

# Cost Automata, Safe Schemes, and Downward Closures

**David Barozzini\***

*Institute of Informatics  
University of Warsaw  
Warsaw, Poland  
dbarozzini@mimuw.edu.pl*

**Thomas Colcombet<sup>†</sup>**

*IRIF-CNRS-Université de Paris  
Paris, France  
thomas.colcombet@irif.fr*

**Lorenzo Clemente\***

*Institute of Informatics  
University of Warsaw  
Warsaw, Poland  
clementelorenzo@gmail.com*

**Pawel Parys\***

*Institute of Informatics  
University of Warsaw  
Warsaw, Poland  
parys@mimuw.edu.pl*

---

**Abstract.** Higher-order recursion schemes are an expressive formalism used to define languages of finite and infinite ranked trees. They extend regular and context-free grammars, and are equivalent in expressive power to the simply typed  $\lambda Y$ -calculus and collapsible pushdown automata. In this work we prove, under a syntactical constraint called safety, decidability of the model-checking problem for recursion schemes against properties defined by alternating B-automata, an extension of alternating parity automata over infinite trees with a boundedness acceptance condition. We then exploit this result to show how to compute downward closures of languages of finite trees recognized by safe recursion schemes.

**Keywords:** Cost logics, cost automata, downward closures, higher-order recursion schemes, safe recursion schemes

## 1. Introduction

Higher-order functions are nowadays widely used not only in functional programming languages such as Haskell and the OCAML family, but also in mainstream languages such as Java, JavaScript, Python, and C++. *Recursion schemes* are faithful and algorithmically manageable abstractions of the control flow of higher-order programs [1]. A deterministic recursion scheme normalizes into a possibly

---

\*Author supported by the National Science Centre, Poland (grant no. 2016/22/E/ST6/00041).

<sup>†</sup>Author supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No.670624), and the DeLTA ANR project (ANR-16-CE40-0007).

infinite Böhm tree, and in this respect recursion schemes can equivalently be presented as simply-typed lambda-terms using a higher-order fixpoint combinator  $Y$  [2]. There are also (nontrivial) inter-reductions between recursion schemes and the equi-expressive formalisms of collapsible higher-order pushdown automata [3] and ordered tree-pushdown automata [4]. In another semantics, also used in this paper, nondeterministic recursion schemes are recognizers of languages of finite trees, and in this view they are also known as higher-order OI grammars [5, 6], generalising indexed grammars [7] (which are recursion schemes of order two) and ordered multi-pushdown automata [8].

The most celebrated algorithmic result in the analysis of recursion schemes is decidability of the model-checking problem against properties expressed in monadic second-order logic (MSO): given a recursion scheme  $\mathcal{G}$  and an MSO sentence  $\varphi$ , one can decide whether the Böhm tree generated by  $\mathcal{G}$  satisfies  $\varphi$  [9]. This fundamental result has been reproved several times, that is, using collapsible higher-order pushdown automata [10], intersection types [11], Krivine machines [12], order-reducing transformations [13], and it has been extended in diverse directions such as global model checking [14], logical reflection [15], effective selection [16], and a transfer theorem via models of lambda-calculus [17]. When the input property is given as an MSO formula, the model-checking problem is non-elementary already for trees of order 0 (regular trees) [18]; when the input property is presented as a parity tree automaton (which is equi-expressive with MSO on trees, but less succinct), the MSO model-checking problem for recursion schemes of order  $n$  is complete for  $n$ -fold exponential time [9]. Despite these hardness results, the model-checking problem can be solved efficiently on multiple nontrivial examples, thanks to the development of several recursion-scheme model checkers [1, 19, 20, 21, 22].

**Unboundedness problems I: Diagonal problem and downward closures.** Recently, an increasing interest has arisen for model checking quantitative properties going beyond the expressive power of MSO. The *diagonal problem* is an example of a quantitative property not expressible in MSO. Over words, the problem asks, for a given set of letters  $\Sigma$  and a language of finite words  $\mathcal{L}$ , whether for every  $n \in \mathbb{N}$  there is a word in  $\mathcal{L}$  where every letter from  $\Sigma$  occurs at least  $n$  times. The diagonal problem for languages of finite words recognized by recursion schemes is decidable [23, 24, 25].

The class of languages of finite words recognized by recursion schemes form a so-called *full trio* (i.e., it is closed under regular transductions) and for full trios decidability of the diagonal problem has interesting algorithmic consequences, such as computability of downward closures [26, 27] and decidability of separability by piecewise testable languages [28].

The problem of computing downward closures is an important problem in its own right. The *downward closure* of a language  $\mathcal{L}$  of finite trees is the set  $\mathcal{L}\downarrow$  of all trees that can be homeomorphically embedded into some tree in  $\mathcal{L}$ . By Higman's lemma [29], the embedding relation on finite ranked trees is a well quasi-order. Consequently, the downward closure  $\mathcal{L}\downarrow$  of an arbitrary set of trees  $\mathcal{L}$  is always a regular language. The downward closure of a language offers a nontrivial regular abstraction thereof: even though the actual count of letters is lost, their limit properties are preserved, as well as their order of appearance. We say that the downward closure is *computable* when a finite automaton for  $\mathcal{L}\downarrow$  can be effectively constructed (which is not true in general). Downward closures are computable for a wide class of languages of finite words such as those recognized by context-free grammars [30, 31, 32], Petri nets [33], stacked counter automata [34], context-free FIFO rewriting systems and OL-systems [35],

second-order pushdown automata [26], higher-order pushdown automata [24], and (possibly unsafe) recursion schemes over words [23]. Over finite trees, it is known that downward closures are computable for the class of regular tree languages [36]. We are not aware of such computability results for other classes of languages of finite trees.

**Unboundedness problems II: B-automata.** In another line of research, B-automata, and among them *alternating B-automata*, have been put forward as a quantitative extension to MSO [37, 38, 39, 40, 41, 42]. They extend alternating automata over infinite trees [43, Chapter 9] by nonnegative integer counters that can be incremented or reset to zero. The extra counters do not constrain the availability of transitions during a run (unlike in other superficially similar models, such as counter machines), but are used in order to define the acceptance condition: an infinite tree is *n-accepted* if  $n$  is a bound on the values taken by the counters during an accepting run of the automaton over it.

The *universality problem* consists in deciding whether for every tree there is a bound  $n$  for which it is  $n$ -accepted. The *boundedness problem* asks whether there exists a bound  $n$  for which all trees are  $n$ -accepted. These two problems are closely related. Their decidability is an important open problem in the field, and proving decidability of the boundedness problem would solve the long standing nondeterministic Mostowski index problem [44]. However, though open in general, the boundedness problem is known to be decidable over finite words [38], finite trees [39], infinite words [40], as well as over infinite trees for its weak [41] and the more general quasi-weak [42] variant.

Another expressive formalism for unboundedness properties beyond MSO is MSO+U, which extends MSO by a novel quantifier “ $UX.\varphi$ ” [45] stating that there exist arbitrarily large finite sets  $X$  satisfying  $\varphi$ . This logic is incomparable with B-automata. The model-checking problem of recursion schemes against its weak fragment WMSO+U, where monadic second-order quantifiers are restricted to finite sets, is decidable [46].

**Contributions.** Our first contribution is decidability of the model-checking problem of properties expressed by alternating B-automata for an expressive class of recursion schemes called *safe recursion schemes*. As generators of infinite trees, safe recursion schemes are equivalent to higher-order pushdown automata without the collapse operation [47] and are strictly less expressive than general (unsafe) recursion schemes [48, Theorem 1.1]. Here, the model-checking problem asks whether a concrete infinite tree (the Böhm tree generated by a safe recursion scheme) is accepted by the B-automaton for some bound. This problem happens to be significantly simpler than the universality/boundedness problems described above. The proof goes by reducing the order of the safe recursion scheme similarly as done by Knapik, Niwiński, and Urzyczyn [47] to show decidability of the MSO model-checking problem, at the expense of making the property automaton two-way. We then rely on the fact that two-way alternating B-automata can effectively be converted to equivalent one-way alternating B-automata [49]. Our result is incomparable with the seminal decidability result of Ong [9], since (1) alternating B-automata are strictly more expressive than MSO, however (2) we obtain it under the more restrictive safety assumption. Whether the safety assumption can be dropped while preserving decidability of the model-checking problem against B-automata properties, thus strictly extending Ong’s result to the more general setting of boundedness properties, remains open.

Our second contribution is to define the following generalization of the diagonal problem from

words to trees: given a language of finite trees  $\mathcal{L}$  and a set of letters  $\Sigma$ , decide whether for every  $n \in \mathbb{N}$  there is a tree  $T \in \mathcal{L}$  such that every letter from  $\Sigma$  occurs at least  $n$  times on every branch of  $T$ . This generalization is designed in order to reduce computation of downward closures to the diagonal problem, in the same fashion as for finite words. Our proof strategy is to represent downward-closed sets of trees  $\mathcal{L}\downarrow$  by simple tree regular expressions, which are a subclass of regular expressions for finite trees [36, 50]. By further analysing and simplifying the structure of these expressions, computation of the downward closure can be reduced to finitely many instances of the diagonal problem. Unlike in the case of finite words, we do not know whether for full trios of finite trees there exists a converse reduction from the diagonal problem to the problem of computing downward closures.

Our third contribution is decidability of the diagonal problem for languages of finite trees recognized by safe recursion schemes (and thus computability of downward closures of those languages). The diagonal problem can directly be expressed in a logic called *weak cost monadic second-order logic* (WCMSO) [41], which extends weak MSO with atomic formulas of the form  $|X| < N$  stating that the cardinality of the monadic variable  $X$  is smaller than  $N$ . Since WCMSO can be translated to alternating B-automata [41], the diagonal problem reduces to the model-checking problem of safe recursion schemes against alternating B-automata, which we have shown decidable in the first part. Note that it seems difficult to express the diagonal problem using alternating B-automata directly, and indeed the fact that alternating B-automata can express all WCMSO properties is nontrivial. It is worth stressing that this connection between these two unboundedness problems (the diagonal problem and model-checking of B-automata) is new and has not been observed before.

This paper is based on a conference paper [51], showing the same results; we add here missing proofs and some examples.

**Outline.** In Section 2, we define recursion schemes and B-automata. In Section 3, we present our first result, namely decidability of model checking of safe recursion schemes against B-automata. In Section 4, we introduce the diagonal problem, and we show how it can be used to compute downward closures. In Section 5, we solve the diagonal problem for safe recursion schemes. We conclude in Section 6 with some open problems.

## 2. Preliminaries

**Recursion schemes.** A *ranked alphabet* is a (usually finite) set  $\mathbb{A}$  of *letters*, together with a function  $\text{rank} : \mathbb{A} \rightarrow \mathbb{N}$ , assigning a *rank* to every letter. When we define trees below, we require that a node labeled by a letter  $a$  has exactly  $\text{rank}(a)$  children. In the sequel, we usually assume some fixed finite ranked alphabet  $\mathbb{A}$ .

The set of (*simple*) *types* is constructed from a unique ground type  $\mathbf{o}$  using a binary operation  $\rightarrow$ ; namely  $\mathbf{o}$  is a type, and if  $\alpha$  and  $\beta$  are types, so is  $\alpha \rightarrow \beta$ . By convention,  $\rightarrow$  associates to the right, that is,  $\alpha \rightarrow \beta \rightarrow \gamma$  is understood as  $\alpha \rightarrow (\beta \rightarrow \gamma)$ . A type  $\mathbf{o} \rightarrow \dots \rightarrow \mathbf{o}$  with  $k$  occurrences of  $\rightarrow$  is also written as  $\mathbf{o}^k \rightarrow \mathbf{o}$ . The *order* of a type  $\alpha$ , denoted  $\text{ord}(\alpha)$  is defined by induction:  $\text{ord}(\mathbf{o}) = 0$  and  $\text{ord}(\alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \mathbf{o}) = \max_i(\text{ord}(\alpha_i)) + 1$  for  $k \geq 1$ .

We coinductively define both *lambda-terms* and a two-argument relation “ $M$  is a *lambda-term* of

type  $\alpha$ ” as follows:<sup>1</sup>

- a letter  $a \in \mathbb{A}$  is a lambda-term of type  $\mathbf{o}^{\text{rank}(a)} \rightarrow \mathbf{o}$ ;
- for every type  $\alpha$  there is a countable set  $\{x, y, \dots\}$  of *variables of type  $\alpha$*  which can be used as lambda-terms of type  $\alpha$ ;
- if  $M$  is a lambda-term of type  $\beta$  and  $x$  a variable of type  $\alpha$ , then  $\lambda x.M$  is a lambda-term of type  $\alpha \rightarrow \beta$ ; this construction is called a *lambda-binder*;
- if  $M$  is a lambda-term of type  $\alpha \rightarrow \beta$ , and  $N$  is a lambda-term of type  $\alpha$ , then  $M N$  is a lambda-term of type  $\beta$ , called an *application*.

Note that this definition is coinductive, meaning that lambda-terms may be infinite. As usual, we identify lambda-terms up to *alpha-conversion* (i.e., renaming of bound variables). Notice that, according to our definition, every lambda-term (and in particular every variable) has a particular type associated with it. We use here the standard notions of *free variable*, *subterm*, (capture-avoiding) *substitution*, and *beta-reduction*. A *closed* lambda-term does not have free variables. For a lambda-term  $M$  of type  $\alpha$ , the *order* of  $M$ , denoted  $\text{ord}(M)$ , is defined as  $\text{ord}(\alpha)$ . A lambda-term  $M$  is a *first-order lambda-term* if every subterm of  $M$  (including  $M$  itself) has order at most 1 and every free variable of  $M$  has order 0. An *applicative term* is a lambda-term not containing lambda-binders (it contains only letters, applications, and variables).

A lambda-term  $M$  is *superficially safe* if all free variables  $x$  thereof satisfy  $\text{ord}(x) \geq \text{ord}(M)$ . A lambda-term  $M$  is *safe* if for every subterm thereof of the form  $K L$  (i.e., an application), the subterm  $L$  is superficially safe.<sup>2</sup> For example, if  $a, x, x'$  are of type  $\mathbf{o}$  and  $y, y'$  are of type  $\mathbf{o} \rightarrow \mathbf{o}$ , then the lambda-term  $(\lambda y.a) (\lambda x.y' a)$  is safe, but the lambda-term  $(\lambda y.a) (\lambda x.x')$  is not safe:  $x'$  is an order-0 free variable in the order-1 subterm  $(\lambda x.x')$  located on the argument position of an application. Intuitively, safety is a syntactic restriction that guarantees that there is no need to rename bound variables when performing substitution, since variable capture is guaranteed not to happen for safe lambda-terms. This simplifies the analysis of lambda-terms, and allows constructions by induction on the order, as done in Knapik et al. [47]. Safe lambda-terms are semantically less expressive than their unrestricted counterpart.

A (*higher-order, deterministic*) *recursion scheme* over the alphabet  $\mathbb{A}$  is a tuple  $\mathcal{G} = \langle \mathbb{A}, \mathcal{N}, X_0, \mathcal{R} \rangle$ , where  $\mathcal{N}$  is a finite set of typed *nonterminals*,  $X_0 \in \mathcal{N}$  is the *initial nonterminal*, and  $\mathcal{R}$  is a function assigning to every nonterminal  $X \in \mathcal{N}$  of type  $\alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \mathbf{o}$  a finite lambda-term of the form  $\lambda x_1. \dots \lambda x_k.K$ , of the same type  $\alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \mathbf{o}$ , in which  $K$  is an applicative term with free variables in  $\mathcal{N} \uplus \{x_1, \dots, x_k\}$ . We refer to  $\mathcal{R}(X)$  as the *rule* for  $X$ . The *order* of a recursion scheme  $\text{ord}(\mathcal{G})$  is the maximum order of its nonterminals.

The lambda-term *represented* by a recursion scheme  $\mathcal{G}$  as above, denoted  $\Lambda(\mathcal{G})$ , is the limit of applying recursively the following operation to  $X_0$ : take an occurrence of some nonterminal  $X$ , and replace it with  $\mathcal{R}(X)$  (the nonterminals should be chosen in a fair way, so that every nonterminal is eventually replaced). Thus,  $\Lambda(\mathcal{G})$  is a (usually infinite) regular lambda-term obtained by unfolding

<sup>1</sup>Cf. the works [52, 53] for analogous definitions in the literature on infinite lambda calculus. Note that we use letters (constants) from a ranked alphabet, which is a minor modification that suits our needs.

