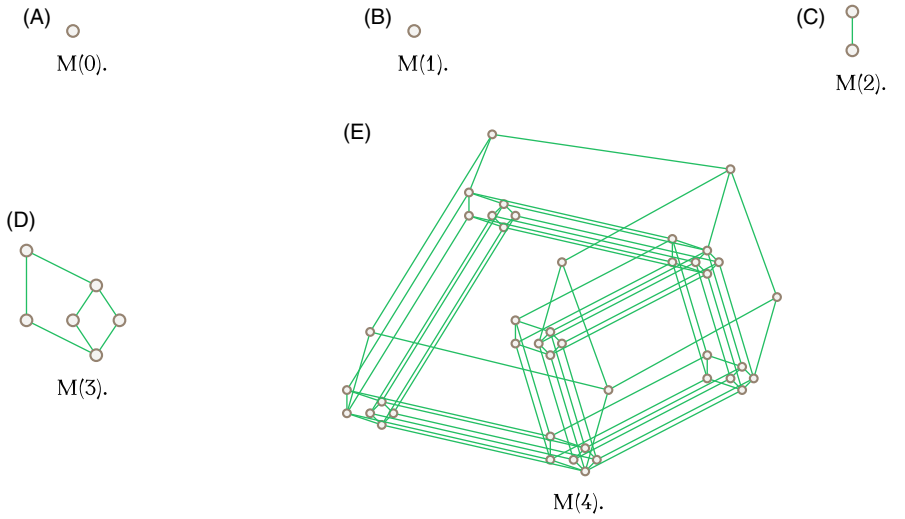


Figure 5. The Hasse diagrams of the Mockingbird lattices M(d) for $d \in [\![4]\!]$.

Proof. By Lemmas 3.2.3 and 3.2.4, the poset $\mathcal{D}^*(\mathfrak{f})$ is isomorphic to a maximal interval of $\mathcal{D}^*(\mathfrak{l}_d)$ for a $d \geqslant 0$. Since $\mathtt{pr}\big(\mathtt{fr}\big(\mathfrak{r}_{d+1}\big)\big) = \mathfrak{l}_d$, and since by Proposition 3.3.3, the poset $\mathcal{P}\big(\mathfrak{r}_{d+1}\big) = \mathrm{M}(d+1)$ is isomorphic to $\mathcal{D}^*(\mathfrak{l}_d)$, the poset $\mathcal{D}^*(\mathfrak{f})$ is isomorphic to a maximal interval of $\mathrm{M}(d+1)$. □

Theorem 3.3.5 justifies the fact that the study of the Mockingbird lattices is universal enough because these lattices contain as maximal interval all duplicative forest lattices.

The Mockingbird lattices are not graded and not self-dual. Since they are not graded, they are not distributive neither. Moreover, they are not semi-distributive. For instance, in M(3), by setting $\mathfrak{t}_1 := \mathbf{M}(\mathbf{MM}(\mathbf{MM}))$, $\mathfrak{t}_2 := \mathbf{MM}(\mathbf{MM})(\mathbf{M}(\mathbf{MM}))$, and $\mathfrak{t}_3 := \mathbf{M}(\mathbf{MM})(\mathbf{MM}(\mathbf{MM}))$, we have $\mathfrak{t}_1 \wedge \mathfrak{t}_2 = \mathfrak{t}_1 \wedge \mathfrak{t}_3$ but $\mathfrak{t}_1 \wedge (\mathfrak{t}_2 \vee \mathfrak{t}_3) \neq \mathfrak{t}_1 \wedge \mathfrak{t}_2$, contradicting one of the required relations to have this property.

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

4. Enumerative Properties

In this last central part of this work, we present some enumerative results concerning the Mockingbird poset and lattices $\mathrm{M}(d)$, $d \geqslant 0$. We enumerate the maximal and minimal elements of the Mockingbird poset by degree and by height, the length of the shortest and longest saturated chains in $\mathrm{M}(d)$, the number of elements in $\mathrm{M}(d)$, the number of edges in the Hasse diagram of $\mathrm{M}(d)$, and the number of intervals of $\mathrm{M}(d)$. All this use the lattice isomorphism between the Mockingbird lattices and the duplicative forest lattices introduced in the previous sections and formal series on terms and on duplicative forests.

4.1 Formal power series

We set here some notions about formal power series and generating series. For the rest of the text, $\mathbb{K}$ is any field of characteristic zero (as $\mathbb{Q}$ for instance).

4.1.1 Formal power series over sets

For any set X, let $\mathbb{K}\langle X\rangle$ be the linear span of X. The dual space of $\mathbb{K}\langle X\rangle$ is denoted by $\mathbb{K}\langle\langle X\rangle\rangle$ and is by definition the space of the maps $\mathbf{F} : X \to \mathbb{K}$, called *X-series*. The coefficient $\mathbf{F}(x)$ of any $x \in X$ is denoted by $\langle x, \mathbf{F}\rangle$. The *support* of $\mathbf{F}$ is the set $\mathrm{Supp}(\mathbf{F}) := \{x \in X : \langle x, \mathbf{F}\rangle \neq 0\}$. The *characteristic series* of any subset X' of X is the series $\mathbf{c}(X')$ having X' as support and such that the coefficient of each $x \in X'$ is 1. For any $k \geqslant 0$, $\mathrm{T}^k\mathbb{K}\langle\langle X\rangle\rangle$ is the k-th tensor power of $\mathbb{K}\langle\langle X\rangle\rangle$. Elements of this space are possibly infinite linear combinations of tensors $x_1 \otimes \cdots \otimes x_k$, where for any $i \in [k]$, $x_i \in X$. The *tensor algebra* of $\mathbb{K}\langle\langle X\rangle\rangle$ is the space

$$\mathrm{T}^*\mathbb{K}\langle\langle X\rangle\rangle := \bigoplus_{k\geqslant 0} \mathrm{T}^k\mathbb{K}\langle\langle X\rangle\rangle. \tag{4.1.1}$$

This space is endowed with the tensor product $\otimes$ so that $(\mathrm{T}^*\mathbb{K}\langle\langle X\rangle\rangle, \otimes)$ is a unital associative algebra, admitting $1 \in \mathbb{K}$ as unit.

A linear map $\phi : \mathrm{T}^{k_1}\mathbb{K}\langle\langle X\rangle\rangle \to \mathrm{T}^{k_2}\mathbb{K}\langle\langle X\rangle\rangle$, $k_1, k_2 \geqslant 0$, is a (k_1, k_2)*-operation* on $\mathbb{K}\langle\langle X\rangle\rangle$. To lighten the notation, when ϕ is a $(2, 1)$-operation on $\mathbb{K}\langle\langle X\rangle\rangle$, we shall sometimes use ϕ as an infix operation by writing $\mathbf{F}_1 \,\phi\, \mathbf{F}_2$ instead of $\phi(\mathbf{F}_1 \otimes \mathbf{F}_2)$ for any $\mathbf{F}_1, \mathbf{F}_2 \in \mathbb{K}\langle\langle X\rangle\rangle$. The *diagonal coproduct* is the $(1, 2)$-operation Δ on $\mathbb{K}\langle\langle X\rangle\rangle$ satisfying $\Delta(x) = x \otimes x$ for any $x \in X$. When X is endowed with an n-ary operation $\bullet : X^n \to X$, $n \geqslant 0$, for any $k \geqslant 0$, the *k-linearization* of $\bullet$ is the (nk, k)-operation $\bullet^{\otimes_k}$ on $\mathbb{K}\langle\langle X\rangle\rangle$ satisfying

$$\bullet^{\otimes_k}\bigl(x_{1,1} \otimes \cdots \otimes x_{1,k} \otimes \cdots \otimes x_{n,1} \otimes \cdots \otimes x_{n,k}\bigr) = \bullet\bigl(x_{1,1}, \ldots, x_{n,1}\bigr) \otimes \cdots \otimes \bullet\bigl(x_{k,1}, \ldots, x_{n,k}\bigr) \tag{4.1.2}$$

for any $x_{i,j} \in X$, $i \in [n]$, and $j \in [k]$. To lighten the notation, we shall write $\bullet^{\otimes}$ for $\bullet^{\otimes_1}$.

4.1.2 Generating series

The space of the usual power series on the formal parameter z is denoted by $\mathbb{K}\langle\langle \mathsf{z}\rangle\rangle$. For any $F, F' \in \mathbb{K}\langle\langle \mathsf{z}\rangle\rangle$, $F\bigl[\mathsf{z} := F'\bigr]$ is the series of $\mathbb{K}\langle\langle \mathsf{z}\rangle\rangle$ obtained by substituting F' for z in F. The *Hadamard product* is the binary operation $\boxtimes$ on $\mathbb{K}\langle\langle \mathsf{z}\rangle\rangle$ defined linearly for any $n_1, n_2 \geqslant 0$ by

$$\mathsf{z}^{n_1} \boxtimes \mathsf{z}^{n_2} := [\, n_1 = n_2 \,]\, z^{n_1}. \tag{4.1.3}$$

The *max product* is the binary operation $\uparrow$ on $\mathbb{K}\langle\langle \mathsf{z}\rangle\rangle$ defined linearly for any $n_1, n_2 \geqslant 0$ by $\mathsf{z}^{n_1} \uparrow \mathsf{z}^{n_2} := \mathsf{z}^{\max\{n_1, n_2\}}$. Observe that for any $F \in \mathbb{K}\langle\langle \mathsf{z}\rangle\rangle$ and $n \geqslant 0$,

$$\bigl\langle \mathsf{z}^n, F \uparrow F\bigr\rangle = \bigl\langle \mathsf{z}^n, F\bigr\rangle^2 + 2\bigl\langle \mathsf{z}^n, F\bigr\rangle \sum_{i\in[n]} \bigl\langle \mathsf{z}^{i-1}, F\bigr\rangle. \tag{4.1.4}$$

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

4.1.3 Formal power series and enumeration

If X is endowed with a map $\omega : X \to \mathbb{N}$, the *$\omega$-enumeration map* is the partial map $\mathrm{en}_\omega : \mathrm{T}^*\mathbb{K}\langle\langle X\rangle\rangle \to \mathbb{K}\langle\langle \mathsf{z}\rangle\rangle$ defined linearly for any $k \geqslant 1$ and $x_1, \ldots, x_k \in X$ by

$$\mathrm{en}_\omega(x_1 \otimes \cdots \otimes x_k) := \mathsf{z}^{\omega(x_1)} \uparrow \ldots \uparrow \mathsf{z}^{\omega(x_k)}. \tag{4.1.5}$$

For any $\mathbf{F} \in \mathrm{T}^*\mathbb{K}\langle\langle X\rangle\rangle$, the generating series $\mathrm{en}_\omega(\mathbf{F})$ is the *ω-enumeration* of $\mathbf{F}$.