<sup>2</sup>Some definitions of safe lambda-terms add the following requirement: if  $K L$  is a subterm of  $M$ , and  $K$  is not an application, then also  $K$  is required to be superficially safe [2, 54]. This does not change anything when it comes to safety of  $\Lambda(\mathcal{G})$  for a recursion scheme  $\mathcal{G}$ : if  $K L$  is a subterm of  $\Lambda(\mathcal{G})$ , and  $K$  is not an application, then  $K$  is either closed or a variable, so it is always superficially safe.

the nonterminals of  $\mathcal{G}$  according to their definition. We remark that when substituting  $\mathcal{R}(X)$  for a nonterminal  $X$  there is no need for any renaming of variables (capture-avoiding substitution), since  $\mathcal{R}(X)$  does not contain free variables other than nonterminals. We only consider recursion schemes for which  $\Lambda(\mathcal{G})$  is well-defined (e.g. by requiring that  $\mathcal{R}(X)$  is not a single nonterminal). A recursion scheme  $\mathcal{G}$  is *safe* if  $\Lambda(\mathcal{G})$  is safe.

A *tree* is a closed applicative term of type  $\circ$ . Note that such a term is coinductively of the form  $a M_1 \cdots M_r$ , where  $a \in \mathbb{A}$  is of rank  $r$ , and where  $M_1, \dots, M_r$  are again trees. Thus, a tree defined this way can be identified with a tree understood in the traditional sense:  $a$  is the label of its root, and  $M_1, \dots, M_r$  are subtrees rooted at the  $r$  children of the root, from left to right. For trees we employ the usual notions of *node*, *root*, *leaf*, *child*, *parent*, *branch*, and *subtree*. A tree is *regular* if it has finitely many distinct subtrees.

The *Böhm tree* of a closed lambda-term  $M$  of type  $\circ$ , denoted  $\text{BT}(M)$ , is the tree defined coinductively as follows: if there is a sequence of beta-reductions from  $M$  to a lambda-term of the form  $a M_1 \dots M_r$  (where  $a \in \mathbb{A}$  is a letter), then  $\text{BT}(M) = a (\text{BT}(M_1)) \dots (\text{BT}(M_r))$ ; otherwise  $\text{BT}(M) = \perp$ , where  $\perp \in \mathbb{A}$  is a distinguished letter of rank 0. It is a classical result that  $\text{BT}(M)$  exists, and is uniquely defined [52, 53]. The *tree generated* by a recursion scheme  $\mathcal{G}$ , denoted  $\text{BT}(\mathcal{G})$ , is  $\text{BT}(\Lambda(\mathcal{G}))$ .

We say that a closed lambda-term  $N$  of type  $\circ$  is *normalizing* if  $\text{BT}(N)$  does not contain the special letter  $\perp$ ; a recursion scheme  $\mathcal{G}$  is *normalizing* if  $\Lambda(\mathcal{G})$  is normalizing. This notion is analogous to productivity in grammars: in a normalizing recursion scheme / lambda-term the reduction process always terminates producing a new node. It is possible to transform every recursion scheme  $\mathcal{G}$  into a normalizing recursion scheme  $\mathcal{G}'$  generating the same tree as  $\mathcal{G}$ , up to renaming  $\perp$  into some non-special letter  $\perp'$  (cf. [55, Section 5]). Moreover, the construction preserves safety and the order.

**Example 2.1.** Consider the ranked alphabet  $\mathbb{A}$  containing two letters  $a, nd$  of rank 2, two letters  $b_1, b_2$  of rank 1, and two letters  $\perp, c$  of rank 0. Let  $\mathcal{G}$  be the recursion scheme consisting of an initial nonterminal  $S$  of order-0 type  $\circ$  and an additional nonterminal  $A$  of order-2 type  $(\circ \rightarrow \circ) \rightarrow (\circ \rightarrow \circ) \rightarrow \circ \rightarrow \circ \rightarrow \circ$ , together with the following two rules:

$$\begin{aligned}\mathcal{R}(S) &= A \ b_1 \ b_2 \ c \ c, \\ \mathcal{R}(A) &= \lambda f. \lambda g. \lambda x. \lambda y. nd \ (a \times y) \ (A \ f \ g \ (f \ x) \ (g \ y)).\end{aligned}$$

Then,  $\text{BT}(\mathcal{G})$  is the infinite non-regular tree

$$nd \ (a \ c \ c) \ (nd \ (a \ (b_1 \ c) \ (b_2 \ c)) \ (nd \ (a \ (b_1 \ (b_1 \ c)) \ (b_2 \ (b_2 \ c))) \ \dots),$$

depicted in Figure 1.

**Recursion schemes as recognizers of languages of finite trees.** The standard semantics of a recursion scheme  $\mathcal{G} = \langle \mathbb{A}, \mathcal{N}, X_0, \mathcal{R} \rangle$  is the single infinite tree  $\text{BT}(\mathcal{G})$  generated by the scheme. An alternative view is to consider a recursion scheme as a recognizer of a language of finite trees  $\mathcal{L}(\mathcal{G})$ . This alternative view is relevant when discussing downward closures of languages of finite trees. We

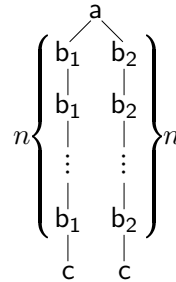


Figure 2. A tree in  $\mathcal{L}(\mathcal{G})$  (Example 2.2)

$$\mathcal{L}(T) = \{U \mid T \rightarrow_{\text{nd}}^* U, \text{ with } U \text{ finite and not containing “nd” or “}\perp\text{”}\}.$$

$$a \underbrace{(b_1 (b_1 (\dots (b_1 c) \dots)))}_n \underbrace{(b_2 (b_2 (\dots (b_2 c) \dots)))}_n \quad \text{for } n \in \mathbb{N},$$

**Alternating B-automata.** We introduce the model of automata used in this paper, namely *alternating one-way/two-way B-automata* over trees (over a ranked alphabet). We consider counters which can be *incremented*  $\mathbf{i}$ , *reset*  $\mathbf{r}$ , or left *unchanged*  $\varepsilon$ . Let  $\Gamma$  be a finite set of *counters* and let  $\mathbb{C} = \{\mathbf{i}, \mathbf{r}, \varepsilon\}$  be the *alphabet of counter actions*. Each counter starts with value zero, and the *value of a sequence* of actions is the supremum of the values achieved during this sequence. For instance  $\mathbf{i}\mathbf{i}\mathbf{r}\varepsilon\mathbf{i}\varepsilon$  has value 2,  $(\mathbf{i}\mathbf{r})^\omega$  has value 1, and  $\mathbf{i}\mathbf{r}\mathbf{i}^2\mathbf{r}\mathbf{i}^3\mathbf{r}\cdots$  has value  $\infty$ . For an infinite sequence of counter actions  $w \in \mathbb{C}^\omega$ , let  $\text{val}(w) \in \mathbb{N} \cup \{\infty\}$  be its *value*. In case of several counters,  $w = c_1c_2\cdots \in (\mathbb{C}^\Gamma)^\omega$ , we take the counter with the maximal value:  $\text{val}(w) = \max_{c \in \Gamma} \text{val}(w(c))$ , where  $w(c) = c_1(c)c_2(c)\cdots$ .

$$\langle \mathbb{A}, Q, q_0, pr, \Gamma, \delta \rangle$$

$$\delta : Q \times \mathbb{A} \rightarrow \mathcal{B}^+(\{\uparrow, \circlearrowleft, \downarrow_1, \downarrow_2, \dots\} \times \mathbb{C}^\Gamma \times Q)$$

mapping a state and a letter  $a$  to a (finite) positive Boolean combination of triples of the form  $(d, c, q)$ ; it is assumed that if  $d = \downarrow_i$  then  $i \leq \text{rank}(a)$ . Such a triple encodes the instruction to send the automaton in the direction  $d$  while performing the action  $c$ , and changing the state to  $q$ . The direction  $\downarrow_i$  denotes moving to the  $i$ -th child,  $\uparrow$  moving to the parent, and  $\circlearrowright$  staying in place. We assume that  $\delta(q, a)$  is written in disjunctive normal form for all  $q$  and  $a$ .

The acceptance of an infinite input tree  $T$  by an alternating B-automaton  $\mathcal{A}$  is defined in terms of a game  $(\mathcal{A}, T)$  between two players, called Eve and Adam. Eve is in charge of disjunctive choices and tries to minimize counter values while satisfying the parity condition. Adam, on the other hand, is in charge of conjunctive choices and tries to either maximize counter values, or to sabotage the parity condition. Since the transition function is given in disjunctive normal form, each turn of the game consists of Eve choosing a disjunct and Adam selecting a single triple  $(d, c, q)$  thereof. In order to deal with the situation that the automaton wants to go up from the root of the tree, we forbid Eve to choose a disjunct containing a triple with direction  $\uparrow$  when the play is in the root. Simultaneously, we assume that  $\delta(q, a)$  for all  $q$  and  $a$  contains a disjunct in which no triple uses the direction  $\uparrow$ , so that from every position there is some move. A *play* of  $\mathcal{A}$  on a tree  $T$  is a sequence  $q_0, (d_1, c_1, q_1), (d_2, c_2, q_2), \dots$  compatible with  $T$  and  $\delta$ :  $q_0$  is the initial state, and for all  $i \in \mathbb{N}$ ,  $(d_{i+1}, c_{i+1}, q_{i+1})$  appears in  $\delta(q_i, T(x_i))$ , where  $x_i$  is the node of  $T$  after following the directions  $d_1, d_2, \dots, d_i$  starting from the root. The *value*  $\text{val}(\pi)$  of such a play  $\pi$  is the value  $\text{val}(c_1 c_2 \dots)$  as defined above if the largest number appearing infinitely often among the priorities  $\text{pr}(q_0), \text{pr}(q_1), \dots$  is even; otherwise,  $\text{val}(\pi) = \infty$ . We say that the play  $\pi$  is *n-winning* (for Eve) if  $\text{val}(\pi) \leq n$ .

A *strategy* for one of the players in the game  $(\mathcal{A}, T)$  is a function that returns the next choice given the history of the play. Note that choosing a strategy for Eve and a strategy for Adam fixes a play in  $(\mathcal{A}, T)$ . We say that a play  $\pi$  is *compatible* with a strategy  $\sigma$  if there is some strategy  $\sigma'$  for the other player such that  $\sigma$  and  $\sigma'$  together yield the play  $\pi$ . A strategy for Eve is *n-winning* if every play compatible with it is *n-winning*. We say that Eve *n-wins* the game if there is some *n-winning* strategy for Eve. The B-automaton *n-accepts* a tree  $T$  if Eve *n-wins* the game  $(\mathcal{A}, T)$ ; it *accepts*  $T$  if it *n-accepts*  $T$  for some  $n \in \mathbb{N}$ . The language *recognized* by  $\mathcal{A}$  is the set of all trees accepted by  $\mathcal{A}$ .

**Example 2.3.** Let  $\mathbb{A}$  be the ranked alphabet containing a letter  $a$  of rank 2 and a letter  $b$  of rank 1. Consider a B-automaton over  $\mathbb{A}$  with one counter and three states  $q_0, q_1, q_2$ , all of priority 0; the state  $q_0$  is initial, and the transitions are

$$\begin{aligned} \delta(q_0, a) &= (\downarrow_1, \varepsilon, q_0) \wedge (\downarrow_2, \varepsilon, q_0), \\ \delta(q_0, b) &= ((\downarrow_1, \varepsilon, q_0) \wedge (\uparrow, i, q_1)) \vee ((\downarrow_1, \varepsilon, q_0) \wedge (\downarrow_1, i, q_2)), \\ \delta(q_1, a) &= (\uparrow, i, q_1) \vee (\circlearrowright, i, q_1), & \delta(q_1, b) &= (\circlearrowright, \varepsilon, q_1), \\ \delta(q_2, a) &= (\downarrow_1, i, q_2) \vee (\downarrow_2, i, q_2), & \delta(q_2, b) &= (\circlearrowright, \varepsilon, q_2). \end{aligned}$$

Here Adam chooses a  $b$ -labeled node  $u$  (using state  $q_0$ ), and then Eve selects a path to some  $b$ -labeled ancestor (using state  $q_1$ ) or descendant (using state  $q_2$ ) of  $u$ ; the counter computes the distance between these two nodes. In consequence, a tree is accepted if there is a bound  $n \in \mathbb{N}$  such that every  $b$ -labeled node has a  $b$ -labeled ancestor or descendant in distance at most  $n$ .

If no  $\delta(q, a)$  uses the direction  $\uparrow$ , then we call  $\mathcal{A}$  *one-way*. Blumensath, Colcombet, Kuperberg, Parys, and Vanden Boom [49, Theorem 6] show that every B-automaton can be made one-way:



**Theorem 2.4.** Given an alternating two-way B-automaton, one can compute an alternating one-way B-automaton that recognizes the same language.

**Proof:**

This essentially follows from the result of Blumensath et al. [49, Theorem 6] modulo some cosmetic changes. Namely, due to some differences in definitions, our Theorem 2.4 is weaker in two aspects and stronger in one aspect than the result of Blumensath et al. [49, Theorem 6]. We elaborate on these differences here.

First, Blumensath et al. [49] do not say that a one-way B-automaton  $\mathcal{A}$  and a two-way B-automaton  $\mathcal{B}$  recognize the same languages, but rather that the cost functions defined by these B-automata are equal (modulo domination equivalence). The latter means that there exists a non-decreasing function  $\alpha: \mathbb{N} \rightarrow \mathbb{N}$  such that if one of the B-automata ( $\mathcal{A}$  or  $\mathcal{B}$ )  $n$ -accepts some tree  $T$ , then the other B-automaton  $\alpha(n)$ -accepts this tree  $T$ . Clearly this is a stronger notion; it implies that the sets of accepted trees are equal.

Second, the definition of one-way B-automata given by Blumensath et al. [49] forbids the usage of the direction  $\circlearrowleft$  (along with  $\uparrow$ ), while we allow to use  $\circlearrowleft$  (only  $\uparrow$  is forbidden). A translation to one-way B-automata becomes only easier if their definition is less restrictive. We remark, however, that we actually need to allow the usage of  $\circlearrowleft$  in order to correctly handle letters of rank 0—we do not want a one-way B-automaton to get stuck in a node without children.

Third, the B-automata of Blumensath et al. [49] work over binary trees, that is, all letters of the alphabet are assumed to be of rank 2, while we allow letters of arbitrary ranks. It is not difficult to believe that the assumption about a binary alphabet is just a technical simplification, and that all the proofs of Blumensath et al. [49] can be repeated for an arbitrary alphabet. Alternatively, it is possible to encode a tree over an arbitrary ranked alphabet into a binary tree, using the first-child next-sibling representation (with some dummy infinite binary tree encoding “no more children”). Such an encoding can easily be incorporated into a B-automaton. Thus, in order to convert a two-way B-automaton  $\mathcal{A}$  into a one-way B-automaton  $\mathcal{B}$ , we can first convert it into a two-way B-automaton  $\mathcal{A}_2$  over a binary alphabet (reading the first-child next-sibling representation of a tree), then convert  $\mathcal{A}_2$  into a one-way B-automaton  $\mathcal{B}_2$  (using the results of Blumensath et al. [49, Theorem 6]), and then convert  $\mathcal{B}_2$  into  $\mathcal{B}$  reading the actual tree instead of its first-child next-sibling representation.  $\square$

**Example 2.5.** In general, the proofs of Blumensath et al. [49] underlying Theorem 2.4 are nontrivial. Nevertheless, in the concrete case of the B-automaton  $\mathcal{A}$  from Example 2.3 it is not difficult to directly construct a one-way B-automaton  $\mathcal{B}$  recognising the same language. The trick is that, instead of going up to a close b-labeled ancestor, already in the ancestor we decide that it will serve as a close b-labeled ancestor for some node. Thus the transitions become

$$\begin{aligned}\delta(q_0, a) &= (\downarrow_1, \varepsilon, q_0) \wedge (\downarrow_2, \varepsilon, q_0), & \delta(q_0, b) &= (\downarrow_1, i, q_1), \\ \delta(q_1, a) &= ((\downarrow_1, i, q_1) \wedge (\downarrow_2, r, q_0)) \vee ((\downarrow_1, r, q_0) \wedge (\downarrow_2, i, q_1)) \vee ((\downarrow_1, i, q_1) \wedge (\downarrow_2, i, q_1)), \\ \delta(q_1, b) &= (\downarrow_1, r, q_0) \vee (\downarrow_1, i, q_1).\end{aligned}$$

As a special case of a result by Colcombet and Göller [56] we obtain the following fact:

**Fact 2.6.** One can decide whether a given B-automaton  $\mathcal{A}$  accepts a given regular tree  $T$ .

**Proof:**

First, thanks to Theorem 2.4, we can assume that  $\mathcal{A}$  is one-way. Next, recall that acceptance of  $T$  is defined in terms of a game  $(\mathcal{A}, T)$ . When  $\mathcal{A}$  is one-way and  $T$  is regular, this game has actually a finite arena. Indeed, for the future of a play, it does not matter what is the current node of  $T$ , it only matters which subtree starts in the current node—and in  $T$  we have finitely many different subtrees. It is not difficult to decide whether such a finite-arena game is  $n$ -won by Eve for some  $n \in \mathbb{N}$ . Nevertheless, instead of showing this directly, we notice that games obtained this way are a special case of games considered by Colcombet and Göller [56], for which they prove decidability.  $\square$

### 3. Model-checking safe recursion schemes against alternating B-automata

In this section we prove the first main theorem of our paper, that is, decidability of the *model-checking problem* of safe recursion schemes against properties described by B-automata:

**Theorem 3.1.** Given an alternating B-automaton  $\mathcal{A}$  and a safe recursion scheme  $\mathcal{G}$ , one can decide whether  $\mathcal{A}$  accepts the tree generated by  $\mathcal{G}$ .

It is worth noticing that this theorem generalizes the result of Knapik et al. [47] on safe recursion schemes from regular (MSO) properties to the more general quantitative realm of properties described by B-automata. On the other hand, our result is incomparable with the celebrated theorem of Ong [9] showing decidability of model checking regular properties of possibly unsafe recursion schemes. Whether model checking of possibly unsafe recursion schemes against properties described by B-automata is decidable remains an open problem.

By Theorem 2.4, every B-automaton can effectively be transformed into an equivalent one-way B-automaton, so it is enough to prove Theorem 3.1 for a one-way B-automaton  $\mathcal{A}$ . The proof of Theorem 3.1 is based on the following lemma, where we use in an essential way the assumption that the recursion scheme is safe:

**Lemma 3.2.** For every safe recursion scheme  $\mathcal{G}$  of order  $m$  and for every alternating one-way B-automaton  $\mathcal{A}$ , one can effectively construct a safe recursion scheme  $\mathcal{G}^\bullet$  of order  $m - 1$  and an alternating two-way B-automaton  $\mathcal{A}'$  such that

$$\mathcal{A} \text{ accepts } \text{BT}(\mathcal{G}) \quad \text{if and only if} \quad \mathcal{A}' \text{ accepts } \text{BT}(\mathcal{G}^\bullet).$$

Notice that the above lemma allows us to decrease the order of a recursion scheme, at the cost of transforming a one-way B-automaton  $\mathcal{A}$  into a two-way B-automaton  $\mathcal{A}'$ .

Before proving Lemma 3.2, let us see how Theorem 3.1 follows from it: Using Lemma 3.2 we can reduce the order of the considered safe recursion scheme by one. We obtain a two-way B-automaton, which we convert back to a one-way B-automaton using Theorem 2.4. It is then sufficient to repeat this process, until we end up with a recursion scheme of order 0. A recursion scheme of order 0 generates a regular tree and, by Fact 2.6, we can decide whether the resulting B-automaton accepts this tree, answering our original question.



Figure 3. The lambda-tree  $T = M^\bullet$  from Example 3.3 (left) and its  $(\mathcal{X}, 2)$ -derived tree  $\llbracket T \rrbracket_{\mathcal{X}, 2}$  (right)

**Lambda-trees.** We now come to the proof of Lemma 3.2. The construction of  $\mathcal{G}^\bullet$  from  $\mathcal{G}$  follows an analogous result for MSO [47, 57], which we generalize to B-automata. We call the construction of  $\mathcal{G}^\bullet$  from  $\mathcal{G}$  *reification*. This is the central idea in Knapik et al. [47, 57], which first proved decidability of MSO model checking of safe recursion schemes. We formally present it in Appendix A.2; here, we illustrate it with some examples.

For a *finite* set  $\mathcal{X}$  of variables of type  $\mathfrak{o}$ , we define a new ranked alphabet  $\mathbb{A}_{\mathcal{X}}$  that contains (1) a letter  $\bar{a}$  of rank 0 for every letter  $a \in \mathbb{A}$ ; (2) a letter  $\bar{x}$  of rank 0 for every variable  $x \in \mathcal{X}$ ; (3) a letter  $\bar{\lambda}x$  of rank 1 for every variable  $x \in \mathcal{X}$ ; (4) a letter  $@$  of rank 2. We remark that  $\mathbb{A}_{\mathcal{X}}$  is a usual finite ranked alphabet. A *lambda-tree* is a tree over the alphabet  $\mathbb{A}_{\mathcal{X}}$ . Reification takes a lambda-term  $M$  and produces a new lambda-term  $M^\bullet$ , in which the maximal order of subterms is strictly smaller. Moreover, when  $M$  is first-order,  $M^\bullet$  is a lambda-tree (i.e., a closed lambda-term of order 0 over the alphabet specified above). Intuitively, order-zero lambda-binders  $\lambda x.K$  (with  $x$  of type  $\mathfrak{o}$ ) and applications  $K L$  with an order-zero argument  $L$  are reified into the syntax: The lambda-binder  $\lambda x.K$  becomes a *letter*  $\bar{\lambda}x$  applied to the recursively reified  $K^\bullet$  and the application  $K L$  becomes also a *letter*  $@$  applied to the recursively reified  $K^\bullet$  and  $L^\bullet$ . This is demonstrated in the next two examples.

**Example 3.3.** Consider the first-order lambda-term (of type  $\mathfrak{o}$ )

$$M = (\lambda x. \lambda y. a \times y) c_1 c_2.$$

In this case  $\mathbb{A}$  contains a letter  $a$  of rank 2 and two letters  $c_1, c_2$  of rank 0, and  $\mathcal{X} = \{x, y\}$ . The new alphabet is thus  $\mathbb{A}_{\mathcal{X}} = \{\bar{a}, \bar{c}_1, \bar{c}_2, \bar{x}, \bar{y}, \bar{\lambda}x, \bar{\lambda}y, @\}$ , where the letter  $@$  is of rank 2, the letters  $\bar{\lambda}x, \bar{\lambda}y$  are of rank 1, and the other letters are of rank 0. The reification of the lambda-term  $M$  is the lambda-tree

$$M^\bullet = @ (@ (\bar{\lambda}x (\bar{\lambda}y (@ (@ \bar{a} \bar{x}) \bar{y}))) \bar{c}_1) \bar{c}_2$$

depicted in Figure 3 (left). Notice that while  $M$  contains actual lambda-binders “ $\lambda x$ ” and “ $\lambda y$ ”, its reification  $M^\bullet$  contains only letters (i.e., no variables and no lambda-binders).

The following example shows how reification is applied to a lambda-term which is not first-order.

**Example 3.4.** Consider the lambda-term

$$M = \lambda f. \lambda x. a \times (f \ x).$$

of type

$$\alpha = (\mathbf{o} \rightarrow \mathbf{o}) \rightarrow \mathbf{o} \rightarrow \mathbf{o}.$$

Applying reification to  $M$  (formally defined in Appendix A) yields the lambda-term

$$M^\bullet = \lambda f^\bullet. \overline{\lambda x} \ (@ \ (@ \ \overline{a} \ \overline{x}) \ (@ \ f^\bullet \ \overline{x}))$$

of the reified type

$$\alpha^\bullet = \mathbf{o} \rightarrow \mathbf{o}.$$

Notice that in this case the reified lambda-term  $M^\bullet$  is not a lambda-tree since the original lambda-term  $M$  is not first-order: Reification is performed only to order-zero lambda-binders, and applications with order-zero arguments; higher-order lambda-binders and applications with higher-order arguments are not reified. In particular, in  $M^\bullet$  we still have a lambda-binder “ $\lambda f^\bullet$ ”, where the variable  $f$  of type  $\beta = \mathbf{o} \rightarrow \mathbf{o}$  has become the variable  $f^\bullet$  of type  $\beta^\bullet = \mathbf{o}$ .

We now define the  $(\mathcal{X}, s)$ -derived tree of a lambda-tree  $T = M^\bullet$ , denoted  $\llbracket T \rrbracket_{\mathcal{X}, s}$ . The construction of the derived tree can be seen as a counterpart of the Böhm tree on the side of reified lambda-terms. Namely, the derived tree is defined in such a way that the derived tree of the reification of  $M$  equals the Böhm tree of  $M$ , that is,  $\llbracket M^\bullet \rrbracket_{\mathcal{X}, s} = \text{BT}(M)$  (c.f. Lemmas 3.7 and A.20). Thus derived trees formally explain how to recover the semantics of  $M$  (its Böhm tree) by only looking at its reification  $M^\bullet$ . This is used later in Lemma 3.8 to state how two-way automata on  $M^\bullet$  can simulate one-way automata on  $M$ .

The definition of the derived tree exploits the fact that a first-order lambda-term  $M$  uses only variables of type  $\mathbf{o}$ . We can thus read the Böhm tree of  $M$  directly, without performing any reduction, just by exploring its reification  $T = M^\bullet$ . Essentially, we walk down through  $T$ , skipping all reified lambda-binders  $\overline{\lambda x}$  and choosing the left branch in all reified applications  $@$ . Whenever we reach some reified variable  $\overline{x}$ , we go up to the corresponding reified lambda-binder  $\overline{\lambda x}$ , then up to the corresponding reified application  $@$ , and then we again start going down in the argument of this application.

Formally, let  $\mathcal{X}$  be a finite set of variables of type  $\mathbf{o}$ , and let  $s \in \mathbb{N}$ . The intended meaning is that  $\mathcal{X}$  contains variables that may potentially appear in the considered lambda-tree  $T$ , and that  $s$  is a bound for the arity of types in the lambda-term represented by  $T$  (the types of all subterms thereof should be of the form  $\mathbf{o}^k \rightarrow \mathbf{o}$  for  $k \leq s$ ). We take<sup>3</sup>  $\text{Dirs}_{\mathcal{X}, s} = \{\Downarrow\} \cup \{\Uparrow_x \mid x \in \mathcal{X}\} \cup \{\Uparrow_i \mid 1 \leq i \leq s\}$ . Intuitively,  $\Downarrow$  means that we go down to the left child of a node labelled by  $@$  or to the unique child of a node labelled by “ $\overline{\lambda x}$ ”,  $\Uparrow_x$  means that we are going up while looking for the value of the variable  $x$ , and  $\Uparrow_i$  means that we are going up while looking for the  $i$ -th argument of an application. For a node  $v$  of  $T$  denote its parent by  $\text{par}(v)$ , and its  $i$ -th child by  $\text{ch}_i(v)$ , where  $1 \leq i \leq k$  and  $k$  is the arity of  $v$ . For  $d \in \text{Dirs}_{\mathcal{X}, s}$ , and for a node  $v$  of  $T$  labeled by  $\zeta \in \mathbb{A}_{\mathcal{X}}$ , we define the  $(\mathcal{X}, s)$ -successor of  $(d, v)$ , when it exists, as

<sup>3</sup>These directions are unrelated with directions in tree automata from Section 2.

1.  $(\Downarrow, ch_1(v))$  if  $d = \Downarrow$  and  $\zeta = \overline{\lambda x}$  (for some  $x$ ) or  $\zeta = @$ ,
2.  $(\Uparrow_x, par(v))$  if  $d = \Downarrow$  and  $\zeta = \overline{x}$  (for some  $x$ ),
3.  $(\Uparrow_x, par(v))$  if  $d = \Uparrow_x$  and  $\zeta \neq \overline{\lambda x}$  (including the case when  $\zeta = \overline{\lambda y}$  for  $y \neq x$ ),
4.  $(\Uparrow_1, par(v))$  if  $d = \Uparrow_x$  and  $\zeta = \overline{\lambda x}$ ,
5.  $(\Uparrow_{i+1}, par(v))$  if  $d = \Uparrow_i$  for  $i < s$  and  $\zeta = \overline{\lambda y}$  (for some  $y$ ),
6.  $(\Uparrow_{i-1}, par(v))$  if  $d = \Uparrow_i$  for  $i > 1$  and  $\zeta = @$ ,
7.  $(\Downarrow, ch_2(v))$  if  $d = \Uparrow_1$  and  $\zeta = @$ .

In particular, the  $(\mathcal{X}, s)$ -successor of  $(d, v)$  is again a pair  $(d', v')$ , where  $d' \in Dirs_{\mathcal{X}, s}$  and  $v'$  is a node of  $T$ . Note that every pair  $(d, v)$  has at most one  $(\mathcal{X}, s)$ -successor, but there may be pairs without any  $(\mathcal{X}, s)$ -successors. In particular, pairs with  $v$  labelled by  $\bar{a}$  do not have a successor.

Rule 1 allows us to go down to the first child in the case of reified lambda-binders  $\overline{\lambda x}$  and reified applications  $@$ . Rule 2 records that we have seen a reified variable  $\overline{x}$  (which is a letter), and thus we need to find its value by going up. Rule 3 climbs the tree upwards as long as we do not see the corresponding reified lambda-binder  $\overline{\lambda x}$ . Rule 4 records that we have seen  $\overline{\lambda x}$  and initializes its level to 1. We now need to find the corresponding application. Rule 5 increments the level and goes up when we encounter a reified lambda-binder  $\overline{\lambda y}$  (which is just a letter), and Rule 6 decrements it for reified applications  $@$ . Finally, when we see a reified application at level 1, we apply Rule 7 which searches for the value of  $\overline{x}$  in the right child.

An  $(\mathcal{X}, s)$ -maximal path from  $(d_1, v_1)$  is a sequence of pairs  $(d_1, v_1), (d_2, v_2), \dots$  in which every  $(d_{i+1}, v_{i+1})$  is the  $(\mathcal{X}, s)$ -successor of  $(d_i, v_i)$ , and which is either infinite or ends in a pair that has no  $(\mathcal{X}, s)$ -successor. For  $d \in Dirs_{\mathcal{X}, s}$ , and for a node  $v$  of  $T$ , we define the  $(\mathcal{X}, s)$ -derived tree from  $(T, d, v)$ , denoted by  $\llbracket T, d, v \rrbracket_{\mathcal{X}, s}$ , by coinduction:

- if the  $(\mathcal{X}, s)$ -maximal path from  $(d, v)$  is finite and ends in  $(\Downarrow, w)$  for a node  $w$  labeled by  $\bar{a}$ , then

$$\llbracket T, d, v \rrbracket_{\mathcal{X}, s} = a \llbracket T, \Uparrow_1, par(w) \rrbracket_{\mathcal{X}, s} \dots \llbracket T, \Uparrow_{rank(a)}, par(w) \rrbracket_{\mathcal{X}, s};$$

- otherwise,  $\llbracket T, d, v \rrbracket_{\mathcal{X}, s} = \perp$ .

The  $(\mathcal{X}, s)$ -derived tree from  $T$  is  $\llbracket T \rrbracket_{\mathcal{X}, s} = \llbracket T, \Downarrow, v_0 \rrbracket_{\mathcal{X}, s}$ , where  $v_0$  is the root of  $T$ . We say that  $T$  is *normalizing* if  $\llbracket T \rrbracket_{\mathcal{X}, s}$  does not contain  $\perp$ .