In the sequel, we shall use the following strategy to enumerate a set X w.r.t. such a map ω: we shall provide a description of $\mathbf{c}(X)$, then deduce a description of $\mathrm{en}_\omega(\mathbf{c}(X))$, and finally deduce from this a formula to compute the coefficients $\langle \mathsf{z}^n, \mathrm{en}_\omega(\mathbf{c}(X))\rangle$, $n \geqslant 0$.

4.2 Maximal, minimal, and isolated terms

We use here series on terms and some operations on these in order to enumerate the closed maximal, closed minimal, and closed isolated elements of the poset $\mathcal{P}$ w.r.t. their degrees or their heights. These enumerative results can be used to establish the probabilities for a term $\mathfrak{t}$ of a given degree or of a given height generated uniformly at random to have one of these properties. In particular, the number of maximal (resp. minimal, isolated) closed terms gives information to compute the probability for the fact that $\mathfrak{t}$ cannot be rewritten into a different term (resp. no term different from $\mathfrak{t}$ rewrites into $\mathfrak{t}$, any sequence of rewriting steps involving $\mathfrak{t}$ does not involve different other terms).

4.2.1 Series of terms and operations

Recall that $\star$ is the binary operation on $\mathfrak{T}(\mathfrak{G})$ such that for any $\mathfrak{t}_1, \mathfrak{t}_2 \in \mathfrak{T}(\mathfrak{G})$, $\mathfrak{t}_1 \star \mathfrak{t}_2$ is the $\mathfrak{G}$-term $\mathfrak{t}_1\mathfrak{t}_2$ having $\mathfrak{t}_1$ as left subtree and $\mathfrak{t}_2$ as right subtree. Observe that for any $\mathbf{F}_1, \mathbf{F}_2 \in \mathbb{K}\langle\langle\mathfrak{T}(\mathfrak{G})\rangle\rangle$,

$$\mathrm{en}_{\mathrm{deg}}\big(\mathbf{F}_1 \star^{\otimes} \mathbf{F}_2\big) = \mathsf{z}\,\mathrm{en}_{\mathrm{deg}}(\mathbf{F}_1)\,\mathrm{en}_{\mathrm{deg}}(\mathbf{F}_2) \tag{4.2.1}$$

and

$$\mathrm{en}_{\mathrm{ht}}\big(\mathbf{F}_1 \star^{\otimes} \mathbf{F}_2\big) = \mathsf{z}\,(\mathrm{en}_{\mathrm{ht}}(\mathbf{F}_1) \uparrow \mathrm{en}_{\mathrm{ht}}(\mathbf{F}_2)). \tag{4.2.2}$$

Observe also that for any $\mathbf{F} \in \mathbb{K}\langle\langle\mathfrak{T}(\mathfrak{G})\rangle\rangle$,

$$\mathrm{en}_{\mathrm{deg}}\big(\star^{\otimes}(\Delta(\mathbf{F}))\big) = \mathsf{z}\left(\mathrm{en}_{\mathrm{deg}}(\mathbf{F})\big[\mathsf{z} := \mathsf{z}^2\big]\right) \tag{4.2.3}$$

and

$$\mathrm{en}_{\mathrm{ht}}\big(\star^{\otimes}(\Delta(\mathbf{F}))\big) = \mathsf{z}\,\mathrm{en}_{\mathrm{ht}}(\mathbf{F}). \tag{4.2.4}$$

4.2.2 Three series of terms

In order to achieve the objectives described above, we begin by providing equations satisfied by the characteristic series $\mathbf{F}_{\mathrm{max}}$ of the closed maximal terms of $\mathcal{P}$, the characteristic series $\mathbf{F}_{\mathrm{min}}$ of the closed minimal terms of $\mathcal{P}$, and the characteristic series $\mathbf{F}_{\mathrm{iso}}$ of the closed isolated terms of $\mathcal{P}$.

Proposition 4.2.1. *The characteristic series* $\mathbf{F}_{\mathrm{max}}$ *satisfies*

$$\mathbf{F}_{\mathrm{max}} = \mathbf{M} + \mathbf{M}\mathbf{M} + \mathbf{F}_{\mathrm{max}} \star^{\otimes} \mathbf{F}_{\mathrm{max}} - \mathbf{M} \star^{\otimes} \mathbf{F}_{\mathrm{max}}. \tag{4.2.5}$$

Proof. By using the description of Proposition 3.1.4 for the set S of closed maximal elements of $\mathcal{P}$, we have

$$\mathbf{F}_{\mathrm{max}} = \mathbf{M} + \sum_{\mathfrak{s}_1,\mathfrak{s}_2 \in S} \mathfrak{s}_1\mathfrak{s}_2 - \sum_{\mathfrak{s}_3 \in S\setminus\{\mathbf{M}\}} \mathbf{M}\mathfrak{s}_3 = \mathbf{M} + \mathbf{F}_{\mathrm{max}} \star^{\otimes} \mathbf{F}_{\mathrm{max}} - \mathbf{M} \star^{\otimes} (\mathbf{F}_{\mathrm{max}} - \mathbf{M}). \tag{4.2.6}$$

This leads to (4.2.5). □

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

Proposition 4.2.2. *The characteristic series* $\mathbf{F}_{\min}$ *satisfies*

$$\mathbf{F}_{\min} = \mathbf{M} + \mathbf{M}\mathbf{M} + \mathbf{F}_{\min} \star^{\otimes} \mathbf{F}_{\min} - \star^{\otimes}(\Delta(\mathbf{F}_{\min})). \tag{4.2.7}$$

Proof. By using the description of Proposition 3.1.4 for the set S of closed minimal elements of $\mathcal{P}$, we have

$$\mathbf{F}_{\min} = \mathbf{M} + \sum_{\mathfrak{s}_1,\mathfrak{s}_2 \in S} \mathfrak{s}_1\mathfrak{s}_2 - \sum_{\mathfrak{s}_3 \in S \setminus \{\mathbf{M}\}} \mathfrak{s}_3\mathfrak{s}_3 = \mathbf{M} + \mathbf{F}_{\min} \star^{\otimes} \mathbf{F}_{\min} - \star^{\otimes}(\Delta(\mathbf{F}_{\min} - \mathbf{M})). \tag{4.2.8}$$

This leads to (4.2.7). □

Proposition 4.2.3. *The characteristic series* $\mathbf{F}_{\mathrm{iso}}$ *satisfies*

$$\mathbf{F}_{\mathrm{iso}} = \mathbf{M} + 2\mathbf{M}\mathbf{M} + \mathbf{F}_{\mathrm{iso}} \star^{\otimes} \mathbf{F}_{\mathrm{iso}} - \mathbf{M} \star^{\otimes} \mathbf{F}_{\mathrm{iso}} - \star^{\otimes}(\Delta(\mathbf{F}_{\mathrm{iso}})). \tag{4.2.9}$$

Proof. By using the description of Proposition 3.1.4 for the set S_1 (resp. S_2) of the closed maximal (resp. minimal) elements of $\mathcal{P}$, the set of the closed isolated elements of $\mathcal{P}$ is the set $S = S_1 \cap S_2$ and we have

$$\begin{aligned} \mathbf{F}_{\mathrm{iso}} &= \mathbf{M} + \sum_{\mathfrak{s}_1,\mathfrak{s}_2 \in S} \mathfrak{s}_1\mathfrak{s}_2 - \sum_{\mathfrak{s}_3 \in S \setminus \{\mathbf{M}\}} \mathbf{M}\mathfrak{s}_3 - \sum_{\mathfrak{s}_4 \in S \setminus \{\mathbf{M}\}} \mathfrak{s}_4\mathfrak{s}_4 \\ &= \mathbf{M} + \mathbf{F}_{\mathrm{iso}} \star^{\otimes} \mathbf{F}_{\mathrm{iso}} - \mathbf{M} \star^{\otimes} (\mathbf{F}_{\mathrm{iso}} - \mathbf{M}) - \star^{\otimes}(\Delta(\mathbf{F}_{\mathrm{iso}} - \mathbf{M})). \end{aligned} \tag{4.2.10}$$

This leads to (4.2.9). □

4.2.3 deg*-enumerations*

A consequence of Proposition 4.2.1 and of (4.2.1) is that the deg-enumeration $\mathsf{D}_{\max}$ of $\mathbf{F}_{\max}$, enumerating the closed maximal elements of $\mathcal{P}$ w.r.t. their degrees, satisfies

$$\mathsf{D}_{\max} = 1 + \mathsf{z} + \mathsf{z}\mathsf{D}_{\max}^2 - \mathsf{z}\mathsf{D}_{\max}. \tag{4.2.11}$$

We deduce from this that the number of these terms of degree $d \geqslant 0$ is $\mathbf{d}_{\max}(d)$ where $\mathbf{d}_{\max}$ is the integer sequence satisfying $\mathbf{d}_{\max}(0) = \mathbf{d}_{\max}(1) = 1$ and, for any $d \geqslant 2$,

$$\mathbf{d}_{\max}(d) = \sum_{i \in [\![d-2]\!]} \mathbf{d}_{\max}(i)\, \mathbf{d}_{\max}(d-1-i). \tag{4.2.12}$$

The first numbers are

$$1, 1, 1, 2, 4, 9, 21, 51, \tag{4.2.13}$$

form Sequence A001006 of Sloane (https://oeis.org/), and are Motzkin numbers. A *Motzkin tree* is a planar rooted tree made of nodes of arity 0, 1, or 2. It is well-known that the sets of these trees having exactly $d \geqslant 1$ nodes are enumerated by Motzkin numbers (see Dulucq and Penaud 2002 for instance). Given a closed maximal term $\mathfrak{t}$, it is easy to see that $\mathtt{fr}(\mathfrak{t})$ contains only black nodes and that it is a Motzkin tree. It is also easy to check that $\mathtt{fr}$ is a bijection between the set of the closed maximal terms of degree $d \geqslant 1$ and the set of the Motzkin trees with d nodes.

Besides, a consequence of Proposition 4.2.2, (4.2.1), and (4.2.3) is that the deg-enumeration $\mathsf{D}_{\min}$ of $\mathbf{F}_{\min}$, enumerating the closed minimal elements of $\mathcal{P}$ w.r.t. their degrees, satisfies

$$\mathsf{D}_{\min} = 1 + \mathsf{z} + \mathsf{z}\mathsf{D}_{\min}^2 - \mathsf{z}\big(\mathsf{D}_{\min}\big[\mathsf{z} := \mathsf{z}^2\big]\big). \tag{4.2.14}$$

We deduce from this that the number of these terms of degree $d \geqslant 0$ is $\mathbf{d}_{\min}(d)$ where $\mathbf{d}_{\min}$ is the integer sequence satisfying $\mathbf{d}_{\min}(0) = \mathbf{d}_{\min}(1) = 1$ and, for any $d \geqslant 2$,

$$\mathbf{d}_{\min}(d) = \sum_{i \in [\![d-1]\!]} \mathbf{d}_{\min}(i)\, \mathbf{d}_{\min}(d-1-i) - \big[d \text{ is odd}\big]\, \mathbf{d}_{\min}\left(\frac{d-1}{2}\right). \tag{4.2.15}$$

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

The first numbers are

$$1, 1, 2, 4, 12, 34, 108, 344 \tag{4.2.16}$$

and form Sequence A343663 of Sloane (https://oeis.org/). A *semi-identity tree* is a binary tree having no subtrees that are brother and equal, unless they are leaves. It is known that the sets of these trees having exactly $d \geqslant 0$ internal nodes are enumerated by the previous sequence. It is immediate to see that the identity map is a bijection between the set of the closed minimal terms of degree $d \geqslant 0$ and the set of the semi-identity trees with exactly d internal nodes.