**Example 3.5.** Let us come back to the lambda-tree  $T$  from Example 3.3 (depicted in Figure 3). Denote the nodes on the leftmost branch of  $T$  by  $v_1, \dots, v_7$  ( $v_1$  is the root, and  $v_7$  is the  $\bar{a}$ -labeled leaf), and the other four nodes with labels  $\bar{c}_2, \bar{c}_1, \bar{y}, \bar{x}$  by  $w_1, w_2, w_3, w_4$ , respectively.

To find the root of the  $(\mathcal{X}, 2)$ -derived tree of  $T$ , we need to follow the  $(\mathcal{X}, 2)$ -maximal path from  $(\Downarrow, v_1)$ . The  $(\mathcal{X}, 2)$ -successor of  $(\Downarrow, v_1)$  is  $(\Downarrow, v_2)$ ; its  $(\mathcal{X}, 2)$ -successor is  $(\Downarrow, v_3)$ , and so on; the path ends in  $(\Downarrow, v_7)$ , which has no  $(\mathcal{X}, 2)$ -successor. Thus the root of the derived tree is labelled with  $a$ . This is shown in the left part of Figure 4.

To find the right child of this root, we need to follow the  $(\mathcal{X}, 2)$ -maximal path from  $(\Uparrow_2, v_6)$ . This path goes through  $(\Uparrow_1, v_5)$ ,  $(\Downarrow, w_3)$ ,  $(\Uparrow_y, v_5)$ ,  $(\Uparrow_y, v_4)$ ,  $(\Uparrow_1, v_3)$ ,  $(\Uparrow_2, v_2)$ ,  $(\Uparrow_1, v_1)$ ,  $(\Downarrow, w_1)$ ; the last pair has no  $(\mathcal{X}, 2)$ -successor, so the right child of the root in the  $(\mathcal{X}, 2)$ -derived tree of  $T$  is labelled by  $c_2$ . Note that the node  $v_5$  is visited twice, with two different directions. This is shown in the right part of Figure 4.

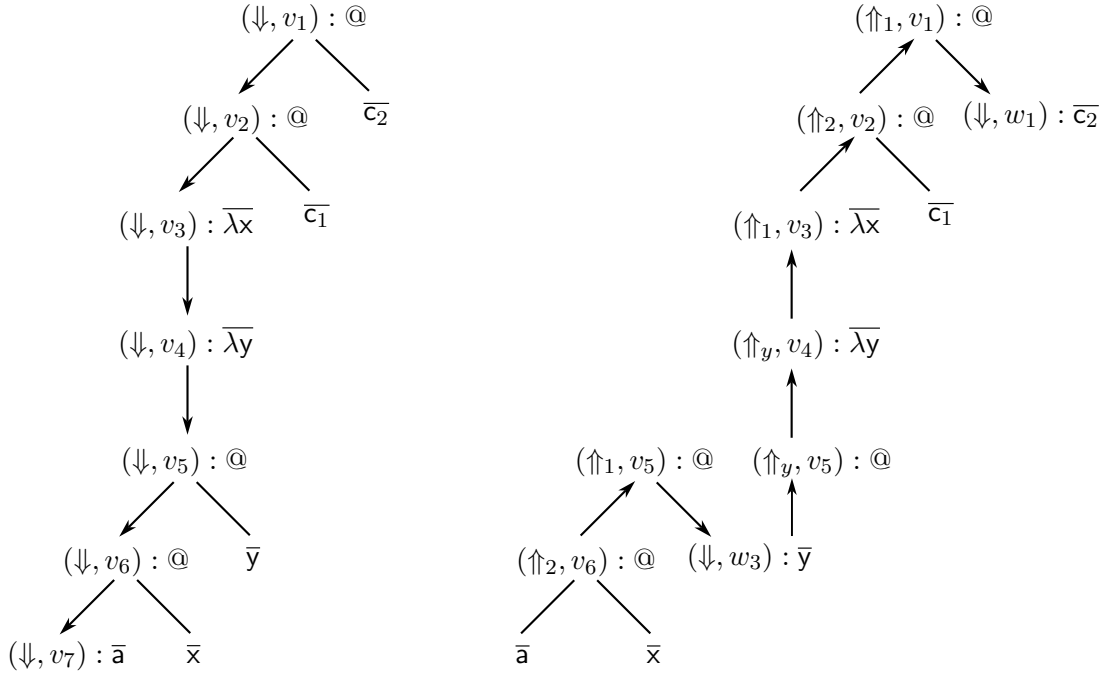


Figure 4. Illustration to Example 3.5. Arrows denote successors in the construction of the derived tree.

The left child can be found in an analogous way, starting from  $(\uparrow_1, v_6)$ . The resulting  $(\mathcal{X}, 2)$ -derived tree  $\llbracket T \rrbracket_{\mathcal{X}, 2}$  is thus a  $c_1 \ c_2$ , depicted in Figure 3 (right).

The next example shows how reification is applied to a whole recursive scheme.

**Example 3.6.** Recall the recursion scheme  $\mathcal{G}$  from Example 2.1, having the following two rules:

$$\begin{aligned} \mathcal{R}(S) &= A \ b_1 \ b_2 \ c \ c, \\ \mathcal{R}(A) &= \lambda f. \lambda g. \lambda x. \lambda y. \text{nd} \ (a \times y) \ (A \ f \ g \ (f \ x) \ (g \ y)). \end{aligned}$$

Applying reification to the recursion scheme  $\mathcal{G}$  one obtains the recursion scheme  $\mathcal{G}^\bullet$  (formally defined in Equality (11) in the appendix) with the following two rules

$$\begin{aligned} \mathcal{R}^\bullet(S^\bullet) &= @ \ ( @ \ (A \ \overline{b_1} \ \overline{b_2}) \ \overline{c}) \ \overline{c}, \\ \mathcal{R}^\bullet(A^\bullet) &= \lambda f. \lambda g. \overline{\lambda x} \ (\overline{\lambda y} \ ( @ \ ( @ \ \overline{\text{nd}} \ ( @ \ ( @ \ \overline{a} \ \overline{x}) \ \overline{y}) \ ( @ \ ( @ \ (A \ f \ g) \ ( @ \ f \ \overline{x}) \ ( @ \ g \ \overline{y}) \))). \end{aligned}$$

The following lemma describes existence of the reified recursion scheme  $\mathcal{G}^\bullet$ , satisfying necessary properties. It crucially relies on the safety assumption.

**Lemma 3.7.** For every normalizing safe recursion scheme  $\mathcal{G}$  of order  $m \geq 1$  one can construct a safe recursion scheme  $\mathcal{G}^\bullet$  of order  $m - 1$ , a finite set of variables  $\mathcal{X}$ , and a number  $s \in \mathbb{N}$  such that

$$\llbracket \text{BT}(\mathcal{G}^\bullet) \rrbracket_{\mathcal{X}, s} = \text{BT}(\mathcal{G}).$$

All the crucial ingredients of the proof of Lemma 3.7 (with some differences in definitions, and with some omitted details) are already contained in papers of Knapik et al. [47, 57]. In the interest of being self-contained, we provide a full proof of Lemma 3.7 in Appendix A. Here, we content ourselves with providing a high-level description of the proof. To construct  $\mathcal{G}^\bullet$ , one needs to replace in  $\mathcal{G}$  every variable  $x$  of type  $\circ$  by  $\bar{x}$ , every lambda-binder concerning such a variable by  $\bar{\lambda}x$ , and every application with an argument of type  $\circ$  by a construct creating a  $@$ -labeled node, as demonstrated in the examples above. Types of subterms change and the order of the recursion scheme decreases by one. While in general computing  $\text{BT}(\Lambda(\mathcal{G}))$  requires one to rename variables in order to perform capture-avoiding substitutions, in the tree generated by the modified recursion scheme we leave the original variable names unchanged. In general (i.e., when the transformation is applied to an arbitrary, possibly unsafe, recursion scheme) this is incorrect due to overlapping variable names and thus possibly unsound substitutions. The assumption that  $\mathcal{G}$  is safe is crucial here: there is no need to rename variables when applying the transformation to a safe recursion scheme. We refer to Appendix A for a detailed proof of Lemma 3.7.

Recall that we are heading towards proving Lemma 3.2. Having Lemma 3.7, it remains to transform a one-way B-automaton  $\mathcal{A}$  operating on the tree generated by  $\mathcal{G}$  into a two-way B-automaton  $\mathcal{A}'$  operating on the lambda-tree generated by  $\mathcal{G}^\bullet$ , as described by the following lemma (as mentioned on page 6, we can assume that  $\mathcal{G}$  is normalizing, which implies that  $\text{BT}(\mathcal{G}^\bullet)$  is normalizing: the tree  $\llbracket \text{BT}(\mathcal{G}^\bullet) \rrbracket_{\mathcal{X},s} = \text{BT}(\mathcal{G})$  does not contain  $\perp$ ):

**Lemma 3.8.** Let  $\mathcal{A}$  be an alternating one-way B-automaton over a finite alphabet  $\mathbb{A}$ , let  $\mathcal{X}$  be a finite set of variables, and let  $s \in \mathbb{N}$ . One can construct an alternating two-way B-automaton  $\mathcal{A}'$  such that for every normalizing lambda-tree  $T$  over  $\mathbb{A}_{\mathcal{X}}$ ,

$$\mathcal{A} \text{ accepts } \llbracket T \rrbracket_{\mathcal{X},s} \quad \text{if and only if} \quad \mathcal{A}' \text{ accepts } T.$$

**Proof:**

The B-automaton  $\mathcal{A}'$  simulates  $\mathcal{A}$  on the lambda-tree. Whenever  $\mathcal{A}$  wants to go down to the  $i$ -th child,  $\mathcal{A}'$  has to follow the  $(\mathcal{X}, s)$ -maximal path from  $(\uparrow_i, v)$  (where  $v$  is the current node). To this end, it has to remember the current pair  $(d, v)$ , and repeatedly find its  $(\mathcal{X}, s)$ -successor. Here  $v$  is always just the current node visited by the B-automaton; the  $d$  component comes from the (finite) set  $\text{Dir}_{\mathcal{X},s}$ , and thus it can be remembered in the state. It is straightforward to encode the definition of an  $(\mathcal{X}, s)$ -successor in transitions of an automaton; we thus omit these tedious details. We do not have to worry about infinite  $(\mathcal{X}, s)$ -maximal paths, because by assumption the  $(\mathcal{X}, s)$ -derived tree does not contain  $\perp$ -labeled nodes.  $\square$

Lemma 3.2 is thus proved by applying Lemma 3.7 and Lemma 3.8. In turn, this proves Theorem 3.1, the main result of this section.

## 4. Downward closures of tree languages

In this section we lay down a method for computation of the downward closure for classes of languages of finite trees closed under linear FTT transductions, which we define in Section 4.2. This

method is analogous to the one of Zetsche [26] for the case of finite words. In Section 4.1 we define the downward closure of languages of finite ranked trees with respect to the embedding well-quasi order and in Section 4.3 we define the simultaneous unboundedness problem for trees and show how computing the downward closure reduces to it. In Section 4.4 we define the diagonal problem for finite trees and show how the previous problem reduces to it. The development of this section is summarised by the following theorem (notions used in its statement are defined in the sequel):

**Theorem 4.1.** Let  $\mathfrak{C}$  be a class of languages of finite trees effectively closed under linear FTT transductions. If the diagonal problem for  $\mathfrak{C}$  is decidable, then downward closures are computable for  $\mathfrak{C}$ .

Let us emphasize that results of this section can be applied to any class of languages of finite trees closed under linear FTT transductions, not just those recognized by safe recursion schemes. In Section 5 we will solve the diagonal problem in the particular case of languages of finite trees recognized by safe recursion schemes, and then we will exploit Theorem 4.1 to show that downward closures are effectively computable for these languages.

#### 4.1. Representations of downward-closed languages

Let  $\sqsubseteq$  be the least relation on finite trees such that (1)  $S \sqsubseteq b T_1 \dots T_r$  if  $S \sqsubseteq T_i$  for some  $i \in \{1, \dots, r\}$ , and (2)  $a S_1 \dots S_r \sqsubseteq b T_1 \dots T_r$  if  $S_i \sqsubseteq T_i$  for all  $i \in \{1, \dots, r\}$ . When  $S \sqsubseteq T$ , we say that  $S$  *homeomorphically embeds into*  $T$ . For a language of finite trees  $\mathcal{L}$ , its *downward closure*, denoted by  $\mathcal{L}_\downarrow$ , is the set of trees  $S$  such that  $S \sqsubseteq T$  for some tree  $T \in \mathcal{L}$ .

**Example 4.2.** The tree  $a c_1 c_2$  embeds into the tree  $b (a (a c_1 c_1) c_2)$ , but it does not embed into the tree  $a (a' c_1 c_2) c_1$ .

**Example 4.3.** The downward closure of the language  $\mathcal{L}(\mathcal{G})$  from Example 2.2 consists of all finite trees of the form

$$\begin{aligned} & \underbrace{b_1 (b_1 (\dots (b_1 c) \dots))}_n && \text{for } n \in \mathbb{N}, \\ & \underbrace{b_2 (b_2 (\dots (b_2 c) \dots))}_m && \text{for } m \in \mathbb{N}, \text{ or} \\ & a \underbrace{(b_1 (b_1 (\dots (b_1 c) \dots)))}_n \underbrace{(b_2 (b_2 (\dots (b_2 c) \dots)))}_m && \text{for } n, m \in \mathbb{N}. \end{aligned}$$

Notice that  $\mathcal{L}(\mathcal{G})$  is a non-regular language of finite trees, while its downward closure above is in fact regular.

**Simple tree regular expressions.** Goubault-Larrecq and Schmitz [36] describe downward-closed sets of trees using *simple tree regular expressions* (*STREs*), which we now introduce.

A *context* is a tree possibly containing one or more occurrences of a special leaf  $\square$ , called a *hole*. Given a context  $C$  and a set of trees  $\mathcal{L}$ , we write  $C[\mathcal{L}]$  for the set of trees obtained from  $C$  by replacing



every occurrence of the hole  $\square$  by some tree from  $\mathcal{L}$ . Different occurrences of  $\square$  are replaced by possibly different trees from  $\mathcal{L}$ . The definition readily extends to a set of contexts  $\mathcal{C}$ , by writing  $\mathcal{C}[\mathcal{L}]$  for  $\bigcup_{C \in \mathcal{C}} C[\mathcal{L}]$ . If  $C$  does not have any  $\square$ , then  $C[\mathcal{L}]$  is just  $\{C\}$ .

An **STRE** is defined according to the following abstract syntax:

$$\begin{aligned} S &::= P + \dots + P, & I &::= C + \dots + C, & S_{\square} &::= \square \mid S. \\ P &::= a^? S \dots S \mid I^*.S, & C &::= a S_{\square} \dots S_{\square}, \end{aligned}$$

These expressions allow empty sums, which are denoted by 0. Subexpressions of the form  $P$ ,  $I$ , and  $C$  are called *pre-products*, *iterators*, and *contexts*, respectively. The word “context” is thus used to describe two different kinds of objects: trees with holes, and expressions of the form  $C$  (denoting sets of trees with holes).

An STRE  $S$  denotes a set of trees  $\llbracket S \rrbracket$  downward-closed for  $\sqsubseteq$ , which is defined recursively as follows:

$$\begin{aligned} \llbracket P_1 + \dots + P_k \rrbracket &= \llbracket P_1 \rrbracket \cup \dots \cup \llbracket P_k \rrbracket, \\ \llbracket a^? S_1 \dots S_r \rrbracket &= \{a T_1 \dots T_r \mid \forall i. T_i \in \llbracket S_i \rrbracket\} \downarrow, \\ \llbracket I^*.S \rrbracket &= \bigcup_{n \in \mathbb{N}} \underbrace{\llbracket I \rrbracket [\dots \llbracket I \rrbracket [\llbracket S \rrbracket]] \dots]}_n, \\ \llbracket C_1 + \dots + C_k \rrbracket &= \llbracket C_1 \rrbracket \cup \dots \cup \llbracket C_k \rrbracket, \\ \llbracket a S_{\square,1} \dots S_{\square,r} \rrbracket &= \{a T_1 \dots T_r \mid \forall i. T_i \in \llbracket S_{\square,i} \rrbracket\} \downarrow, \\ \llbracket \square \rrbracket &= \{\square\}. \end{aligned}$$

Two STREs  $S, T$  are *equivalent* if  $\llbracket S \rrbracket = \llbracket T \rrbracket$ . Since the sets  $\llbracket S_i \rrbracket$  are downward closed, we can see that if all  $\llbracket S_i \rrbracket$  are nonempty, then

$$\llbracket a^? S_1 \dots S_r \rrbracket = \{a T_1 \dots T_r \mid \forall i. T_i \in \llbracket S_i \rrbracket\} \cup \llbracket S_1 \rrbracket \cup \dots \cup \llbracket S_r \rrbracket. \quad (\star)$$

If, however,  $\llbracket S_i \rrbracket = \emptyset$  for some  $i \in \{1, \dots, r\}$ , then  $\llbracket a^? S_1 \dots S_r \rrbracket = \emptyset$ . We have the same property also for  $\llbracket a S_{\square,1} \dots S_{\square,r} \rrbracket$ .

**Example 4.4.** The set  $\llbracket (a \ b \ \square)^*.c^? \rrbracket$  (where  $a$  is of rank 2, and  $b, c$  are of rank 0) consists of trees of the form either  $b$ , or  $c$ , or  $a \ b \ (a \ b \ (\dots (a \ b \ x) \dots))$ , where  $x$  is either  $b$  or  $c$ .

The following lemma is shown by Goubault-Larrecq and Schmitz [36, Proposition 18]:

**Lemma 4.5.** For every downward-closed set of trees  $\mathcal{L}$  there exists an STRE  $S$  such that  $\mathcal{L} = \llbracket S \rrbracket$  (and, vice versa, every STRE  $S$  denotes a downward-closed set of trees  $\llbracket S \rrbracket$ ).  $\square$

**Products.** Among all STREs, Goubault-Larrecq and Schmitz [36] distinguish *products*, which describe *ideals* of trees. Because every downward-closed set of trees is a finite union of ideals, such a set can be described by a finite list of products; this is the idea staying behind Corollary 4.6 below.

In order to define products, Goubault-Larrecq and Schmitz [36] give a way of simplifying STREs by means of a rewrite relation  $\rightarrow_1$ . Intuitively, the idea is to move the operator “+” inside-out as much as possible, and a product is a STRE where no more rewriting can be done. A context  $a S_{\square,1} \dots S_{\square,r}$  is *linear* if at most one  $S_{\square,i}$  is a hole  $\square$ , and it is *full* if  $r \geq 1$  and all the  $S_{\square,j}$ ’s are holes  $\square$ . An iterator  $C_1 + \dots + C_k$  is *linear* (*full*) if all the  $C_i$ ’s are linear contexts (full contexts, respectively). Assuming that “+” is commutative and associative, we define the rewrite relation  $\rightarrow_1$  as follows:

$$P + P' \rightarrow_1 P' \quad \text{if } \llbracket P \rrbracket \subseteq \llbracket P' \rrbracket, \quad (1)$$

$$C + C' \rightarrow_1 C' \quad \text{if } \llbracket C \rrbracket \subseteq \llbracket C' \rrbracket, \quad (2)$$

$$0^*.S \rightarrow_1 S, \quad (3)$$

$$a^? S_1 \dots S_{i-1} 0 S_{i+1} \dots S_r \rightarrow_1 0, \quad (4)$$

$$a S_{\square,1} \dots S_{\square,i-1} 0 S_{\square,i+1} \dots S_{\square,r} \rightarrow_1 0, \quad (5)$$

$$I^*.0 \rightarrow_1 0 \quad \text{if } I \text{ is full}, \quad (6)$$

$$(I + (a S_1 \dots S_r))^*.S \rightarrow_1 I^*. (S + (a^? S_1 \dots S_r)), \quad (7)$$

$$a^? S_1 \dots S_{i-1} (S_i + S'_i) S_{i+1} \dots S_r \rightarrow_1$$

$$a^? S_1 \dots S_{i-1} S_i S_{i+1} \dots S_r + a^? S_1 \dots S_{i-1} S'_i S_{i+1} \dots S_r, \quad (8)$$

$$a S_{\square,1} \dots S_{\square,i-1} (S_{\square,i} + S'_{\square,i}) S_{\square,i+1} \dots S_{\square,r} \rightarrow_1$$

$$a S_{\square,1} \dots S_{\square,i-1} S_{\square,i} S_{\square,i+1} \dots S_{\square,r} + a S_{\square,1} \dots S_{\square,i-1} S'_{\square,i} S_{\square,i+1} \dots S_{\square,r}, \quad (9)$$

$$I^*. (S + S') \rightarrow_1 I^*.S + I^*.S' \quad \text{if } I \text{ is linear}. \quad (10)$$

We allow to apply  $\rightarrow_1$  for subexpressions of an STRE, that is, we write  $S \rightarrow_1 S'$  also when  $S'$  is obtained from  $S$  by replacing some its subexpression  $R$  with  $R'$  such that  $R \rightarrow_1 R'$ .

A *product* is a pre-product  $P$  that is a normal form with respect to  $\rightarrow_1$ , that is, there is no  $P'$  such that  $P \rightarrow_1 P'$ . We know that the rewrite relation  $\rightarrow_1$  preserves the denotation of STRE [36, Fact 19], and that every STRE has a normal form with respect to  $\rightarrow_1$  [36, Lemma 20]. The following corollary is immediate:

**Corollary 4.6.** Every STRE  $S$  is equivalent to a sum of products  $P_1 + \dots + P_k$ . □

**Pure products.** Since the definition of a product is rather indirect, we introduce a stronger notion of *pure products*, which is defined as a syntactic restriction of STREs. Such a definition is more convenient for our purposes. Simultaneously, it still allows us to obtain a decomposition result stated in Lemma 4.7, which is an analogue of Corollary 4.6 for pure products instead of products.

A *pure product* is defined according to the following abstract syntax:

$$\begin{aligned} P &::= a^? P \dots P \mid I^*.P, & C &::= a P_{\square} \dots P_{\square}, \\ I &::= C + \dots + C, & P_{\square} &::= \square \mid P, \end{aligned}$$

where the sum of contexts is nonempty, and where in a context  $C = a P_{\square,1} \dots P_{\square,r}$  it is required that at least one  $P_{\square,i}$  is a hole  $\square$ . The semantics  $\llbracket P \rrbracket$  of pure products is inherited from STRE. Notice, however, that  $\llbracket P \rrbracket$  is always nonempty, so we can use Formula  $(\star)$  to define  $\llbracket a^? P_1 \dots P_r \rrbracket$  and  $\llbracket a P_{\square,1} \dots P_{\square,r} \rrbracket$ .

Formally, a pure product needs not be a product: a pure product is allowed to contain an iterator  $C + C'$  with  $\llbracket C \rrbracket \subseteq \llbracket C' \rrbracket$ , to which Rule (2) can be applied. Nevertheless, by replacing every such sum  $C + C'$  with  $C'$  we can obtain an equivalent pure product that is a product (it is easy to see that no rule other than Rule (2) can be applied to a pure product). Thus, it is justified to say that, morally, the notion of a pure product strengthens the notion of a product.

Based on the results of Goubault-Larrecq and Schmitz [36], in the remaining part of this subsection we deduce the following lemma:

**Lemma 4.7.** Every set of trees  $\mathcal{L}$  downward-closed for  $\sqsubseteq$  can be represented as  $\mathcal{L} = \llbracket P_1 \rrbracket \cup \dots \cup \llbracket P_k \rrbracket$ , in which  $P_1, \dots, P_k$  are pure products.

This decomposition result strengthens the results of Goubault-Larrecq and Schmitz [36] by showing that pure products (instead of just products) suffice in order to represent downward-closed sets of trees.

**From products to pure products.** In Lemma 4.8 we show how to convert an arbitrary product into a pure product. Lemma 4.7 is then a direct consequence of Lemma 4.5, Corollary 4.6, and Lemma 4.8.

**Lemma 4.8.** For every product  $P$  one can create an equivalent pure product  $P'$ .

**Proof:**

The proof is by induction on the size of  $P$ . Before starting the actual proof, let us observe two facts, which we use implicitly below. First, every subterm of  $P$  that is a pre-product is actually a product (i.e., it cannot be rewritten by  $\rightarrow_1$ ). Second, if we replace a product subexpression of  $P$  by some equivalent product, then the resulting STRE is still a product (i.e., it cannot be rewritten by  $\rightarrow_1$ ).

Coming now to the proof, suppose that  $P$  is of the form  $a^? S_1 \dots S_r$ . Then, because  $P$  is a product, that is, because it cannot be rewritten by  $\rightarrow_1$ , we can observe that all the  $S_i$ 's are products. Indeed, if some  $S_i$  was a sum of two or more products (or 0), then  $P$  could be rewritten using Rule (8) (or Rule (4), respectively). By the induction assumption for every product  $S_i$  we can create an equivalent pure product  $S'_i$ ; as  $P'$  we take  $a^? S'_1 \dots S'_r$ .

Next, suppose that  $P$  is of the form  $I^*.S$ . Consider a context  $C = a S_{\square,1} \dots S_{\square,r}$ , being a component of  $I$ . We can observe that all  $S_{\square,i}$  are either  $\square$  or products. Indeed, if some  $S_{\square,i}$  was a sum of two or more products (or 0), then  $C$  could be rewritten using Rule (9) (or Rule (5), respectively). As previously, using the induction assumption we can replace every product  $S_{\square,i}$  that is not a hole  $\square$  by an equivalent pure product  $S'_{\square,i}$ . Applying this to every context  $C$  in  $I$ , we obtain a new iterator  $I_\circ$  in which every STRE subterm is a single product. Likewise, writing  $S = P_1 + \dots + P_k$ , we can replace every product  $P_i$  by an equivalent pure product  $P'_i$ . This way, we obtain a product  $P^\circ = I_\circ^*.(P'_1 + \dots + P'_k)$  equivalent to  $P$ . Observe also that there is at least one context in  $I_\circ$ , and that every context in  $I_\circ$  contains a hole (because Rules (3) and (7) cannot be applied to  $P^\circ$ ), as required in our definition of a pure product. Thus, when  $k = 1$ ,  $P^\circ$  is a pure product, hence it can be taken as  $P'$ . It remains to deal with the situation when  $k \neq 1$ .

One possibility is that  $k = 0$ . Then  $I_\circ$  is not full (otherwise Rule (6) could be applied to  $P^\circ$ ), which means that in  $I_\circ$  there is a context  $C' = a S'_{\square,1} \dots S'_{\square,r}$  such that  $S'_{\square,j} \neq \square$  for some  $j \in \{1, \dots, r\}$ .

Fix one such  $C'$  and  $j$ , and define  $P' := I_o^* \cdot S'_{\square, j}$ . Then  $P'$  is a pure product. Clearly  $\llbracket P^\circ \rrbracket \subseteq \llbracket P' \rrbracket$ , because  $\llbracket 0 \rrbracket \subseteq \llbracket S'_{\square, j} \rrbracket$ . On the other hand,  $\llbracket S'_{\square, i} \rrbracket \neq \emptyset$  for all  $i \in \{1, \dots, r\}$ , because  $S'_{\square, i}$  is either a hole or a pure product, and it can be easily seen (by induction on its structure) that a pure product always denotes a nonempty set; in consequence  $\llbracket S'_{\square, j} \rrbracket \subseteq \llbracket C' \rrbracket[\emptyset] \subseteq \llbracket I_o \rrbracket[\emptyset]$ , so also  $\llbracket P' \rrbracket \subseteq \llbracket P^\circ \rrbracket$ .

Another possibility is that  $k \geq 2$ . Then  $I_o$  is not linear (if  $I_o$  were linear, then Rule (10) could be applied to  $P^\circ$ ), which means that in  $I_o$  there is a context  $C' = a S'_{\square, 1} \dots S'_{\square, r}$  with two or more holes. Fix one such  $C'$ . For simplicity, we show the proof assuming that the first  $\ell$  among  $S'_{\square, i}$  are holes, and the remaining  $r - \ell$  among  $S'_{\square, i}$  are products (i.e., are not holes); the general situation can be handled in the same way, but writing it down would require us to use intricate indices. We define

$$\begin{aligned} R_1 &= a^? P'_1 P'_1 \dots P'_1 S'_{\square, \ell+1} \dots S'_{\square, r} & \text{and} \\ R_i &= a^? R_{i-1} P'_i \dots P'_i S'_{\square, \ell+1} \dots S'_{\square, r} & \text{for } i \in \{2, \dots, k\}, \end{aligned}$$

and we define  $P' := I_o^* \cdot R_k$ . Notice that  $P'$  is a pure product. On the one hand,  $\llbracket P'_i \rrbracket \subseteq \llbracket R_k \rrbracket$  for every  $i \in \{1, \dots, k\}$  (it is important here that there are at least two holes, so  $P'_i$  actually appears in  $R_i$ ), so  $\llbracket P^\circ \rrbracket \subseteq \llbracket P' \rrbracket$ . Let us see the opposite inclusion. First, by definition,  $\llbracket P'_i \rrbracket \subseteq \llbracket I_o^* \cdot (P'_1 + \dots + P'_k) \rrbracket = \llbracket P^\circ \rrbracket$  for all  $i \in \{1, \dots, k\}$ . Second, because  $R_i$  for  $i \in \{2, \dots, k\}$  is obtained by substituting  $P'_i$  and  $R_{i-1}$  for all holes in  $C'$ , we have  $\llbracket R_i \rrbracket \subseteq \llbracket C' \rrbracket[\llbracket P'_i \rrbracket \cup \llbracket R_{i-1} \rrbracket]$ ; likewise  $\llbracket R_1 \rrbracket \subseteq \llbracket C' \rrbracket[\llbracket P'_1 \rrbracket]$ . Then, by induction on  $i \in \{1, \dots, k\}$  we have that  $\llbracket R_i \rrbracket \subseteq \llbracket P^\circ \rrbracket$ : indeed, due to the above observation and the induction hypothesis (if  $i \geq 2$ ) we have that  $\llbracket R_i \rrbracket \subseteq \llbracket C' \rrbracket[\llbracket P^\circ \rrbracket] \subseteq \llbracket I_o \rrbracket[\llbracket P^\circ \rrbracket] \subseteq \llbracket P^\circ \rrbracket$ . In particular,  $\llbracket R_k \rrbracket \subseteq \llbracket P^\circ \rrbracket$ , so also  $\llbracket P' \rrbracket \subseteq \llbracket P^\circ \rrbracket$ .  $\square$

## 4.2. Transductions

A (nondeterministic) *finite tree transducer* (FTT) is a tuple  $\mathcal{A} = (\mathbb{A}_{in}, \mathbb{A}_{out}, S, p^I, \Delta)$ , where  $\mathbb{A}_{in}$ ,  $\mathbb{A}_{out}$  are the input and output alphabets (finite, ranked),  $S$  is a finite set of *control states*,  $p^I \in S$  is an *initial state*, and  $\Delta$  is a finite set of *transition rules* of the form either  $(p, a x_1 \dots x_r) \rightarrow V$  or  $(p, x) \rightarrow V$ , where  $p \in S$  is a control state,  $a \in \mathbb{A}_{in}$  is a letter of rank  $r$ , and  $V$  is a finite tree over the alphabet  $\mathbb{A}_{out} \cup (S \times \{x_1, \dots, x_r\})$  or  $\mathbb{A}_{out} \cup (S \times \{x\})$ , respectively. Here  $x, x_1, \dots, x_r$  are just some special symbols, and the rank of all the pairs from  $S \times \{x_1, \dots, x_r\}$  or  $S \times \{x\}$  is 0. An FTT is *linear* if for each rule of the form  $(p, a x_1 \dots x_r) \rightarrow V$  and for each  $i \in \{1, \dots, r\}$ , in  $V$  there is at most one letter from  $S \times \{x_i\}$ , and moreover for each rule of the form  $(p, x) \rightarrow V$ , in  $V$  there is at most one letter from  $S \times \{x\}$ .