Besides, a consequence of Proposition 4.2.3, (4.2.1), and (4.2.3) is that the deg-enumeration $\mathsf{D}_{\mathrm{iso}}$ of $\mathbf{F}_{\mathrm{iso}}$, enumerating the closed isolated elements of $\mathcal{P}$ w.r.t. their degrees, satisfies

$$\mathsf{D}_{\mathrm{iso}} = 1 + 2\mathsf{z} + \mathsf{z}\mathsf{D}_{\mathrm{iso}}^2 - \mathsf{z}\mathsf{D}_{\mathrm{iso}} - \mathsf{z}\big(\mathsf{D}_{\mathrm{iso}}\big[\mathsf{z} := \mathsf{z}^2\big]\big). \tag{4.2.17}$$

We deduce from this that the number of these terms of degree $d \geqslant 0$ is $\mathbf{d}_{\mathrm{iso}}(d)$ where $\mathbf{d}_{\mathrm{iso}}$ is the integer sequence satisfying $\mathbf{d}_{\mathrm{iso}}(0) = \mathbf{d}_{\mathrm{iso}}(1) = 1$ and, for any $d \geqslant 2$,

$$\mathbf{d}_{\mathrm{iso}}(d) = \sum_{i \in [\![d-2]\!]} \mathbf{d}_{\mathrm{iso}}(i)\, \mathbf{d}_{\mathrm{iso}}(d-1-i) - \big[\, d \text{ is odd} \,\big]\, \mathbf{d}_{\mathrm{iso}}\left(\frac{d-1}{2}\right). \tag{4.2.18}$$

The first numbers are

$$1, 1, 1, 1, 3, 5, 13, 29, 71, 171, 427, 1067, 2709. \tag{4.2.19}$$

This sequence does not appear in Sloane (https://oeis.org/) for the time being.

4.2.4 ht-*enumerations*

Due to the fact that $\mathcal{C}$ is hierarchical and, by Proposition 3.1.2, $\mathcal{P}$ is rooted, Proposition 2.2.3 implies that $\mathcal{P}$ satisfies the properties exposed at the very end of Section 2.2.2. In particular, in each $\equiv$-equivalence class of closed terms, there is exactly one closed maximal term, exactly one closed minimal term, and these terms have the same height. For this reason, the ht-enumerations of $\mathbf{F}_{\max}$ and $\mathbf{F}_{\min}$ are equal and enumerate also the $\equiv$-equivalence classes of closed terms w.r.t. the heights of their elements. By denoting by H_{mm} this generating series, by Propositions 4.2.1 and 4.2.2, H_{mm} satisfies

$$\mathsf{H}_{\mathrm{mm}} = 1 + \mathsf{z} + \mathsf{z}(\mathsf{H}_{\mathrm{mm}} \uparrow \mathsf{H}_{\mathrm{mm}}) - \mathsf{z}\mathsf{H}_{\mathrm{mm}}. \tag{4.2.20}$$

We deduce from this and (4.1.4) that the number of these $\equiv$-equivalence classes of terms of height $h \geqslant 0$ is $\mathbf{h}_{\mathrm{mm}}(h)$ where $\mathbf{h}_{\mathrm{mm}}$ is the integer sequence satisfying $\mathbf{h}_{\mathrm{mm}}(0) = \mathbf{h}_{\mathrm{mm}}(1) = 1$ and, for any $h \geqslant 2$,

$$\mathbf{h}_{\mathrm{mm}}(h) = \mathbf{h}_{\mathrm{mm}}(h-1)^2 - \mathbf{h}_{\mathrm{mm}}(h-1) + 2\mathbf{h}_{\mathrm{mm}}(h-1) \sum_{i \in [h-1]} \mathbf{h}_{\mathrm{mm}}(i-1). \tag{4.2.21}$$

The first numbers are

$$1, 1, 2, 10, 170, 33490, 1133870930, 1285739648704587610 \tag{4.2.22}$$

and form Sequence A063573 of Sloane (https://oeis.org/).

Besides, a consequence of Proposition 4.2.3, (4.2.2), and (4.2.4) is that the ht-enumeration $\mathsf{H}_{\mathrm{iso}}$ of $\mathbf{F}_{\mathrm{iso}}$, enumerating the closed isolated elements of $\mathcal{P}$ w.r.t. their heights, satisfies

$$\mathsf{H}_{\mathrm{iso}} = 1 + 2z + z(\mathsf{H}_{\mathrm{iso}} \uparrow \mathsf{H}_{\mathrm{iso}}) - 2z\mathsf{H}_{\mathrm{iso}}. \tag{4.2.23}$$

We deduce from this and (4.1.4) that the number of these terms of height $h \geqslant 0$ is $\mathbf{h}_{\mathrm{iso}}(h)$ where $\mathbf{h}_{\mathrm{iso}}$ is the integer sequence satisfying $\mathbf{h}_{\mathrm{iso}}(0) = \mathbf{h}_{\mathrm{iso}}(1) = 1$ and, for any $h \geqslant 2$,

$$\mathbf{h}_{\mathrm{iso}}(h) = \mathbf{h}_{\mathrm{iso}}(h-1)^2 - 2\mathbf{h}_{\mathrm{iso}}(h-1) + 2\mathbf{h}_{\mathrm{iso}}(h-1) \sum_{i \in [h-1]} \mathbf{h}_{\mathrm{iso}}(i-1). \tag{4.2.24}$$

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

The first numbers are

$$1, 1, 1, 3, 21, 651, 457653, 210065930571 \tag{4.2.25}$$

and form Sequence A001699 of Sloane (https://oeis.org/), which enumerates binary trees w.r.t. the height (but with a shift). Therefore, there is a one-to-one correspondence between the set of the closed isolated elements of $\mathcal{P}$ of height $h \geqslant 1$ and the set of the binary trees of height $h-1$.

4.3 Shortest and longest saturated chains

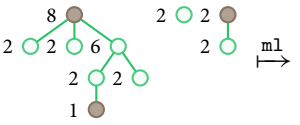
Let $\mathtt{ml} : \mathcal{D}^* \to \mathbb{N}$ be the statistics defined for any $\ell \geqslant 0$, $\mathfrak{f}_1, \ldots, \mathfrak{f}_\ell \in \mathcal{D}^*$, and $\mathfrak{f} \in \mathcal{D}^*$ by

$$\mathtt{ml}(\mathfrak{f}_1 \centerdot \ldots \centerdot \mathfrak{f}_\ell) := 1 - \ell + \sum_{i \in [\ell]} \mathtt{ml}(\mathfrak{f}_i), \tag{4.3.1a}$$

$$\mathtt{ml}(\bullet(\mathfrak{f})) := \mathtt{ml}(\mathfrak{f}), \tag{4.3.1b}$$

$$\mathtt{ml}(\circ(\mathfrak{f})) := 2\mathtt{ml}(\mathfrak{f}). \tag{4.3.1c}$$

For instance, we have



(4.3.2)

where the integer decoration of each node refers to the image by $\mathtt{ml}$ of the duplicative tree rooted at this node.

Proposition 4.3.1. *For any* $\mathfrak{t} \in \mathfrak{T}_0(\mathfrak{G})$,

(i) a shortest saturated chain from $\mathfrak{t}$ *to the maximal element of* $\mathcal{P}(\mathfrak{t})$ *has as length the number of* $\circ$ *in* $fr(\mathfrak{t})$;
(ii) a longest saturated chain from $\mathfrak{t}$ *to the maximal element of* $\mathcal{P}(\mathfrak{t})$ *has length* $ml(fr(\mathfrak{t}))$.

Proof. Let $\mathfrak{f} := \mathtt{fr}(\mathfrak{t})$. By Proposition 3.3.3, the posets $\mathcal{P}(\mathfrak{t})$ and $\mathcal{D}^*(\mathfrak{f})$ are isomorphic. Therefore, to prove the statement, we consider the lengths of the shortest and longest saturated chains in $\mathcal{D}^*(\mathfrak{f})$ from $\mathfrak{f}$ to $\mathfrak{g}$ where $\mathfrak{g}$ is the maximal element of $\mathcal{D}^*(\mathfrak{f})$.

First, immediately from the definition of $\ll$, a shortest saturated chain from $\mathfrak{f}$ to $\mathfrak{g}$ consists in selecting at each step a white node that has no white nodes as descendants and turn it into black. This establishes (i).

Let us prove (ii) by structural induction on $\mathfrak{f}$. We use Lemma 3.2.1 in order to exhibit a longest saturated chain from $\mathfrak{f}$ to $\mathfrak{g}$. If $\mathfrak{f} = \mathfrak{f}_1 \centerdot \ldots \centerdot \mathfrak{f}_\ell$, $\ell \geqslant 0$, a longest saturated chain from $\mathfrak{f}$ passes through the duplicative forest $\mathfrak{f}'_1 \centerdot \mathfrak{f}_2 \centerdot \ldots \centerdot \mathfrak{f}_\ell$ and then passes through $\mathfrak{f}'_1 \centerdot \mathfrak{f}'_2 \centerdot \ldots \centerdot \mathfrak{f}_\ell$, …, and ends at $\mathfrak{f}'_1 \centerdot \mathfrak{f}'_2 \centerdot \ldots \centerdot \mathfrak{f}'_\ell = \mathfrak{g}$, where for any $i \in [\ell]$, $\mathfrak{f}'_i$ is the greatest element of $\mathcal{D}^*(\mathfrak{f}_i)$. By induction hypothesis, $\mathtt{ml}(\mathfrak{f}_i)$ is the length of a longest saturated chain from $\mathfrak{f}_i$ to $\mathfrak{f}'_i$. We deduce that the length of the former chain from $\mathfrak{f}$ to $\mathfrak{g}$ is $1 + \sum_{i \in [\ell]} (\mathtt{ml}(\mathfrak{f}_i) - 1)$. Therefore, (4.3.1a) is consistent. If $\mathfrak{f} = \bullet(\mathfrak{f}')$, $\mathfrak{f}' \in \mathcal{D}^*$, a longest saturated saturated chain from $\mathfrak{f}$ to $\mathfrak{g}$ has the same length as a longest saturated chain from $\mathfrak{f}'$ to $\mathfrak{g}'$ where $\mathfrak{g}'$ is the greatest element of $\mathcal{D}^*(\mathfrak{f}')$. Hence, by induction hypothesis, $\mathtt{ml}(\mathfrak{f})$ is consistent with (4.3.1b). If $\mathfrak{f} = \circ(\mathfrak{f}')$, $\mathfrak{f}' \in \mathcal{D}^*$, a longest saturated chain from $\mathfrak{f}$ to $\mathfrak{g}$ passes through the covering $\bullet(\mathfrak{f}' \centerdot \mathfrak{f}')$ of $\mathfrak{f}$. Therefore, by induction hypothesis and by (4.3.1b) and (4.3.1a), a longest saturated chain from $\mathfrak{f}$ to $\mathfrak{g}$ has length $1 + \mathtt{ml}(\bullet(\mathfrak{f}' \centerdot \mathfrak{f}')) = 1 + 1 - 2 + 2\mathtt{ml}(\mathfrak{f}') = 2\mathtt{ml}(\mathfrak{f}')$, which is consistent with (4.3.1c). □

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

By Proposition 4.3.1, for any $d \geqslant 1$, in $\mathrm{M}(d)$, shortest saturated chains are of length d and longest saturated chains are of length 2^{d-1}.

4.4 Elements, covering pairs, and intervals

We use here series on duplicative forests and a palette of operations on these to enumerate the elements, covering pairs, and intervals of the Mockingbird lattices. We obtain for these three recursive formulas leading to expressions for the generating series of these numbers. These enumerative results can be used to establish the probability for a pair $(\mathrm{t}_1, \mathrm{t}_2)$ of terms both of a given height and generated uniformly at random to have one of these properties. In particular, the number of covering pairs (resp. of intervals) gives information to compute the probability for the fact that t_1 can be rewritten in one step (resp. in any number of steps) into t_2.

4.4.1 Series of duplicative forests and operations

We shall consider in the sequel the k-linearization $\centerdot^{\otimes_k}$, $k \geqslant 0$, of the concatenation product $\centerdot$ of duplicative forests. Observe that for any $\mathbf{F}_1, \mathbf{F}_2 \in \mathbb{K}\langle\langle \mathcal{D}^* \rangle\rangle$, we have in particular

$$\mathrm{en}_{\mathrm{ht}}\big(\mathbf{F}_1 \centerdot^{\otimes} \mathbf{F}_2\big) = \mathrm{en}_{\mathrm{ht}}(\mathbf{F}_1) \uparrow \mathrm{en}_{\mathrm{ht}}(\mathbf{F}_2). \tag{4.4.1}$$

We shall also consider the k-linearization $\square^{\otimes_k}$, $k \geqslant 0$ of the grafting product $\square$. Observe that for any $\mathbf{F} \in \mathbb{K}\langle\langle \mathcal{D}^* \rangle\rangle$, we have in particular

$$\mathrm{en}_{\mathrm{ht}}\big(\square^{\otimes}(\mathbf{F})\big) = \mathrm{z}\,\mathrm{en}_{\mathrm{ht}}(\mathbf{F}). \tag{4.4.2}$$

For any $k \geqslant 1$ and $u \in \{\circ, \bullet\}^k$, the *merging product* is the $\big(k + |u|_\bullet, k\big)$-operation mg_u on $\mathbb{K}\langle\langle \mathcal{D}^* \rangle\rangle$ satisfying, for any $\mathfrak{f}_1, \ldots, \mathfrak{f}_{k+|u|_\bullet} \in \mathcal{D}^*$,

$$\mathrm{mg}_\circ(\mathfrak{f}_1) = \circ(\mathfrak{f}_1), \tag{4.4.3a}$$

$$\mathrm{mg}_{\circ u'}(\mathfrak{f}_1 \otimes \cdots \otimes \mathfrak{f}_k) = \mathrm{mg}_\circ(\mathfrak{f}_1) \otimes \mathrm{mg}_{u'}(\mathfrak{f}_2 \otimes \cdots \otimes \mathfrak{f}_k), \tag{4.4.3b}$$

$$\mathrm{mg}_\bullet(\mathfrak{f}_1 \otimes \mathfrak{f}_2) = \bullet(\mathfrak{f}_1 \centerdot \mathfrak{f}_2), \tag{4.4.3c}$$

$$\mathrm{mg}_{\bullet u'}(\mathfrak{f}_1 \otimes \cdots \otimes \mathfrak{f}_k) = \mathrm{mg}_\bullet(\mathfrak{f}_1 \otimes \mathfrak{f}_2) \otimes \mathrm{mg}_{u'}(\mathfrak{f}_3 \otimes \cdots \otimes \mathfrak{f}_k), \tag{4.4.3d}$$

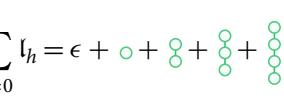
where $u' \in \{\circ, \bullet\}^*$. For instance, with $k = 3$,

$\mathrm{mg}_{\circ\bullet\bullet}($  $) =$  . (4.4.4)

Intuitively, this product consists in grafting a $\circ$ onto one forest or a $\bullet$ onto a forest and its right neighbor following the letters of u. Observe that for any $u \in \{\circ, \bullet\}^*$ and $\mathbf{F} \in \mathrm{T}^{k+|u|_\bullet}\mathbb{K}\langle\langle \mathcal{D}^* \rangle\rangle$, we have

$$\mathrm{en}_{\mathrm{ht}}\big(\mathrm{mg}_u(\mathbf{F})\big) = \mathrm{z}\,\mathrm{en}_{\mathrm{ht}}(\mathbf{F}). \tag{4.4.5}$$

Let us finally define the *series of ladders* as the $\mathcal{D}^*$-series

$\mathbf{ld} := \sum_{h \geqslant 0} \mathfrak{l}_h = \epsilon +$  $+ \cdots$. (4.4.6)

Observe that

$$\mathbf{ld} = \epsilon + \circ^{\otimes}(\mathbf{ld}). \tag{4.4.7}$$

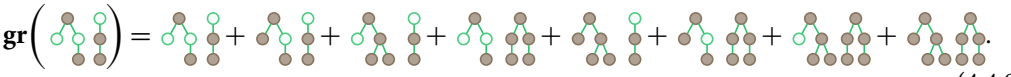
https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

4.4.2 Number of elements

Let **gr** be the (1, 1)-operation on $\mathbb{K}\langle\langle\mathcal{D}^*\rangle\rangle$ satisfying, for any $\mathfrak{f} \in \mathcal{D}^*$,

$$\mathbf{gr}(\mathfrak{f}) = \sum_{\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f})} \mathfrak{f}'. \tag{4.4.8}$$

By definition, $\mathbf{gr}(\mathfrak{f})$ is the characteristic series of $\mathcal{D}^*(\mathfrak{f})$ and is therefore also the formal sum of all duplicative forests $\mathfrak{f}'$ such that $\mathfrak{f} \ll \mathfrak{f}'$. For instance,



(4.4.9)

Proposition 4.4.1. *For any $\ell \geqslant 0$, $\mathfrak{f}_1, \ldots, \mathfrak{f}_\ell \in \mathcal{D}^*$, and $\mathfrak{f} \in \mathcal{D}^*$,*

$$\mathbf{gr}(\mathfrak{f}_1 \centerdot \cdots \centerdot \mathfrak{f}_\ell) = \mathbf{gr}(\mathfrak{f}_1) \centerdot^{\otimes} \cdots \centerdot^{\otimes} \mathbf{gr}(\mathfrak{f}_\ell), \tag{4.4.10a}$$

$$\mathbf{gr}(\bullet(\mathfrak{f})) = \bullet^{\otimes}(\mathbf{gr}(\mathfrak{f})), \tag{4.4.10b}$$

$$\mathbf{gr}(\circ(\mathfrak{f})) = \circ^{\otimes}(\mathbf{gr}(\mathfrak{f})) + \bullet^{\otimes}(\mathbf{gr}(\mathfrak{f} \centerdot \mathfrak{f})). \tag{4.4.10c}$$

Proof. This follows by structural induction on duplicative forests and is a consequence of Lemma 3.2.1 describing a necessary and sufficient condition for the fact that a duplicative forest $\mathfrak{f}'$ belong to $\mathcal{D}^*(\mathfrak{f})$ and the definitions of the operations $\centerdot^{\otimes}$, $\circ^{\otimes}$, and $\bullet^{\otimes}$ on $\mathcal{D}^*$-series. □

Observe that $\mathbf{gr}(\mathbf{ld})$ is the characteristic series of the set $\mathcal{L}$ (defined in Section 3.2.3). Moreover, since for any $h \geqslant 0$, all elements of $\mathcal{D}^*(\mathfrak{l}_h)$ have h as height, we have

$$\mathrm{en}_{\mathrm{ht}}(\mathbf{gr}(\mathbf{ld})) = \sum_{h \geqslant 0} \#\mathcal{D}^*(\mathfrak{l}_h) \mathsf{z}^h, \tag{4.4.11}$$

so that $\mathrm{en}_{\mathrm{ht}}(\mathbf{gr}(\mathbf{ld}))$ is the generating series of the cardinalities of the lattices $\mathcal{D}^*(\mathfrak{l}_h)$, enumerated w.r.t. $h \geqslant 0$.