An FTT  $\mathcal{A} = (\mathbb{A}_{in}, \mathbb{A}_{out}, S, p^I, \Delta)$  reading a tree  $T$  over the alphabet  $\mathbb{A}_{in}$  starts in the state  $p^I$  at the root of  $T$ . Then, when  $\mathcal{A}$  is in a state  $p$  at the root of a subtree  $T' = a T_1 \dots T_r$  of  $T$ , it can use a rule of the form  $(p, a x_1, \dots, x_r) \rightarrow V$  from  $\Delta$ ; it produces a tree starting like  $V$ , but leaves of the form  $(q, x_i)$  are replaced by the output of running  $\mathcal{A}$  in the state  $q$  at the root of  $T_i$ . Alternatively,  $\mathcal{A}$  can use a rule of the form  $(p, x) \rightarrow V$ ; then it produces a tree starting like  $V$ , but leaves of the form  $(q, x)$  are replaced by the output of running  $\mathcal{A}$  in the state  $q$  at the same node (i.e., this is an  $\varepsilon$ -transition producing some output). In this way, an FTT  $\mathcal{A}$  defines a relation between finite trees, also denoted  $\mathcal{A}$ ; for a fully formal definition see Comon et al. [58, Section 6.4.2]. For a language  $\mathcal{L}$  we write  $\mathcal{A}(\mathcal{L})$  for the set of trees  $U$  such that  $(T, U) \in \mathcal{A}$  for some  $T \in \mathcal{L}$ . A function that maps  $\mathcal{L}$  to  $\mathcal{A}(\mathcal{L})$  for some linear FTT  $\mathcal{A}$  is called a *linear FTT transduction*.

We now recall two easy facts about linear FTT transductions. The first fact says that taking downward closures is an FTT transduction:

**Fact 4.9.** Given a finite ranked alphabet  $\mathbb{A}$  one create a linear FTT  $\mathcal{A}$  such that for every language  $\mathcal{L}$  of finite trees over  $\mathbb{A}$ , the language  $\mathcal{A}(\mathcal{L})$  equals  $\mathcal{L}\downarrow$ .

**Proof:**

It is enough to take  $\mathcal{A} = (\mathbb{A}, \mathbb{A}, \{p\}, p, \Delta)$  with a single state  $p$ , where  $\Delta$  consists of the following rules for every letter  $a \in \mathbb{A}$  of rank  $r$ :

$$\begin{aligned} (p, a \ x_1 \ \dots \ x_r) &\rightarrow a \ (p, x_1) \ \dots \ (p, x_r), & \text{and} \\ (p, a \ x_1 \ \dots \ x_r) &\rightarrow (p, x_i), & \text{for all } i \in \{1, \dots, r\}. \end{aligned}$$

Such a transducer can convert every tree  $T \in \mathcal{L}$  into every tree  $S$  that homeomorphically embeds into  $T$ .  $\square$

The second fact says that linear FTT transductions can implement intersections with regular languages:

**Fact 4.10.** Given (a finite tree automaton recognizing) a regular language  $\mathcal{R}$  of finite trees over a finite ranked alphabet  $\mathbb{A}$ , one create a linear FTT  $\mathcal{A}$  such that for every language  $\mathcal{L}$  of finite trees over  $\mathbb{A}$ , the language  $\mathcal{A}(\mathcal{L})$  equals  $\mathcal{L} \cap \mathcal{R}$ .

**Proof:**

We are given an automaton that accepts a tree  $T$  if  $T \in \mathcal{R}$  (and rejects it otherwise), and we want to construct a linear FTT  $\mathcal{A}$  that converts a tree  $T$  into itself if  $T \in \mathcal{R}$  (and does not allow to produce any output tree otherwise). Creating  $\mathcal{A}$  out of the automaton is just a matter of changing the syntax: we take to  $\mathcal{A}$  all transition rules of the automaton, enhancing them so that the input tree is produced again on the output.  $\square$

### 4.3. The simultaneous unboundedness problem for trees

We say that a pure product  $P$  is *diversified*, if no letter appears in  $P$  more than once. The *simultaneous unboundedness problem (SUP)* for a class  $\mathfrak{C}$  of finite trees asks, given a diversified pure product  $P$  and a language  $\mathcal{L} \in \mathfrak{C}$  such that  $\mathcal{L} \subseteq \llbracket P \rrbracket$ , whether  $\llbracket P \rrbracket \subseteq \mathcal{L}\downarrow$ .

**Remark 4.11.** This is a generalization of SUP over finite words. In the latter problem, one is given a language of finite words  $\mathcal{L}$  such that  $\mathcal{L} \subseteq a_1^* \dots a_k^*$ , and must check whether  $a_1^* \dots a_k^* \subseteq \mathcal{L}\downarrow$ . A word in  $a_1^* \dots a_k^*$  can be represented as a linear tree by interpreting  $a_1, \dots, a_k$  as unary letters and by appending a new leaf  $e$  at the end. Thus  $a_1^* \dots a_k^*$  can be represented as the language of the diversified pure product  $(a_1 \ \square)^* \cdot (a_2 \ \square)^* \cdot \dots \cdot (a_k \ \square)^* \cdot e$ .

Every pure product  $P$  can be made diversified by adding additional marks to letters appearing in  $P$ . Namely, for each letter  $a$  appearing  $k$  times in  $P$ , we consider “marked” letters  $a_1, \dots, a_k$ , and for each occurrence of  $a$  in  $P$  we substitute a different letter  $a_i$ . To specify a correspondence between the original pure product  $P$  and the resulting diversified pure product  $P'$  we define a  $cl(\cdot)$  operation: when  $X'$  is an object (e.g., a pure product, a context, a tree, etc.) over such an extended alphabet, we write  $cl(X')$  for the object obtained from  $X'$  by removing marks from its labels (i.e., replacing back all  $a_1, \dots, a_k$  by  $a$ ). We also define  $cl(\mathcal{L}') = \{cl(T') \mid T' \in \mathcal{L}'\}$  for a set of trees  $\mathcal{L}'$ . In particular, when  $P'$  is obtained by adding marks to all letters in  $P$ , we have  $cl(P') = P$ . We have the following claim:

**Claim 4.12.**  $\llbracket cl(X') \rrbracket = cl(\llbracket X' \rrbracket)$  whenever  $X$  is an STRE, a pure product, a context, or an iterator over the extended alphabet.

**Proof:**

The claim follows by a straightforward induction on the size of  $X'$ , because the  $cl(\cdot)$  operation commutes with all constructs appearing in the definition of  $\llbracket \cdot \rrbracket$ , namely  $cl(\mathcal{L}'_1 \cup \mathcal{L}'_2) = cl(\mathcal{L}'_1) \cup cl(\mathcal{L}'_2)$ ,  $cl(\mathcal{L}'\downarrow) = (cl(\mathcal{L}'))\downarrow$ , and  $cl(\mathcal{C}'[\mathcal{L}']) = cl(\mathcal{C}')[cl(\mathcal{L}')]$ .  $\square$

Following Zetsche [26], we can reduce computation of the downward closure to SUP:

**Lemma 4.13.** Let  $\mathfrak{C}$  be a class of languages of finite trees closed under linear FTT transductions. One can compute a finite tree automaton recognizing the downward closure of a given language from  $\mathfrak{C}$  if and only if SUP is decidable for  $\mathfrak{C}$ .

**Proof:**

If downward closures are computable, then one can compute a finite tree automaton recognizing  $\mathcal{L}\downarrow$ . Moreover, given a (diversified) pure product  $P$ , one can easily construct a finite tree automaton recognizing  $\llbracket P \rrbracket$ , following the inductive definition of  $\llbracket \cdot \rrbracket$ . Having these two automata, one can check whether  $\llbracket P \rrbracket \subseteq \mathcal{L}\downarrow$ : language inclusion for finite tree automata is decidable [58, Section 1.7].

For the other direction, assume that SUP is decidable for  $\mathfrak{C}$  and let  $\mathcal{L} \in \mathfrak{C}$ . The downward closure  $\mathcal{L}^d := \mathcal{L}\downarrow$  is effectively in  $\mathfrak{C}$  since it can be obtained as a linear FTT transduction of  $\mathcal{L}$  by Fact 4.9. Thus, it is enough to compute a finite tree automaton recognizing the downward-closed language  $\mathcal{L}^d$ . Furthermore, by Lemma 4.7  $\mathcal{L}^d$  equals  $\llbracket P_1 \rrbracket \cup \dots \cup \llbracket P_k \rrbracket$  for some (unknown) pure products  $P_1, \dots, P_k$ , and a finite tree automaton recognizing  $\llbracket P_1 \rrbracket \cup \dots \cup \llbracket P_k \rrbracket$  can be easily computed out of  $P_1, \dots, P_k$ . In consequence, it suffices to guess these pure products and check whether the equality  $\mathcal{L}^d = \llbracket P_1 \rrbracket \cup \dots \cup \llbracket P_k \rrbracket$  indeed holds.

We start by showing how to decide whether  $\mathcal{L}^d \subseteq \llbracket P_1 \rrbracket \cup \dots \cup \llbracket P_k \rrbracket$ . Firstly,  $\mathcal{R} := \llbracket P_1 \rrbracket \cup \dots \cup \llbracket P_k \rrbracket$  is (effectively) a regular language, and consequently its complement  $\mathcal{R}^c$  is also regular. In consequence,  $\mathcal{M} := \mathcal{L}^d \cap \mathcal{R}^c$  is effectively in  $\mathfrak{C}$ , because it can be obtained from  $\mathcal{L}^d$  by intersecting it with  $\mathcal{R}^c$ , which is a linear FTT transduction by Fact 4.10. Secondly, emptiness of any language  $\mathcal{M} \in \mathfrak{C}$  is decidable by reducing to SUP, since it suffices to apply to it the linear FTT  $\mathcal{A}$  that ignores the input and outputs all trees of the form  $a(a \dots (a e) \dots)$  (for some fixed letters  $a$  of rank 1 and  $e$  of rank 0), and to compare the result with the diversified pure product  $P := (a \square)^*.e^?$ . Indeed,

$\mathcal{A}(\mathcal{M}) = \mathcal{A}(\mathcal{M})\downarrow = \llbracket P \rrbracket$  if  $\mathcal{M}$  is nonempty, and  $\mathcal{A}(\mathcal{M}) = \mathcal{A}(\mathcal{M})\downarrow = \emptyset$  if  $\mathcal{M}$  is empty; thus, on the one hand,  $\mathcal{A}(\mathcal{M}) \subseteq \llbracket P \rrbracket$  and, on the other hand,  $\mathcal{M}$  is nonempty if and only if  $\llbracket P \rrbracket \subseteq \mathcal{A}(\mathcal{M})\downarrow$ .

For the other inclusion  $\llbracket P_1 \rrbracket \cup \dots \cup \llbracket P_k \rrbracket \subseteq \mathcal{L}^d$  we can equivalently check whether  $\llbracket P_i \rrbracket \subseteq \mathcal{L}^d$  for all  $i \in \{1, \dots, k\}$ , which implies that it suffices to show decidability of checking the containment  $\llbracket P \rrbracket \subseteq \mathcal{L}^d$  for a single pure product  $P$ . We make  $P$  diversified by adding additional marks to letters appearing in  $P$ . As described before Claim 4.12, we achieve this by unambiguously replacing the  $i$ -th occurrence of letter  $a$  with the new letter  $a_i$ . Let  $P'$  be the resulting diversified pure product. We also create a corresponding linear FTT  $\mathcal{A}$ ; it replaces every label  $a$  in the input tree by an arbitrary letter among the corresponding letters  $a_i$  (for every occurrence of  $a$  we choose a mark  $i$  independently). We obtain  $\llbracket P \rrbracket = cl(\llbracket P' \rrbracket) = \{cl(T') \mid T' \in \llbracket P' \rrbracket\}$  by Claim 4.12, and  $\mathcal{A}(\mathcal{L}^d) = \{T' \mid cl(T') \in \mathcal{L}^d\}$  by definition, which gives us the following equivalence:

$$\llbracket P \rrbracket \subseteq \mathcal{L}^d \iff \llbracket P' \rrbracket \subseteq \mathcal{A}(\mathcal{L}^d).$$

Thus, instead of checking whether  $\llbracket P \rrbracket \subseteq \mathcal{L}^d$ , we can check whether  $\llbracket P' \rrbracket \subseteq \mathcal{A}(\mathcal{L}^d)$ . Finally, we consider a language  $\mathcal{L}' := \mathcal{A}(\mathcal{L}^d) \cap \llbracket P' \rrbracket$ , which can be obtained from  $\mathcal{A}(\mathcal{L}^d)$  by a linear FTT transduction (cf. Fact 4.10), and thus which is effectively in  $\mathfrak{C}$ . Then, on the one hand,  $\mathcal{L}' \subseteq \llbracket P' \rrbracket$  and, on the other hand,  $\llbracket P' \rrbracket \subseteq \mathcal{A}(\mathcal{L}^d)$  if and only if  $\llbracket P' \rrbracket \subseteq \mathcal{L}'$ . Recall that  $\mathcal{L}^d$  and  $\llbracket P' \rrbracket$  are downward closed. It does not matter whether we first remove some parts of a tree and then we add marks to labels, or we first add marks to labels and then we remove some part of a tree, so  $\mathcal{A}(\mathcal{L}^d)$  and  $\mathcal{L}'$  are downward closed as well (and hence  $\mathcal{L}'\downarrow = \mathcal{L}'$ ). It follows that checking whether  $\llbracket P' \rrbracket \subseteq \mathcal{L}'$  is an instance of SUP.  $\square$

**Remark 4.14.** Pure products for trees correspond to expressions of the form  $a_0^? A_1^* a_1^? \dots A_k^* a_k^?$  for words (where  $A_i$  are sets of letters). In SUP for words simpler expressions of the form  $b_1^* \dots b_k^*$  suffice. This is not possible for trees: (1) expressions of the form  $a^? P_1 P_2$  cannot be removed since they are responsible for branching, and (2) reducing the two contexts in  $((a P_1 \square) + (b P_2 \square))^* . P_3$  to a single one would require changing trees of the form  $a T_1 (b T_2 T_3)$  into trees of the form  $c T_1 T_2 T_3$ , which is not a linear FTT transduction.

#### 4.4. The diagonal problem for trees

In SUP for words, instead of checking whether  $a_1^* \dots a_k^* \subseteq \mathcal{L}\downarrow$ , one can equivalently check whether, for each  $n \in \mathbb{N}$ , there is a word in  $\mathcal{L}\downarrow \cap a_1^* \dots a_k^*$  containing at least  $n$  occurrences of every letter  $a_i$ , where  $i \in \{1, \dots, k\}$ . The latter problem (for an arbitrary language  $\mathcal{L}'$  in place of  $\mathcal{L}\downarrow \cap a_1^* \dots a_k^*$ ) is known as the *diagonal problem* for words. In this section, we define an analogous diagonal problem for trees, and we show how to reduce SUP to it.

Given a set of letters  $\Sigma$ , we say that a language of finite trees  $\mathcal{L}$  is  $\Sigma$ -*diagonal* if, for every  $n \in \mathbb{N}$ , there is a tree  $T \in \mathcal{L}$  such that for every letter  $a \in \Sigma$  and every branch  $B$  of  $T$  there are at least  $n$  occurrences  $a$  in  $B$ . The *diagonal problem* for a class  $\mathfrak{C}$  of finite trees asks, given a language  $\mathcal{L} \in \mathfrak{C}$  and a set of letters  $\Sigma$ , whether  $\mathcal{L}$  is  $\Sigma$ -diagonal.

**Versatile trees.** Contrary to the case of words, the presence of sums in our expressions creates some complications in reducing from SUP to the diagonal problem. Namely, suppose that we want to check whether  $\llbracket ((a \square) + (b \square))^* . c \rrbracket \subseteq \mathcal{L} \downarrow$ . This question is not equivalent to checking whether  $\mathcal{L} \downarrow \cap \llbracket ((a \square) + (b \square))^* . c \rrbracket$  contains trees with arbitrarily many  $a$  and  $b$ . Indeed, it is possible that  $\mathcal{L} \downarrow$  contains trees of the form  $a (a (\dots (a (b (b (\dots (b c) \dots))) \dots))$  with arbitrarily many  $a$  and  $b$ , but this does not yet mean that it contains arbitrarily large trees of the form  $a (b (a (b (\dots (a (b c)) \dots)))$ . Denote the latter tree with  $n$  occurrences of  $a$  by  $T_n$ ; the original question is rather equivalent to checking whether  $\mathcal{L} \downarrow \cap \{T_n \mid n \in \mathbb{N}\}$  contains trees with arbitrarily many  $a$  and  $b$ . This is the case, because every tree in  $\llbracket ((a \square) + (b \square))^* . c \rrbracket$  can be embedded in a large enough tree  $T_n$  (e.g.,  $b (b (a c))$  embeds in  $T_3 = a (b (a (b (a (b c))))$ ).