Theorem 4.4.2. *The series* $\mathbf{gr}(\mathbf{ld})$ *satisfies*

$$\mathbf{gr}(\mathbf{ld}) = \epsilon + \circ^{\otimes}(\mathbf{gr}(\mathbf{ld})) + \bullet^{\otimes}\big(\mathbf{gr}\big(\centerdot^{\otimes}(\Delta(\mathbf{ld}))\big)\big). \tag{4.4.12}$$

Proof. By (4.4.7) and by Relation (4.4.10c) of Proposition 4.4.1, we have

$$\mathbf{gr}(\mathbf{ld}) = \mathbf{gr}(\epsilon + \circ^{\otimes}(\mathbf{ld})) = \epsilon + \sum_{h \geqslant 0} \big(\circ^{\otimes}\big(\mathbf{gr}(\mathfrak{l}_h)\big) + \bullet^{\otimes}\big(\mathbf{gr}(\mathfrak{l}_h \centerdot \mathfrak{l}_h)\big)\big) \tag{4.4.13}$$

and the relation of the statement follows. □

Proposition 4.4.3. *The* ht*-enumeration* H_{gr} *of* $\mathbf{gr}(\mathbf{ld})$ *satisfies*

$$\mathsf{H}_{\mathrm{gr}} = 1 + \mathsf{z}\mathsf{H}_{\mathrm{gr}} + \mathsf{z}\big(\mathsf{H}_{\mathrm{gr}} \boxtimes \mathsf{H}_{\mathrm{gr}}\big). \tag{4.4.14}$$

Proof. By (4.4.1) and by Relation (4.4.10a) of Proposition 4.4.1, we have

$$\begin{aligned} \mathrm{en}_{\mathrm{ht}}(\mathbf{gr}(\centerdot^{\otimes}(\Delta(\mathbf{ld})))) &= \sum_{h \geqslant 0} \mathrm{en}_{\mathrm{ht}}\big(\mathbf{gr}(\mathfrak{l}_h \centerdot \mathfrak{l}_h)\big) \\ &= \sum_{h \geqslant 0} \mathrm{en}_{\mathrm{ht}}\big(\mathbf{gr}(\mathfrak{l}_h) \centerdot^{\otimes} \mathbf{gr}(\mathfrak{l}_h)\big) \\ &= \sum_{h \geqslant 0} \big(\mathrm{en}_{\mathrm{ht}}\big(\mathbf{gr}(\mathfrak{l}_h)\big) \uparrow \mathrm{en}_{\mathrm{ht}}\big(\mathbf{gr}(\mathfrak{l}_h)\big)\big). \end{aligned} \tag{4.4.15}$$

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

Since for any $h \geqslant 0$, all the duplicative forests appearing in $\mathbf{gr}(\mathfrak{l}_h)$ have h as height, the last member of (4.4.15) is equal to $\mathsf{H}_{\mathrm{gr}} \boxtimes \mathsf{H}_{\mathrm{gr}}$. Now, by using this identity together with (4.4.2) and Theorem 4.4.2, we obtain the stated expression for H_{gr}. □

By Proposition 3.3.3, for any $d \geqslant 1$, the cardinality of $\mathrm{M}(d)$ is $\mathbf{h}_{\mathrm{gr}}(h) := \left\langle \mathsf{z}^h, \mathsf{H}_{\mathrm{gr}} \right\rangle$ where $h = d - 1$. By Proposition 4.4.3, $\mathbf{h}_{\mathrm{gr}}$ is the integer sequence satisfying $\mathbf{h}_{\mathrm{gr}}(0) = 1$ and, for any $h \geqslant 1$,

$$\mathbf{h}_{\mathrm{gr}}(h) = \mathbf{h}_{\mathrm{gr}}(h-1) + \mathbf{h}_{\mathrm{gr}}(h-1)^2. \tag{4.4.16}$$

The sequence of the cardinalities of $\mathrm{M}(d)$, $d \geqslant 0$, starts by

$$1, 1, 2, 6, 42, 1806, 3263442, 10650056950806 \tag{4.4.17}$$

and forms Sequence A007018 of Sloane (https://oeis.org/). It is known that the sets of the planar rooted trees $\mathfrak{s}$ made of nodes of arity 0, 1, or 2 and such that any path connecting the root of $\mathfrak{s}$ to any leaf of $\mathfrak{s}$ consists in $d \geqslant 1$ nodes are enumerated by the previous sequence. Let ϕ be the map sending any such planar rooted tree $\mathfrak{s}$ to the duplicative tree obtained by replacing each node of $\mathfrak{s}$ having a single child (resp. two children) by a white (resp. black) node and by removing all nodes which have no children. It is easy to see that ϕ is a bijection between the set of the planar rooted trees described above and $\mathcal{D}^*(\mathfrak{l}_h)$ where $h = d - 1$.

4.4.3 Number of covering pairs

Let **cv** be the (1, 1)-operation on $\mathbb{K}\langle\langle \mathcal{D}^* \rangle\rangle$ satisfying, for any $\mathfrak{f} \in \mathcal{D}^*$,

$$\mathbf{cv}(\mathfrak{f}) = \sum_{\substack{\mathfrak{f}' \in \mathcal{D}^* \\ \mathfrak{f} \twoheadrightarrow \mathfrak{f}'}} \mathfrak{f}'. \tag{4.4.18}$$

Immediately from the definition of the covering relation $\twoheadrightarrow$ of $\ll$, it follows that for any $\ell \geqslant 0$, $\mathfrak{f}_1, \ldots, \mathfrak{f}_\ell \in \mathcal{D}^*$, and $\mathfrak{f} \in \mathcal{D}^*$,

$$\mathbf{cv}(\mathfrak{f}_1 \centerdot \cdots \centerdot \mathfrak{f}_\ell) = \sum_{i \in [\ell]} \mathfrak{f}_1 \centerdot \cdots \centerdot \mathfrak{f}_{i-1} \centerdot^{\otimes} \mathbf{cv}(\mathfrak{f}_i) \centerdot^{\otimes} \mathfrak{f}_{i+1} \centerdot \cdots \centerdot \mathfrak{f}_\ell, \tag{4.4.19a}$$

$$\mathbf{cv}(\bullet(\mathfrak{f})) = \bullet^{\otimes}(\mathbf{cv}(\mathfrak{f})), \tag{4.4.19b}$$

$$\mathbf{cv}(\circ(\mathfrak{f})) = \circ^{\otimes}(\mathbf{cv}(\mathfrak{f})) + \bullet^{\otimes}(\mathfrak{f} \centerdot \mathfrak{f}). \tag{4.4.19c}$$

Let **ni** be the (1, 1)-operation on $\mathbb{K}\langle\langle \mathcal{D}^* \rangle\rangle$ satisfying, for any $\mathfrak{f} \in \mathcal{D}^*$,

$$\mathbf{ni}(\mathfrak{f}) = \mathbf{cv}(\mathbf{gr}(\mathfrak{f})). \tag{4.4.20}$$

By a straightforward computation, we obtain

$$\mathbf{ni}(\mathfrak{f}) = \sum_{\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f})} \#\left\{\mathfrak{f}'' \in \mathcal{D}^*(\mathfrak{f}) : \mathfrak{f}'' \twoheadrightarrow \mathfrak{f}'\right\} \mathfrak{f}', \tag{4.4.21}$$

so that the coefficient of each $\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f})$ in $\mathbf{ni}(\mathfrak{f})$ is the number of duplicative forests admitting $\mathfrak{f}'$ as covering in $\mathcal{D}^*(\mathfrak{f})$. For instance (see at the same time Figure 3),

$\mathbf{ni}($  $) =$ $+ 3$  $+ 4$ $.$ (4.4.22)

Proposition 4.4.4. *For any $\ell \geqslant 0$, $\mathfrak{f}_1, \ldots, \mathfrak{f}_\ell \in \mathcal{D}^*$, and $\mathfrak{f} \in \mathcal{D}^*$,*

$$\mathbf{ni}(\mathfrak{f}_1 \centerdot \cdots \centerdot \mathfrak{f}_\ell) = \sum_{i \in [\ell]} \mathbf{gr}(\mathfrak{f}_1 \centerdot \cdots \centerdot \mathfrak{f}_{i-1}) \centerdot^{\otimes} \mathbf{ni}(\mathfrak{f}_i) \centerdot^{\otimes} \mathbf{gr}(\mathfrak{f}_{i+1} \centerdot \cdots \centerdot \mathfrak{f}_\ell), \tag{4.4.23a}$$

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

$$\mathbf{ni}(\bullet(\mathfrak{f})) = \bullet^{\otimes}(\mathbf{ni}(\mathfrak{f})), \tag{4.4.23b}$$

$$\mathbf{ni}(\circ(\mathfrak{f})) = \circ^{\otimes}(\mathbf{ni}(\mathfrak{f})) + \bullet^{\otimes}(\mathbf{ni}(\mathfrak{f} \centerdot \mathfrak{f})) + \bullet^{\otimes}\bigl(\centerdot^{\otimes}\bigl(\Delta\bigl(\mathbf{gr}(\mathfrak{f})\bigr)\bigr)\bigr). \tag{4.4.23c}$$

Proof. Relations (4.4.23a) and (4.4.23b) are direct consequences of Relations (4.4.19a) and (4.4.19b), and of Proposition 4.4.1. By Proposition 4.4.1 and Relations (4.4.19b) and (4.4.19c), we have

$$\begin{aligned}\mathbf{ni}(\circ(\mathfrak{f})) &= \mathbf{cv}\bigl(\circ^{\otimes}(\mathbf{gr}(\mathfrak{f})) + \bullet^{\otimes}(\mathbf{gr}(\mathfrak{f} \centerdot \mathfrak{f}))\bigr)\\ &= \sum_{\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f})} \mathbf{cv}\bigl(\circ(\mathfrak{f}')\bigr) + \bullet^{\otimes}(\mathbf{ni}(\mathfrak{f} \centerdot \mathfrak{f}))\\ &= \sum_{\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f})} \bigl(\circ^{\otimes}\bigl(\mathbf{cv}(\mathfrak{f}')\bigr) + \bullet^{\otimes}\bigl(\mathfrak{f}' \centerdot \mathfrak{f}'\bigr)\bigr) + \bullet^{\otimes}(\mathbf{ni}(\mathfrak{f} \centerdot \mathfrak{f})).\end{aligned} \tag{4.4.24}$$

This establishes (4.4.23c). □

Observe that

$$\mathrm{Supp}(\mathbf{ni}(\mathbf{ld})) = \mathcal{L} \setminus \bigl\{\mathfrak{l}_h : h \geqslant 0\bigr\} \tag{4.4.25}$$

and that the coefficient of each duplicative forest $\mathfrak{f}$ of this set is the number of duplicative forests of this same set covered by $\mathfrak{f}$. Moreover, since for any $h \geqslant 0$, all elements of $\mathcal{D}^*(\mathfrak{l}_h)$ have h as height,

$$\mathrm{en}_{\mathrm{ht}}(\mathbf{ni}(\mathbf{ld})) = \sum_{h \geqslant 0} \sum_{\mathfrak{f} \in \mathcal{D}^*(\mathfrak{l}_h)} \#\bigl\{\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{l}_h) : \mathfrak{f}' \twoheadrightarrow \mathfrak{f}\bigr\}\, \mathsf{z}^h, \tag{4.4.26}$$

so that $\mathrm{en}_{\mathrm{ht}}(\mathbf{ni}(\mathbf{ld}))$ is the generating series of the number of edges of the Hasse diagrams of the lattices $\mathcal{D}^*(\mathfrak{l}_h)$, enumerated w.r.t. $h \geqslant 0$.

Theorem 4.4.5. *The series* $\mathbf{ni}(\mathbf{ld})$ *satisfies*

$$\mathbf{ni}(\mathbf{ld}) = \circ^{\otimes}(\mathbf{ni}(\mathbf{ld})) + \bullet^{\otimes}\bigl(\mathbf{ni}\bigl(\centerdot^{\otimes}(\Delta(\mathbf{ld}))\bigr)\bigr) + \bullet^{\otimes}\bigl(\centerdot^{\otimes}\bigl(\Delta\bigl(\mathbf{gr}(\mathbf{ld})\bigr)\bigr)\bigr). \tag{4.4.27}$$

Proof. By (4.4.7) and by Relation (4.4.23c) of Proposition 4.4.4, we have

$$\begin{aligned}\mathbf{ni}(\mathbf{ld}) &= \mathbf{ni}\bigl(\epsilon + \circ^{\otimes}(\mathbf{ld})\bigr)\\ &= 0 + \sum_{h \geqslant 0} \mathbf{ni}(\circ(\mathfrak{l}_h))\\ &= \sum_{h \geqslant 0} \bigl(\circ^{\otimes}(\mathbf{ni}(\mathfrak{l}_h)) + \bullet^{\otimes}(\mathbf{ni}(\mathfrak{l}_h \centerdot \mathfrak{l}_h)) + \bullet^{\otimes}\bigl(\centerdot^{\otimes}\bigl(\Delta\bigl(\mathbf{gr}(\mathfrak{l}_h)\bigr)\bigr)\bigr)\bigr)\end{aligned} \tag{4.4.28}$$

and the relation of the statement follows. □

Proposition 4.4.6. *The* ht*-enumeration* H_{ni} *of* $\mathbf{ni}(\mathbf{ld})$ *satisfies*

$$\mathsf{H}_{\mathrm{ni}} = \mathsf{z}\mathsf{H}_{\mathrm{ni}} + \mathsf{z}\mathsf{H}_{\mathrm{gr}} + 2\mathsf{z}\bigl(\mathsf{H}_{\mathrm{ni}} \boxtimes \mathsf{H}_{\mathrm{gr}}\bigr). \tag{4.4.29}$$

Proof. Observe first that we have

$$\begin{aligned}\mathrm{en}_{\mathrm{ht}}\bigl(\centerdot^{\otimes}(\Delta(\mathbf{gr}(\mathbf{ld})))\bigr) &= \sum_{h \geqslant 0} \sum_{\mathfrak{f} \in \mathcal{D}^*(\mathfrak{l}_h)} \mathrm{en}_{\mathrm{ht}}(\mathfrak{f} \centerdot \mathfrak{f})\\ &= \sum_{h \geqslant 0} \sum_{\mathfrak{f} \in \mathcal{D}^*(\mathfrak{l}_h)} \mathrm{en}_{\mathrm{ht}}(\mathfrak{f})\\ &= \mathsf{H}_{\mathrm{gr}}.\end{aligned} \tag{4.4.30}$$

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

Moreover, by (4.4.1) and by Relation (4.4.23a) of Proposition 4.4.4, we have

$$\begin{aligned}\mathrm{en}_{\mathrm{ht}}\big(\mathbf{ni}\big(\boldsymbol{\cdot}^{\otimes}(\Delta(\mathbf{ld}))\big)\big) &= \sum_{h\geqslant 0} \mathrm{en}_{\mathrm{ht}}(\mathbf{ni}(\mathfrak{l}_h \boldsymbol{\cdot} \mathfrak{l}_h)) \\ &= \sum_{h\geqslant 0} \mathrm{en}_{\mathrm{ht}}\big(\mathbf{gr}(\mathfrak{l}_h) \boldsymbol{\cdot}^{\otimes} \mathbf{ni}(\mathfrak{l}_h) + \mathbf{ni}(\mathfrak{l}_h) \boldsymbol{\cdot}^{\otimes} \mathbf{gr}(\mathfrak{l}_h)\big) \\ &= \sum_{h\geqslant 0} \big(\mathrm{en}_{\mathrm{ht}}\big(\mathbf{gr}(\mathfrak{l}_h)\big) \uparrow \mathrm{en}_{\mathrm{ht}}(\mathbf{ni}(\mathfrak{l}_h)) + \mathrm{en}_{\mathrm{ht}}(\mathbf{ni}(\mathfrak{l}_h)) \uparrow \mathrm{en}_{\mathrm{ht}}\big(\mathbf{gr}(\mathfrak{l}_h)\big)\big).\end{aligned} \tag{4.4.31}$$

Since for any $h \geqslant 0$, all duplicative forests appearing in $\mathbf{gr}(\mathfrak{l}_h)$ and in $\mathbf{ni}(\mathfrak{l}_h)$ have h as height, the last member of (4.4.31) is equal to $\mathsf{H}_{\mathrm{gr}} \boxtimes \mathsf{H}_{\mathrm{ni}} + \mathsf{H}_{\mathrm{ni}} \boxtimes \mathsf{H}_{\mathrm{gr}}$. Now by using this identity together with (4.4.30), (4.4.2), and Theorem 4.4.5, we obtain the stated expression for H_{ni}. □

By Proposition 3.3.3, for any $d \geqslant 1$, the number of edges in the Hasse diagram of $\mathrm{M}(d)$ is $\mathbf{h}_{\mathrm{ni}}(h) := \left\langle \mathsf{z}^h, \mathsf{H}_{\mathrm{ni}} \right\rangle$ where $h = d - 1$. By Proposition 4.4.6, $\mathbf{h}_{\mathrm{ni}}$ is the integer sequence satisfying $\mathbf{h}_{\mathrm{ni}}(0) = 0$ and, for any $h \geqslant 1$,

$$\mathbf{h}_{\mathrm{ni}}(h) = \mathbf{h}_{\mathrm{ni}}(h-1) + \mathbf{h}_{\mathrm{gr}}(h-1) + 2\mathbf{h}_{\mathrm{ni}}(h-1)\,\mathbf{h}_{\mathrm{gr}}(h-1), \tag{4.4.32}$$

where $\mathbf{h}_{\mathrm{gr}}$ is the integer sequence defined in Section 4.4.2. The sequence of the number of edges of the Hasse diagram of $\mathrm{M}(d)$, $d \geqslant 0$, starts by

$$0, 0, 1, 7, 97, 8287, 29942737, 195432804247687. \tag{4.4.33}$$

This sequence does not appear in Sloane (https://oeis.org/) for the time being.

4.4.4 Number of intervals

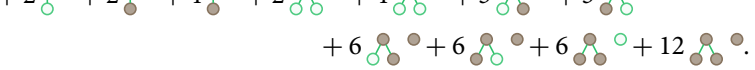
Let $\mathbf{ns}$ be the $(1, 1)$-operation on $\mathbb{K}\langle\langle \mathcal{D}^* \rangle\rangle$ satisfying, for any $\mathfrak{f} \in \mathcal{D}^*$,

$$\mathbf{ns}(\mathfrak{f}) = \mathbf{gr}(\mathbf{gr}(\mathfrak{f})). \tag{4.4.34}$$

By a straightforward computation, we obtain

$$\mathbf{ns}(\mathfrak{f}) = \sum_{\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f})} \#\big[\mathfrak{f}, \mathfrak{f}'\big]\mathfrak{f}', \tag{4.4.35}$$

so that the coefficient of each $\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f})$ in $\mathbf{ns}(\mathfrak{f})$ is the number of duplicative forests smaller than or equal as $\mathfrak{f}'$ in $\mathcal{D}^*(\mathfrak{f})$. For instance (see at the same time Figure 3),

$\mathbf{ns}($  $) =$

(4.4.36)

Contrary to what we have undertaken in Sections 4.4.2 and 4.4.3 where we have provided direct recursive expressions to compute $\mathbf{gr}(\mathfrak{f})$ and $\mathbf{ni}(\mathfrak{f})$, we fail to provide similar expressions for $\mathbf{ns}(\mathfrak{f})$. The trick here consists in considering first a slightly different series depending on a parameter $k \geqslant 1$ which can be seen as a catalytic parameter. With this in mind, let for any $k \geqslant 1$, $\mathbf{md}_k$ be the $(1, k)$-operation on $\mathbb{K}\langle\langle \mathcal{D}^* \rangle\rangle$ satisfying, for any $\mathfrak{f} \in \mathcal{D}^*$,

$$\mathbf{md}_k(\mathfrak{f}) = \sum_{\substack{\mathfrak{g}_1,\dots,\mathfrak{g}_k \in \mathcal{D}^*(\mathfrak{f}) \\ \mathfrak{g}_1 \wedge \dots \wedge \mathfrak{g}_k = \mathfrak{f}}} \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k, \tag{4.4.37}$$

where $\wedge$ is the meet operation of the duplicative forest lattices introduced in Section 3.2.3. We call $\mathbf{md}_k(\mathfrak{f})$ the *k-meet decomposition* of $\mathfrak{f}$. Observe that $\mathbf{md}_1$ is the identity map.

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

Lemma 4.4.7. *For any $k \geqslant 1$ and $\mathfrak{f} \in \mathcal{D}^*$,*

$$\mathbf{md}_k(\mathbf{ns}(\mathfrak{f})) = \sum_{\substack{\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f}) \\ \mathfrak{g}_1, \ldots, \mathfrak{g}_k \in \mathcal{D}^*(\mathfrak{f}')}} \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k. \tag{4.4.38}$$

Proof. We have

$$\begin{aligned}
\mathbf{md}_k(\mathbf{ns}(\mathfrak{f})) &= \sum_{\substack{\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f}) \\ \mathfrak{f}'' \in \mathcal{D}^*(\mathfrak{f}')}} \sum_{\substack{\mathfrak{g}_1, \ldots, \mathfrak{g}_k \in \mathcal{D}^*(\mathfrak{f}) \\ \mathfrak{g}_1 \wedge \ldots \wedge \mathfrak{g}_k = \mathfrak{f}''}} \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k \\
&= \sum_{\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f})} \sum_{\substack{\mathfrak{g}_1, \ldots, \mathfrak{g}_k \in \mathcal{D}^*(\mathfrak{f}) \\ \mathfrak{g}_1 \wedge \ldots \wedge \mathfrak{g}_k = \mathfrak{f}'}} \#\big[\mathfrak{f}, \mathfrak{f}'\big] \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k \\
&= \sum_{\mathfrak{g}_1, \ldots, \mathfrak{g}_k \in \mathcal{D}^*(\mathfrak{f})} \#[\mathfrak{f}, \mathfrak{g}_1 \wedge \ldots \wedge \mathfrak{g}_k] \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k \\
&= \sum_{\mathfrak{g}_1, \ldots, \mathfrak{g}_k \in \mathcal{D}^*(\mathfrak{f})} \sum_{\substack{\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f}) \\ \mathfrak{f}' \ll \mathfrak{g}_1, \ldots, \mathfrak{f}' \ll \mathfrak{g}_k}} \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k,
\end{aligned} \tag{4.4.39}$$

showing the stated identity. □

Proposition 4.4.8. *For any $k \geqslant 1$, $\ell \geqslant 0$, $\mathfrak{f}_1, \ldots, \mathfrak{f}_\ell \in \mathcal{D}^*$, and $\mathfrak{f} \in \mathcal{D}^*$,*

$$\mathbf{md}_k(\mathbf{ns}(\mathfrak{f}_1 \centerdot \ldots \centerdot \mathfrak{f}_\ell)) = \mathbf{md}_k(\mathbf{ns}(\mathfrak{f}_1)) \centerdot^{\otimes_k} \ldots \centerdot^{\otimes_k} \mathbf{md}_k(\mathbf{ns}(\mathfrak{f}_\ell)), \tag{4.4.40a}$$

$$\mathbf{md}_k(\mathbf{ns}(\bullet(\mathfrak{f}))) = \bullet^{\otimes_k}(\mathbf{md}_k(\mathbf{ns}(\mathfrak{f}))), \tag{4.4.40b}$$

$$\mathbf{md}_k(\mathbf{ns}(\circ(\mathfrak{f}))) = \sum_{u \in \{\circ, \bullet\}^k} \mathrm{mg}_u\big(\mathbf{md}_{k+|u|_\bullet}(\mathbf{ns}(\mathfrak{f}))\big) + \bullet^{\otimes_k}(\mathbf{md}_k(\mathbf{ns}(\mathfrak{f} \centerdot \mathfrak{f}))). \tag{4.4.40c}$$