We thus deal with sums by considering trees like  $T_n$ , which we call *versatile trees*. Intuitively, in order to obtain a versatile tree of a pure product  $P$ , for every sum  $I = C_1 + \dots + C_k$  in  $P$  we fix some order of the contexts  $C_1, \dots, C_k$ , and we allow the contexts to be appended only in this order. Formally, the set  $\langle P \rangle$  of versatile trees of a pure product  $P$  is defined by induction on the structure of  $P$ :

$$\begin{aligned} \langle I^* . P \rangle &= \bigcup_{n \in \mathbb{N}} \langle I \rangle [\underbrace{(\langle I \rangle \cup \{\square\}) \dots (\langle I \rangle \cup \{\square\})}_{n} [\langle P \rangle]] \dots], \\ \langle a^? P_1 \dots P_r \rangle &= \langle a P_1 \dots P_r \rangle, \\ \langle C_1 + \dots + C_k \rangle &= \langle C_1 \rangle [\dots [\langle C_k \rangle] \dots], \\ \langle a P_{\square,1} \dots P_{\square,r} \rangle &= \{a T_1 \dots T_r \mid \forall i. T_i \in \langle P_{\square,i} \rangle\}, \\ \langle \square \rangle &= \{\square\}. \end{aligned}$$

For example, if  $I = (a S_1 \square \square) + (b \square S_2)$ , then  $\langle I \rangle = \{a S_1 (b \square S_2) (b \square S_2)\}$ ; in particular, we have  $b (a S_1 \square \square) S_2 \notin \langle I \rangle$ . Notice that the roots of all trees in  $\langle P \rangle$  have the same label; denote this label by  $\text{root}(P)$ .

**From SUP to the diagonal problem.** Assuming that  $P$  is diversified, for a number  $n \in \mathbb{N}$  we say that a tree  $T$  is *n-large with respect to  $P$*  if, for every subexpression of  $P$  of the form  $I^* . P'$ , above every occurrence of  $\text{root}(P')$  in the tree  $T$  there are at least  $n$  ancestors labeled by  $\text{root}(I^* . P')$ . In other words, for  $T \in \langle P \rangle$  this means that in  $T$  every context appearing in  $P$  was appended at least  $n$  times, on all branches where it was possible to append it. Clearly  $\langle P \rangle \subseteq \llbracket P \rrbracket$ . On the other hand, every tree from  $\llbracket P \rrbracket$  can be embedded into every versatile tree which is large enough. We thus obtain the following lemma:

**Lemma 4.15.** For every diversified pure product  $P$ , and for every sequence of trees  $T_1, T_2, \dots \in \langle P \rangle$  such that every  $T_n$  is  $n$ -large,

$$\{T_n \mid n \in \mathbb{N}\} \downarrow = \llbracket P \rrbracket.$$

□

Using versatile trees we can reduce SUP to the diagonal problem:

**Lemma 4.16.** Let  $\mathcal{C}$  be a class of languages of finite trees closed under linear FTT transductions. SUP for  $\mathcal{C}$  reduces to the diagonal problem for  $\mathcal{C}$ .



**Proof:**

In an instance of SUP we are given a diversified pure product  $P$  and a language  $\mathcal{L} \in \mathfrak{C}$ . Consider the language of trees  $\mathcal{L}' = \mathcal{L} \downarrow \cap \langle P \rangle$ . Clearly  $\langle P \rangle$  is regular, so  $\mathcal{L}' \in \mathfrak{C}$  by Facts 4.9 and 4.10. The following claim is a direct consequence of Lemma 4.15:

**Claim 4.17.**  $\langle P \rangle \subseteq \mathcal{L} \downarrow$  if and only if for every  $n \in \mathbb{N}$  there is a tree in  $\mathcal{L}'$  that is  $n$ -large with respect to  $P$ .  $\square$

We have reduced to a problem which is very similar to the diagonal problem, except that we should put no requirement on the number of occurrences of  $\text{root}(I^*.P')$  for branches not containing an occurrence of  $\text{root}(P')$ . In order to fix this, let  $\mathcal{L}''$  be the set of trees  $T''$  obtained from some tree  $T'$  of  $\mathcal{L}'$  by the following procedure: whenever a branch of  $T'$  does not contain an occurrence of  $\text{root}(P')$ , then the leaf finishing this branch can be replaced by an arbitrarily large tree with internal nodes labeled by  $\text{root}(I^*.P')$ . Let  $\Sigma$  be the set of root labels of the form  $\text{root}(I^*.P')$  for every subexpression  $I^*.P'$  of  $P$ . The following claim is a direct consequence of the definition:

**Claim 4.18.**  $\mathcal{L}''$  is  $\Sigma$ -diagonal if and only if for every  $n \in \mathbb{N}$  there is a tree in  $\mathcal{L}'$  which is  $n$ -large with respect to  $P$ .  $\square$

The operation mapping  $\mathcal{L}'$  to  $\mathcal{L}''$  can be realized as a linear FTT transduction, and thus  $\mathcal{L}'' \in \mathfrak{C}$ . This completes the reduction from SUP to the diagonal problem.  $\square$

**Remark 4.19.** Another formulation of the diagonal problem for languages of finite trees [24, 23, 25] requires that, for every  $n \in \mathbb{N}$ , there is a tree  $T \in \mathcal{L}$  containing at least  $n$  occurrences of every letter  $a \in \Sigma$  (not necessarily on the same branch, unlike in our case). Such a formulation of the diagonal problem seems too weak to compute downward closures for languages of finite trees.

The main result of this section, Theorem 4.1, stating that the downward closure computation reduces to the diagonal problem, follows at once from Lemma 4.13 and Lemma 4.16 above.

## 5. Languages of safe recursion schemes

In the previous section, we have developed a general machinery allowing one to compute downward closures for classes of languages of finite trees closed under linear FTT transductions. In this section, we apply this machinery to the particular case of languages recognized by safe recursion schemes. The following is the main theorem of this section:

**Theorem 5.1.** Finite tree automata recognizing downward closures of languages of finite trees recognized by safe recursion schemes are computable.

In order to prove the theorem we need to recall a formalism necessary to express the diagonal problem in logic.

**Cost logics.** *Cost monadic logic (CMSO)* was introduced by Colcombet [59] as a quantitative extension of monadic second-order logic (*MSO*). As usual, the logic can be defined over any relational structure, but we restrict our attention to CMSO over trees. In addition to *first-order variables* ranging over nodes of a tree and *monadic second-order variables* (also called *set variables*) ranging over sets of nodes, CMSO uses a single additional variable  $N$ , called the *numeric variable*, which ranges over  $\mathbb{N}$ . The atomic formulas in CMSO are those from MSO (the membership relation  $x \in X$  and relations  $a(x, x_1, \dots, x_r)$  asserting that  $a \in \mathbb{A}$  of rank  $r$  is the label at node  $x$  with children  $x_1, \dots, x_r$  from left to right), as well as a new predicate  $|X| < N$ , where  $X$  is any set variable and  $N$  is the numeric variable. Arbitrary CMSO formulas are built inductively by applying Boolean connectives and by quantifying (existentially or universally) over first-order or set variables. We require that predicates of the form  $|X| < N$  appear positively in the formula (i.e., within the scope of an even number of negations). We regard  $N$  as a parameter. As usual, a *sentence* is a formula without first-order or monadic free variables; however, the parameter  $N$  is allowed to occur in a sentence. If we fix a value  $n \in \mathbb{N}$  for  $N$ , the semantics of  $|X| < N$  is what one would expect: the predicate holds when  $X$  has cardinality smaller than  $n$ . We say that a sentence  $\varphi$  *n-accepts* a tree  $T$  if it holds in  $T$  when  $n$  is used as a value of  $N$ ; it *accepts*  $T$  if it  $n$ -accepts  $T$  for some  $n \in \mathbb{N}$ .

*Weak cost monadic logic (WCMSO)* for short) is the variant of CMSO where the second-order quantification is restricted to finite sets. Vanden Boom [41, Theorem 2] proves that WCMSO is effectively equivalent to a subclass of alternating B-automata, called *weak B-automata*. Thanks to Theorem 3.1, we obtain the following corollary:

**Corollary 5.2.** Given a WCMSO formula  $\varphi$  and a safe recursion scheme  $\mathcal{G}$ , one can decide whether  $\varphi$  accepts the tree generated by  $\mathcal{G}$ .  $\square$

**Remark 5.3.** The same holds for a more expressive logic called *quasi-weak cost monadic logic (QWCMSO)* [49], whose expressive power lies between WCMSO and the CMSO. Indeed, Blumensath et al. [49, Theorem 2] prove that QWCMSO is effectively equivalent to a subclass of alternating B-automata called *quasi-weak B-automata*, and thus by Theorem 3.1 even model checking of safe recursion schemes against QWCMSO properties is decidable.

**Solving the diagonal problem.** By Theorem 4.1 all we need to do in order to obtain Theorem 5.1 is to show that (1) the diagonal problem is decidable for languages recognized by safe recursion schemes, and (2) the class of these languages is effectively closed under linear FTT transductions. We start by proving the former:

**Lemma 5.4.** The diagonal problem is decidable for the class of languages of finite trees recognized by safe recursion schemes.

Recall that in the diagonal problem we are given a safe recursion scheme  $\mathcal{G}$  and a set of letters  $\Sigma$ , and we have to determine whether for every  $n \in \mathbb{N}$  there is a tree  $T \in \mathcal{L}(\mathcal{G})$  such that there are at least  $n$  occurrences of every letter  $a \in \Sigma$  on every branch of  $T$  (we say that such a tree  $T$  is *n-large with respect to  $\Sigma$* ). In order to obtain decidability of this problem, given a set of letters  $\Sigma$ , we write a WCMSO sentence  $\varphi_\Sigma$  that  $n$ -accepts an (infinite) tree  $T$  if and only if no tree in  $\mathcal{L}(T)$  is  $n$ -large

with respect to  $\Sigma$ . Consequently,  $\varphi_\Sigma$  accepts  $T$  if for some  $n$  no tree in  $\mathcal{L}(T)$  is  $n$ -large with respect to  $\Sigma$ , that is, if  $\mathcal{L}(T)$  is not  $\Sigma$ -diagonal. Thus, in order to solve the diagonal problem, it is enough to check whether  $\varphi_\Sigma$  accepts  $\text{BT}(\mathcal{G})$  (recall that  $\mathcal{L}(\mathcal{G})$  is defined as  $\mathcal{L}(\text{BT}(\mathcal{G}))$ ), which is decidable by Corollary 5.2. It remains to construct the aforementioned sentence  $\varphi_\Sigma$ .

First, observe that the process of producing a finite tree recognized by  $\mathcal{G}$  from the infinite tree  $\text{BT}(\mathcal{G})$  generated by  $\mathcal{G}$  is expressible by a formula of WCMSO (actually, by a first-order formula):

**Lemma 5.5.** There is a WCMSO formula  $\text{tree}(X)$  that holds in a tree  $T$  if and only if  $X$  is instantiated to a set of nodes of a tree  $T' \in \mathcal{L}(T)$ , together with their nd-labeled ancestors.

**Proof:**

The formula simply says that

- $X$  is finite,
- the root of the tree belongs to  $X$ ,
- no node  $x \in X$  is  $\perp$ -labeled,
- for every nd-labeled node  $x \in X$ , exactly one among the children of  $x$  belongs to  $X$ ,
- for every node  $x \in X$  with label other than nd, all children of  $x$  belong to  $X$ , and
- if  $x \notin X$ , then no child of  $x$  belongs to  $X$ .

All the above statements can easily be expressed in WCMSO. □

Using  $\text{tree}(X)$  we construct the desired formula  $\varphi_\Sigma$ , and thus we finish the proof of Lemma 5.4:

**Lemma 5.6.** Given a set of letters  $\Sigma$ , one can compute a WCMSO sentence  $\varphi_\Sigma$  that, for every  $n \in \mathbb{N}$ ,  $n$ -accepts a tree  $T$  if and only if no tree in  $\mathcal{L}(T)$  is  $n$ -large with respect to  $\Sigma$ .

**Proof:**

We can reformulate the property as follows: for every tree  $T' \in \mathcal{L}(T)$  there is a letter  $a \in \Sigma$ , and a leaf  $x$  that has less than  $n$   $a$ -labeled ancestors. This is expressed by the following formula of WCMSO (where  $\text{leaf}(x)$  states that the node  $x$  is a leaf,  $a(x)$  that  $x$  has label  $a$ , and  $z \leq x$  that  $z$  is an ancestor of  $x$ , all being easily expressible):

$$\forall X. \left( \text{tree}(X) \rightarrow \bigvee_{a \in \Sigma} \exists x \exists Z. (x \in X \wedge \text{leaf}(x) \wedge \forall z. (z \leq x \wedge a(z) \rightarrow z \in Z) \wedge |Z| < N) \right). \quad \square$$

**Closure under transductions.** Finally, we show closure under linear FTT transductions, which allows us to apply the results of the previous section to safe recursion schemes:

**Lemma 5.7.** The class of languages of finite trees recognized by safe recursion schemes is effectively closed under linear FTT transductions.

Observe that Theorem 5.1 is a direct consequence of Theorem 4.1 and Lemmas 5.4 and 5.7. It thus remains to prove Lemma 5.7. A very similar result, albeit without the safety assumption, has been proved by Clemente, Parys, Salvati, and Walukiewicz [23, Theorem 2.1]:

**Lemma 5.8.** The class of languages of finite trees recognized by recursion schemes is effectively closed under linear FTT transductions.  $\square$

Notice that Lemma 5.7 does not follow from Lemma 5.8, since we need to additionally show that applying a linear FTT transduction to a language recognized by a safe recursion scheme preserves safety. Essentially the same construction as in the proof of Lemma 5.8 [60, Appendix A] already achieves this, albeit some modifications are needed. We now argue how to modify the proof in three aspects:

1. The proof uses the fact that higher-order recursion schemes *with states* (as introduced in that proof) are convertible to equivalent higher-order recursion schemes by increasing the arity of nonterminals. It is a simple observation that such a translation preserves safety.
2. The proof of Clemente et al. [60, Appendix A] uses the notion of *normalized* recursion schemes, wherein every rule is assumed to be of the form

$$\mathcal{R}(X) = \lambda x_1. \dots \lambda x_p. h (Y_1 x_1 \dots x_p) \dots (Y_r x_1 \dots x_p),$$

where  $h$  is either a variable  $x_i$ , or a nonterminal, or a letter, and the  $Y_j$ 's are nonterminals. This normal form is used only to simplify the presentation and is in no way essential. This is important since putting a recursion scheme in such a normal form does not preserve safety. Indeed, a subterm  $Y_j x_1 \dots x_p$  replaces some subterm  $M_j$  appearing originally in the rule for  $X$ ; if some variable  $x_i$  was not used in  $M_j$ , we could have  $\text{ord}(x_i) < \text{ord}(M_j)$  (the latter equals  $\text{ord}(Y_j x_1 \dots x_p)$ ), which violates safety of the normalized rule. On the other hand, by safety we have  $\text{ord}(x_i) \geq \text{ord}(M_j)$  if  $x_i$  appeared in  $M_j$ . Therefore, we modify the definition of the normal form to allow removal of selected variables, that is, to allow rules of the form

$$\mathcal{R}(X) = \lambda x_1. \dots \lambda x_p. h (Y_1 x_{i_{1,1}} \dots x_{i_{1,k_1}}) \dots (Y_r x_{i_{r,1}} \dots x_{i_{r,k_r}}).$$

By leaving in each subterm  $Y_j x_{i_{j,1}} \dots x_{i_{j,k_j}}$  only variables used in the replaced subterm  $M_j$ , we obtain a normalized recursion scheme that is safe.