Proof. Relations (4.4.40a) and (4.4.40b) are direct consequences of Lemmas 3.2.1 and 4.4.7. By Lemmas 3.2.1 and 4.4.7,

$$\begin{aligned}
\mathbf{md}_k(\mathbf{ns}(\circ(\mathfrak{f}))) &= \sum_{\substack{\mathfrak{f}' \in \mathcal{D}^*(\circ(\mathfrak{f})) \\ \mathfrak{g}_1, \ldots, \mathfrak{g}_k \in \mathcal{D}^*(\mathfrak{f}')}} \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k \\
&= \sum_{\substack{\circ(\mathfrak{f}') \in \mathcal{D}^*(\circ(\mathfrak{f})) \\ \mathfrak{g}_1, \ldots, \mathfrak{g}_k \in \mathcal{D}^*(\circ(\mathfrak{f}'))}} \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k + \sum_{\substack{\bullet(\mathfrak{f}') \in \mathcal{D}^*(\circ(\mathfrak{f})) \\ \mathfrak{g}_1, \ldots, \mathfrak{g}_k \in \mathcal{D}^*(\bullet(\mathfrak{f}'))}} \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k \\
&= \sum_{\substack{\mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f}) \\ \mathfrak{g}_1, \ldots, \mathfrak{g}_k \in \mathcal{D}^*(\circ(\mathfrak{f}'))}} \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k + \sum_{\substack{\mathfrak{f}'_1, \mathfrak{f}'_2 \in \mathcal{D}^*(\mathfrak{f}) \\ \mathfrak{g}_1, \ldots, \mathfrak{g}_k \in \mathcal{D}^*(\bullet(\mathfrak{f}'_1 \mathfrak{f}'_2))}} \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k \\
&= \sum_{\substack{u \in \{\circ, \bullet\}^k \\ \mathfrak{f}' \in \mathcal{D}^*(\mathfrak{f}) \\ \mathfrak{g}_1, \ldots, \mathfrak{g}_k \in \mathcal{D}^*(\circ(\mathfrak{f}')) \\ \mathrm{rts}(\mathfrak{g}_1 \ldots \mathfrak{g}_k) = u}} \mathfrak{g}_1 \otimes \cdots \otimes \mathfrak{g}_k + \sum_{\substack{\mathfrak{f}'_1, \mathfrak{f}'_2 \in \mathcal{D}^*(\mathfrak{f}) \\ \mathfrak{g}'_1, \ldots, \mathfrak{g}'_k \in \mathcal{D}^*(\mathfrak{f}'_1) \\ \mathfrak{g}''_1, \ldots, \mathfrak{g}''_k \in \mathcal{D}^*(\mathfrak{f}'_2)}} \bullet\big(\mathfrak{g}'_1 \centerdot \mathfrak{g}''_1\big) \otimes \cdots \otimes \bullet\big(\mathfrak{g}'_k \centerdot \mathfrak{g}''_k\big) \\
&= \sum_{u \in \{\circ, \bullet\}^k} \mathrm{mg}_u\big(\mathbf{md}_{k+|u|_\bullet}(\mathbf{ns}(\mathfrak{f}))\big) + \bullet^{\otimes_k}(\mathbf{md}_k(\mathbf{ns}(\mathfrak{f} \centerdot \mathfrak{f}))),
\end{aligned} \tag{4.4.41}$$

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

where for any duplicative forest $\mathfrak{f}$, $\mathrm{rts}(\mathfrak{f})$ is the word on $\{\circ, \bullet\}$ containing from left to right the roots of the trees forming $\mathfrak{f}$. This shows (4.4.40c). □

Observe that $\mathrm{Supp}(\mathbf{ns}(\mathbf{ld})) = \mathcal{L}$ and that the coefficient of each duplicative forest $\mathfrak{f}$ of this set is the number of duplicative forests of this same set smaller than or equal as $\mathfrak{f}$. Moreover, since for any $h \geqslant 0$, all elements of $\mathcal{D}^*(\mathfrak{l}_h)$ have h as height,

$$\mathrm{en}_{\mathrm{ht}}(\mathbf{ns}(\mathbf{ld})) = \sum_{h \geqslant 0} \sum_{\mathfrak{f} \in \mathcal{D}^*(\mathfrak{l}_h)} \#[\mathfrak{l}_h, \mathfrak{f}]\, \mathsf{z}^h, \tag{4.4.42}$$

so that $\mathrm{en}_{\mathrm{ht}}(\mathbf{ns}(\mathbf{ld}))$ is the generating series of the number of intervals of the lattices $\mathcal{D}^*(\mathfrak{l}_h)$, enumerated w.r.t. $h \geqslant 0$.

Theorem 4.4.9. *The series* $\mathbf{ns}(\mathbf{ld})$ *satisfies* $\mathbf{ns}(\mathbf{ld}) = \mathbf{md}_1(\mathbf{ns}(\mathbf{ld}))$ *where, for any* $k \geqslant 1$, *the series* $\mathbf{md}_k(\mathbf{ns}(\mathbf{ld}))$ *satisfies*

$$\mathbf{md}_k(\mathbf{ns}(\mathbf{ld})) = \underbrace{\epsilon \otimes \cdots \otimes \epsilon}_{k \text{ terms}} + \sum_{u \in \{\circ,\bullet\}^k} \mathrm{mg}_u\big(\mathbf{md}_{k+|u|_\bullet}(\mathbf{ns}(\mathbf{ld}))\big) + \bullet^{\otimes_{\mathrm{k}}}\big(\mathbf{md}_k\big(\mathbf{ns}\big(\boldsymbol{.}^{\otimes}(\Delta(\mathbf{ld}))\big)\big)\big). \tag{4.4.43}$$

Proof. First, the relation $\mathbf{ns}(\mathbf{ld}) = \mathbf{md}_1(\mathbf{ns}(\mathbf{ld}))$ holds since $\mathbf{md}_1$ is the identity map. By (4.4.7) and by Proposition 4.4.8, we have

$$\begin{aligned}\mathbf{md}_k(\mathbf{ns}(\mathbf{ld})) &= \mathbf{md}_k\big(\mathbf{ns}\big(\epsilon + \circ^{\otimes}(\mathbf{ld})\big)\big) \\ &= \underbrace{\epsilon \otimes \cdots \otimes \epsilon}_{k \text{ terms}} + \sum_{h \geqslant 0} \left(\sum_{u \in \{\circ,\bullet\}^k} \mathrm{mg}_u\big(\mathbf{md}_{k+|u|_\bullet}(\mathbf{ns}(\mathfrak{l}_h))\big) + \bullet^{\otimes_{\mathrm{k}}}(\mathbf{md}_k(\mathbf{ns}(\mathfrak{l}_h \boldsymbol{.} \mathfrak{l}_h))) \right)\end{aligned} \tag{4.4.44}$$

and the relation of the statement follows. □

Proposition 4.4.10. *The* ht*-enumeration* H_{ns} *of* $\mathbf{ns}(\mathbf{ld})$ *satisfies* $\mathsf{H}_{\mathrm{ns}} = \mathsf{H}_{\mathrm{ns}}^{(1)}$ *where, for any* $k \geqslant 1$, $\mathsf{H}_{\mathrm{ns}}^{(k)}$ *is the* ht*-enumeration of* $\mathbf{md}_k(\mathbf{ns}(\mathbf{ld}))$ *which satisfies*

$$\mathsf{H}_{\mathrm{ns}}^{(k)} = 1 + \mathsf{z}\Big(\mathsf{H}_{\mathrm{ns}}^{(k)} \boxtimes \mathsf{H}_{\mathrm{ns}}^{(k)}\Big) + \mathsf{z} \sum_{i \in [\![k]\!]} \binom{k}{i} \mathsf{H}_{\mathrm{ns}}^{(k+i)}. \tag{4.4.45}$$

Proof. By Relation (4.4.40a) of Proposition 4.4.8, we have

$$\begin{aligned}\mathrm{en}_{\mathrm{ht}}\big(\mathbf{md}_k\big(\mathbf{ns}\big(\boldsymbol{.}^{\otimes}(\Delta(\mathbf{ld}))\big)\big)\big) &= \sum_{h \geqslant 0} \mathrm{en}_{\mathrm{ht}}(\mathbf{md}_k(\mathbf{ns}(\mathfrak{l}_h \boldsymbol{.} \mathfrak{l}_h))) \\ &= \sum_{h \geqslant 0} \mathrm{en}_{\mathrm{ht}}\big(\mathbf{md}_k(\mathbf{ns}(\mathfrak{l}_h)) \boldsymbol{.}^{\otimes_{\mathrm{k}}} \mathbf{md}_k(\mathbf{ns}(\mathfrak{l}_h))\big) \\ &= \sum_{h \geqslant 0} (\mathrm{en}_{\mathrm{ht}}(\mathbf{md}_k(\mathbf{ns}(\mathfrak{l}_h))) \uparrow \mathrm{en}_{\mathrm{ht}}(\mathbf{md}_k(\mathbf{ns}(\mathfrak{l}_h)))).\end{aligned} \tag{4.4.46}$$

Since for any $h \geqslant 0$, all the duplicative forests of the tensors $\mathfrak{f}_1 \otimes \cdots \otimes \mathfrak{f}_k$ appearing in $\mathbf{md}_k(\mathbf{ns}(\mathfrak{l}_h))$ have h as height, the last member of (4.4.46) is equal to $\mathsf{H}_{\mathrm{ns}}^{(k)} \boxtimes \mathsf{H}_{\mathrm{ns}}^{(k)}$. Moreover, by (4.4.5), for any $u \in \{\circ, \bullet\}^k$,

$$\mathrm{en}_{\mathrm{ht}}\big(\mathrm{mg}_u\big(\mathbf{md}_{k+|u|_\bullet}(\mathbf{ns}(\mathbf{ld}))\big)\big) = \mathsf{z}\, \mathrm{en}_{\mathrm{ht}}\big(\mathbf{md}_{k+|u|_\bullet}(\mathbf{ns}(\mathbf{ld})\big) = \mathsf{z}\mathsf{H}_{\mathrm{ns}}^{(k+|u|_\bullet)}. \tag{4.4.47}$$

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

Now, by using these two identities together with Theorem 4.4.9, and by the fact that the number of words of length k on $\{\circ, \bullet\}$ having exactly $i \in [\![k]\!]$ occurrences of $\bullet$ is $\binom{k}{i}$, we obtain the stated expression for $\mathsf{H}_{\mathrm{ns}}^{(k)}$. Finally, we have $\mathsf{H}_{\mathrm{ns}} = \mathsf{H}_{\mathrm{ns}}^{(1)}$ since $\mathbf{ns}(\mathbf{ld}) = \mathbf{md}_1(\mathbf{ns}(\mathbf{ld}))$. □

By Proposition 3.3.3, for any $d \geqslant 1$, the number of intervals in $\mathrm{M}(d)$ is $\mathbf{h}_{\mathrm{ns}}^{(1)}(h) := \left\langle \mathsf{z}^h, \mathsf{H}_{\mathrm{ns}} \right\rangle$ where $h = d - 1$. By Proposition 4.4.10, for any $k \geqslant 1$, $\mathbf{h}_{\mathrm{ns}}^{(k)}$ is the integer sequence satisfying $\mathbf{h}_{\mathrm{ns}}^{(k)}(0) = 1$ and, for any $h \geqslant 1$,

$$\mathbf{h}_{\mathrm{ns}}^{(k)}(h) = \mathbf{h}_{\mathrm{ns}}^{(k)}(h-1)^2 + \sum_{i \in [\![k]\!]} \binom{k}{i} \mathbf{h}_{\mathrm{ns}}^{(k+i)}(h-1). \tag{4.4.48}$$

The sequence of the cardinalities of $\mathrm{M}(d)$, $d \geqslant 0$, starts by

$$1, 1, 3, 17, 371, 144513, 20932611523, 438176621806663544657. \tag{4.4.49}$$

This sequence does not appear in Sloane (https://oeis.org/) for the time being.

5. Open Questions and Future Work

We have studied in this work a CLS having a lot of rich combinatorial properties despite its simplicity. This can be considered as the prototypical example for this kind of investigation. We could expect to have also similar combinatorial properties (related to poset and lattice properties, and quantitative data) for more complex CLS. In addition to the obvious possible future investigations consisting in studying in a similar fashion some other CLS, we describe here three open questions raised by this work.

The description of minimal and maximal elements of $\mathcal{P}$ uses a notion of pattern avoidance in terms (see Proposition 3.1.4). This is a general fact: when a CLS $(\mathfrak{G}, \to)$ has the poset property, its minimal (resp. maximal) elements are the terms avoiding some terms deduced from the ones appearing as right-hand (resp. left-hand) members of $\to$. Such an enumerative problem has been considered in Khoroshkin and Piontkovski (2015), Giraudo (2020) for the particular case of linear terms. The question here concerns the general enumeration of terms avoiding a set of other terms wherein multiple occurrences of a same variable are allowed.

We have shown that the Mockingbird CLS has the poset property and is rooted (Proposition 3.1.2), and has the lattice property (Theorem 3.3.4) by employing some specific reasoning from the definition of the Mockingbird basic combinator $\mathbf{M}$. A question here concerns the existence of a general criterion to decide if a CLS has the poset (resp. lattice) property and if it is rooted. All this seems independent from the property of a CLS to be hierarchical because, among others, the CLS containing a basic combinator $\mathbf{X}$ such that $\mathbf{X}$ has order 3 and $\mathfrak{t}_{\mathbf{X}} := \mathsf{x}_3(\mathsf{x}_1\mathsf{x}_1\mathsf{x}_2)$ has not the poset property because we have $\mathfrak{t} \Rightarrow \mathfrak{t}'$ and $\mathfrak{t}' \Rightarrow \mathfrak{t}$ with $\mathfrak{t} := \mathbf{X}\,\mathbf{X}\,\mathbf{X}(\mathbf{X}\,\mathbf{X}(\mathbf{X}\,\mathbf{X}))$ and $\mathfrak{t}' := \mathbf{X}\,\mathbf{X}(\mathbf{X}\,\mathbf{X})(\mathbf{X}\,\mathbf{X}\,\mathbf{X})$ while $\mathbf{X}$ is hierarchical.

Finally, in Proposition 2.2.3, we have shown that being hierarchical is a sufficient condition for a CLS $\mathcal{C}$ to be locally finite. The question in this context consists in obtaining a necessary and sufficient condition for this last property. Observe that being nonerasing is necessary because, by assuming that $\mathcal{C}$ contains a basic combinator $\mathbf{X}$ such that $\mathbf{X}$ has order $n \geqslant 2$ and that there is $i \in [n]$ such that x_i does not appear in $\mathfrak{t}_{\mathbf{X}}$, we have $\mathbf{X}\,\mathsf{x}_1 \ldots \mathsf{x}_{i-1}\mathfrak{s}\mathsf{x}_{i+1} \ldots \mathsf{x}_n \Rightarrow \mathfrak{t}_{\mathbf{X}}$ for any term $\mathfrak{s}$, showing that the connected component of $\mathfrak{t}_{\mathbf{X}}$ in $\mathrm{G}_{\mathcal{C}}$ is infinite.

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press

References

Barendregt, H. P. (1981). *The Lambda Calculus, its Syntax and Semantics*, Studies in Logic and the Foundations of Mathematics, vol. 103, Elsevier. Revised edition.

Barendregt, H., Endrullis, J., Klop, J. W. and Waldmann, J. (2017). Dance of the starlings. In: *Raymond Smullyan on Self Reference*, Outstanding Contributions to Logic, vol. 14, Cham, Springer, 67–111.

Bendkowski, M., Grygiel, K. and Zaionc, M. (2017). On the likelihood of normalization in combinatory logic. *Journal of Logic and Computation* **27** (7) 2251–2269.

Bimbó, K. (2011). *Combinatory Logic: Pure, Applied and Typed*, Discrete Mathematics and Its Applications, New York, Taylor & Francis.

Bezem, M., Klop, J. W., de Vrijer, R. and Terese. (2003). *Term Rewriting Systems*, Cambridge Tracts in Theoretical Computer Science, Cambridge, Cambridge University Press.

Berry, G. and Lévy, J.-J. (1979). Minimal and optimal computations of recursive programs. *Journal of the ACM* **26** (1) 148–175.

Baader, F. and Nipkow, T. (1998). *Term Rewriting and All That*, Cambridge, Cambridge University Press.

Baril, J. L. and Pallo, J. M. (2006). The phagocyte lattice of Dyck words. *Order* **23** (2–3) 97–107.

Baril, J. L. and Pallo, J. M. (2008). The pruning-grafting lattice of binary trees. *Theoretical Computer Science* **409** (3) 382–393.

Curry, H. B. and Feys, R. (1958). *Combinatory Logic*, vol. 1, Amsterdam, North-Holland Publishing Company.

Curry, H. B. (1930). Grundlagen der Kombinatorischen Logik. *American Journal of Mathematics* **52** (4) 789–834.

David, R., Grygiel, K., Kozik, J., Raffalli, C., Theyssier, G. and Zaionc, M. (2013). Asymptotically almost all λ-terms are strongly normalizing. *Logical Methods in Computer Science* **9** (1) 1.

Dulucq, S. and Penaud, J.-G. (2002). Interprétation bijective d'une récurrence des nombres de Motzkin. *Discrete Mathematics* **256** (3) 671–676.

Giraudo, S. (2020). Tree series and pattern avoidance in syntax trees. *Journal of Combinatorial Theory A* **176**.

Giraudo, S. (2022). Mockingbird lattices. *Formal Power Series and Algebraic Combinatorics* **86B** (3).

Hindley, J. R. and Seldin, J. P. (2008). *Lambda-Calculus and Combinators, an Introduction*, Cambridge, Cambridge University Press.

Khoroshkin, A. and Piontkovski, D. (2015). On generating series of finitely presented operads. *Journal of Algebra* **426** 377–429.

Kreweras, G. (1972). Sur les partitions non croisées d'un cycle. *Discrete Mathematics* **1** (4) 333–350.

Probst, D. and Studer, T. (2001). How to normalize the Jay. *Theoretical Computer Science* **254** (1–2) 677–681.

Rosen, B. K. (1973). Tree-manipulating systems and Church-Rosser theorems. *Journal of the Association for Computing Machinery* **20** 160–187.

Schönfinkel, M. (1924). Über die Bausteine der mathematischen Logik. *Mathematische Annalen* **92** 305–316.

https://oeis.org/ Sloane, N. J. A. The On-Line Encyclopedia of Integer Sequences. https://oeis.org/.

Smullyan, R. (1985). *To Mock a Mockingbird*, New York, Alfred A. Knopf, Inc.

Stanley, R. P. (1975). The Fibonacci lattice. *Fibonacci Quarterly* **13** (3) 215–232.

Statman, R. (1989). The word problem for Smullyan's lark combinator is decidable. *Journal of Symbolic Computation* **7** (2) 103–112.

Statman, R. (2000). On the word problem for combinators. In: *11th International Conference on Rewriting Techniques and Applications*, Lecture Notes in Computer Science, vol. 1833, 203–213.

Stanley, R. P. (2011). *Enumerative Combinatorics*, 2nd ed., vol. 1, Cambridge, Cambridge University Press.

Statman, R. (2011). To type a Mockingbird. *Unpublished*. Draft available at http://tlca.di.unito.it/PAPER/TypeMock.pdf.

Statman, R. (2017). Some tweets about Mockingbirds. In: *Raymond Smullyan on Self Reference*, Outstanding Contributions to Logic, vol. 14, Cham, Springer, 113–122.

Sprenger, M. and Wymann-Böni, M. (1993). How to decide the lark. *Theoretical Computer Science* **110** (2) 419–432.

Tamari, D. (1962). The algebra of bracketings and their enumeration. *Nieuw Archief voor Wiskunde* **10** (3) 131–146.

Taylor, W. (1993). Abstract clone theory. In: *Algebras and Orders*, NATO Advanced Science Institutes, Series C: Mathematical, vol. 389, Dordrecht, Kluwer Academic Publishers, 507–530.

Venturini Zilli, M. (1984). Reduction graphs in the lambda calculus. *Theoretical Computer Science* **29** (3) 251–275.

Waldmann, J. (2000). The combinator **S**. *Information and Computation* **159** (1–2) 2–21.

Wolfram, S. (2021). *Combinators: A Centennial View*, Champaign, Wolfram Media Inc.

Cite this article: Giraudo S (2022). The combinator **M** and the Mockingbird lattice. *Mathematical Structures in Computer Science* **32**, 271–299. https://doi.org/10.1017/S0960129522000354

https://doi.org/10.1017/S0960129522000354 Published online by Cambridge University Press