3. The proof [60, Lemma A.3] uses also the *MSO-reflection property* of recursion schemes. In order to define this property consider a tree  $T$  and an MSO formula  $\varphi(x)$  with one free first-order variable. We define  $T_\varphi$  to be the tree obtained from  $T$  by enhancing its labels: for every node  $v$  of  $T$ , we change its label from  $a$  to  $(a, b_{\varphi,v})$ , where  $b_{\varphi,v} \in \{\mathbf{tt}, \mathbf{ff}\}$  says whether  $\varphi(v)$  is true ( $\mathbf{tt}$ ) or false ( $\mathbf{ff}$ ) in  $T$ . The MSO-reflection property says that given a recursion scheme  $\mathcal{G}$  generating a tree  $T$  and given an MSO formula  $\varphi(x)$  one can compute a recursion scheme  $\mathcal{G}'$  generating the tree  $T_\varphi$ . It was shown [15, Corollary 2] that recursion schemes indeed have the MSO-reflection property.

While switching to safe recursion schemes one needs a similar property, where both the input and the output recursion schemes are safe (this way we have a stronger conclusion under stronger assumptions). It is a folklore result that such an MSO-reflection property for safe recursion schemes holds as well. Let us support this statement in three ways:

- It is remarked by Carayol and Serre [15, Remark 5] that even a stronger property, called MSO-selection, holds for safe recursion schemes.

- To obtain a proof of the MSO-reflection property for safe recursion schemes one can take the original proof of this property for all schemes [15], and observe that the construction in this proof preserves safety. The proof uses collapsible pushdown automata, where safety corresponds to absence of collapse operations; it is thus enough to see that no collapse operations are introduced if no such operations were present on input.
- Carayol and Wöhrle [61] prove that the class of trees generated by deterministic higher-order pushdown automata is effectively closed under MSO-markings, which is essentially the same as MSO-reflection. Although Carayol and Wöhrle [61] work with edge-labeled trees, it is a routine to transfer their results to our setting of node-labeled trees (and to change MSO-markings into MSO-reflection). Moreover, a tree can be generated by a deterministic higher-order pushdown automaton if and only if it can be generated by a safe recursion scheme (see Knapik et al. [47, Theorems 5.1 and 5.3]; note, however, that the authors use the word “grammar” for a recursion scheme). Thus, the result of Carayol and Wöhrle [61] implies the desired MSO-reflection property for safe recursion schemes.

## 6. Conclusions

A tantalising direction for further work is to drop the safety assumption from Theorem 3.1, that is, to establish decidability of the model-checking problem against B-automata for trees generated by (not necessarily safe) recursion schemes. We also leave open whether downward closures are computable for this more expressive class. Another direction for further work is to analyse the complexity of the considered diagonal problem. The related problem described in Remark 4.19 is  $k$ -EXP-complete for languages of finite trees recognized by recursion schemes of order  $k$  [25], and thus not harder than the nonemptiness problem [9]. Does the same upper bound hold for the more general diagonal problem that we consider in this paper? Zetsche [62] has shown that the downward closure inclusion problem is  $\text{co-}k\text{-NEXP-hard}$  for languages of finite trees recognized by safe recursion schemes of order  $k$ . Is it possible to obtain a matching upper bound?

## References

- [1] Kobayashi N. Model Checking Higher-Order Programs. *J. ACM*, 2013. **60**(3):20:1–20:62. doi:10.1145/2487241.2487246.
- [2] Salvati S, Walukiewicz I. Simply Typed Fixpoint Calculus and Collapsible Pushdown Automata. *Math. Struct. Comput. Sci.*, 2016. **26**(7):1304–1350. doi:10.1017/S0960129514000590.
- [3] Hague M, Murawski AS, Ong CL, Serre O. Collapsible Pushdown Automata and Recursion Schemes. *ACM Trans. Comput. Log.*, 2017. **18**(3):25:1–25:42. doi:10.1145/3091122.
- [4] Clemente L, Parys P, Salvati S, Walukiewicz I. Ordered Tree-Pushdown Systems. In: Harsha P, Ramalingam G (eds.), 35th IARCS Annual Conference on Foundation of Software Technology and Theoretical Computer Science, FSTTCS 2015, December 16–18, 2015, Bangalore, India, volume 45 of *LIPICs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015 pp. 163–177. doi:10.4230/LIPICs.FSTTCS.2015.163.

- [5] Damm W. The IO- and OI-Hierarchies. *Theor. Comput. Sci.*, 1982. **20**:95–207. doi:10.1016/0304-3975(82)90009-3.
- [6] Kobele GM, Salvati S. The IO and OI Hierarchies Revisited. *Inf. Comput.*, 2015. **243**:205–221. doi:10.1016/j.ic.2014.12.015.
- [7] Aho AV. Indexed Grammars—An Extension of Context-Free Grammars. *J. ACM*, 1968. **15**(4):647–671. doi:10.1145/321479.321488.
- [8] Breveglieri L, Cherubini A, Citrini C, Crespi-Reghizzi S. Multi-push-down Languages and Grammars. *Int. J. Found. Comput. Sci.*, 1996. **7**(3):253–292. doi:10.1142/S0129054196000191.
- [9] Ong CL. On Model-Checking Trees Generated by Higher-Order Recursion Schemes. In: 21th IEEE Symposium on Logic in Computer Science (LICS 2006), 12–15 August 2006, Seattle, WA, USA, Proceedings. IEEE Computer Society, 2006 pp. 81–90. doi:10.1109/LICS.2006.38.
- [10] Hague M, Murawski AS, Ong CL, Serre O. Collapsible Pushdown Automata and Recursion Schemes. In: Proceedings of the Twenty-Third Annual IEEE Symposium on Logic in Computer Science, LICS 2008, 24–27 June 2008, Pittsburgh, PA, USA. IEEE Computer Society, 2008 pp. 452–461. doi:10.1109/LICS.2008.34.
- [11] Kobayashi N, Ong CL. A Type System Equivalent to the Modal Mu-Calculus Model Checking of Higher-Order Recursion Schemes. In: Proceedings of the 24th Annual IEEE Symposium on Logic in Computer Science, LICS 2009, 11–14 August 2009, Los Angeles, CA, USA. IEEE Computer Society, 2009 pp. 179–188. doi:10.1109/LICS.2009.29.
- [12] Salvati S, Walukiewicz I. Krivine Machines and Higher-Order Schemes. *Inf. Comput.*, 2014. **239**:340–355. doi:10.1016/j.ic.2014.07.012.
- [13] Parys P. Higher-Order Model Checking Step by Step. In: Bansal N, Merelli E, Worrell J (eds.), 48th International Colloquium on Automata, Languages, and Programming, ICALP 2021, July 12–16, 2021, Glasgow, Scotland (Virtual Conference), volume 198 of *LIPIcs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021 pp. 140:1–140:16. doi:10.4230/LIPIcs.ICALP.2021.140.
- [14] Broadbent CH, Ong CL. On Global Model Checking Trees Generated by Higher-Order Recursion Schemes. In: de Alfaro L (ed.), Foundations of Software Science and Computational Structures, 12th International Conference, FOSSACS 2009, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2009, York, UK, March 22–29, 2009. Proceedings, volume 5504 of *Lecture Notes in Computer Science*. Springer, 2009 pp. 107–121. doi:10.1007/978-3-642-00596-1\_9.
- [15] Broadbent CH, Carayol A, Ong CL, Serre O. Recursion Schemes and Logical Reflection. In: Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science, LICS 2010, 11–14 July 2010, Edinburgh, United Kingdom. IEEE Computer Society, 2010 pp. 120–129. doi:10.1109/LICS.2010.40.
- [16] Carayol A, Serre O. Collapsible Pushdown Automata and Labeled Recursion Schemes: Equivalence, Safety and Effective Selection. In: Proceedings of the 27th Annual IEEE Symposium on Logic in Computer Science, LICS 2012, Dubrovnik, Croatia, June 25–28, 2012. IEEE Computer Society, 2012 pp. 165–174. doi:10.1109/LICS.2012.73.
- [17] Salvati S, Walukiewicz I. A Model for Behavioural Properties of Higher-Order Programs. In: Kreutzer S (ed.), 24th EACSL Annual Conference on Computer Science Logic, CSL 2015, September 7–10, 2015, Berlin, Germany, volume 41 of *LIPIcs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015 pp. 229–243. doi:10.4230/LIPIcs.CSL.2015.229.

- [18] Stockmeyer LJ. The Complexity of Decision Problems in Automata Theory and Logic. Ph.D. thesis, MIT, 1974.
- [19] Broadbent CH, Kobayashi N. Saturation-Based Model Checking of Higher-Order Recursion Schemes. In: Rocca SRD (ed.), Computer Science Logic 2013 (CSL 2013), CSL 2013, September 2–5, 2013, Torino, Italy, volume 23 of *LIPICs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2013 pp. 129–148. doi:10.4230/LIPICs.CSL.2013.129.
- [20] Kobayashi N. A Practical Linear Time Algorithm for Trivial Automata Model Checking of Higher-Order Recursion Schemes. In: Hofmann M (ed.), Foundations of Software Science and Computational Structures – 14th International Conference, FOSSACS 2011, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2011, Saarbrücken, Germany, March 26–April 3, 2011. Proceedings, volume 6604 of *Lecture Notes in Computer Science*. Springer, 2011 pp. 260–274. doi:10.1007/978-3-642-19805-2\_18.
- [21] Neatherway RP, Ong CL. TravMC2: Higher-Order Model Checking for Alternating Parity Tree Automata. In: Rungta N, Tkachuk O (eds.), 2014 International Symposium on Model Checking of Software, SPIN 2014, Proceedings, San Jose, CA, USA, July 21–23, 2014. ACM, 2014 pp. 129–132. doi:10.1145/2632362.2632381.
- [22] Ramsay SJ, Neatherway RP, Ong CL. A Type-Directed Abstraction Refinement Approach to Higher-Order Model Checking. In: Jagannathan S, Sewell P (eds.), The 41st Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL '14, San Diego, CA, USA, January 20–21, 2014. ACM, 2014 pp. 61–72. doi:10.1145/2535838.2535873.
- [23] Clemente L, Parys P, Salvati S, Walukiewicz I. The Diagonal Problem for Higher-Order Recursion Schemes is Decidable. In: Grohe M, Koskinen E, Shankar N (eds.), Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5–8, 2016. ACM, 2016 pp. 96–105. doi:10.1145/2933575.2934527.
- [24] Hague M, Kochems J, Ong CL. Unboundedness and Downward Closures of Higher-Order Pushdown Automata. In: Bodík R, Majumdar R (eds.), Proceedings of the 43rd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2016, St. Petersburg, FL, USA, January 20–22, 2016. ACM, 2016 pp. 151–163. doi:10.1145/2837614.2837627.
- [25] Parys P. A Type System Describing Unboundedness. *Discret. Math. Theor. Comput. Sci.*, 2020. **22**(4). doi:10.23638/DMTCS-22-4-2.
- [26] Zetsche G. An Approach to Computing Downward Closures. In: Halldórsson MM, Iwama K, Kobayashi N, Speckmann B (eds.), Automata, Languages, and Programming – 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6–10, 2015, Proceedings, Part II, volume 9135 of *Lecture Notes in Computer Science*. Springer, 2015 pp. 440–451. doi:10.1007/978-3-662-47666-6\_35.
- [27] Czerwiński W, Martens W, van Rooijen L, Zeitoun M, Zetsche G. A Characterization for Decidable Separability by Piecewise Testable Languages. *Discret. Math. Theor. Comput. Sci.*, 2017. **19**(4). doi:10.23638/DMTCS-19-4-1.
- [28] Czerwiński W, Martens W, van Rooijen L, Zeitoun M. A Note on Decidable Separability by Piecewise Testable Languages. In: Kosowski A, Walukiewicz I (eds.), Fundamentals of Computation Theory – 20th International Symposium, FCT 2015, Gdańsk, Poland, August 17–19, 2015, Proceedings, volume 9210 of *Lecture Notes in Computer Science*. Springer, 2015 pp. 173–185. doi:10.1007/978-3-319-22177-9\_14.
- [29] Higman G. Ordering by Divisibility in Abstract Algebras. *Proc. London Math. Soc.*, 1952. **s3-2**(1):326–336. doi:10.1112/plms/s3-2.1.326.

- [30] Bachmeier G, Luttenberger M, Schlund M. Finite Automata for the Sub- and Superword Closure of CFLs: Descriptive and Computational Complexity. In: Dediu A, Formenti E, Martín-Vide C, Truthe B (eds.), *Language and Automata Theory and Applications – 9th International Conference, LATA 2015, Nice, France, March 2–6, 2015, Proceedings*, volume 8977 of *Lecture Notes in Computer Science*. Springer, 2015 pp. 473–485. doi:10.1007/978-3-319-15579-1\_37.
- [31] Courcelle B. On Constructing Obstruction Sets of Words. *Bull. EATCS*, 1991. **44**:178–186.
- [32] van Leeuwen J. Effective Constructions in Well-Partially-Ordered Free Monoids. *Discret. Math.*, 1978. **21**(3):237–252. doi:10.1016/0012-365X(78)90156-5.
- [33] Habermehl P, Meyer R, Wimmel H. The Downward-Closure of Petri Net Languages. In: Abramsky S, Gavioille C, Kirchner C, auf der Heide FM, Spirakis PG (eds.), *Automata, Languages and Programming, 37th International Colloquium, ICALP 2010, Bordeaux, France, July 6–10, 2010, Proceedings, Part II*, volume 6199 of *Lecture Notes in Computer Science*. Springer, 2010 pp. 466–477. doi:10.1007/978-3-642-14162-1\_39.
- [34] Zetsche G. Computing Downward Closures for Stacked Counter Automata. In: Mayr EW, Ollinger N (eds.), *32nd International Symposium on Theoretical Aspects of Computer Science, STACS 2015, March 4–7, 2015, Garching, Germany*, volume 30 of *LIPIcs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015 pp. 743–756. doi:10.4230/LIPIcs.STACS.2015.743.
- [35] Abdulla PA, Boasson L, Bouajjani A. Effective Lossy Queue Languages. In: Orejas F, Spirakis PG, van Leeuwen J (eds.), *Automata, Languages and Programming, 28th International Colloquium, ICALP 2001, Crete, Greece, July 8–12, 2001, Proceedings*, volume 2076 of *Lecture Notes in Computer Science*. Springer, 2001 pp. 639–651. doi:10.1007/3-540-48224-5\_53.
- [36] Goubault-Larrecq J, Schmitz S. Deciding Piecewise Testable Separability for Regular Tree Languages. In: Chatzigiannakis I, Mitzenmacher M, Rabani Y, Sangiorgi D (eds.), *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11–15, 2016, Rome, Italy*, volume 55 of *LIPIcs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2016 pp. 97:1–97:15. doi:10.4230/LIPIcs.ICALP.2016.97.
- [37] Bojańczyk M, Colcombet T. Bounds in  $\omega$ -Regularity. In: *21th IEEE Symposium on Logic in Computer Science (LICS 2006)*, 12–15 August 2006, Seattle, WA, USA, Proceedings. IEEE Computer Society, 2006 pp. 285–296. doi:10.1109/LICS.2006.17.
- [38] Colcombet T. The Theory of Stabilisation Monoids and Regular Cost Functions. In: Albers S, Marchetti-Spaccamela A, Matias Y, Nikolettseas SE, Thomas W (eds.), *Automata, Languages and Programming, 36th International Colloquium, ICALP 2009, Rhodes, Greece, July 5–12, 2009, Proceedings, Part II*, volume 5556 of *Lecture Notes in Computer Science*. Springer, 2009 pp. 139–150. doi:10.1007/978-3-642-02930-1\_12.
- [39] Colcombet T, Löding C. Regular Cost Functions Over Finite Trees. In: *Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science, LICS 2010, 11–14 July 2010, Edinburgh, United Kingdom*. IEEE Computer Society, 2010 pp. 70–79. doi:10.1109/LICS.2010.36.
- [40] Kuperberg D, Vanden Boom M. On the Expressive Power of Cost Logics Over Infinite Words. In: Czumaj A, Mehlhorn K, Pitts AM, Wattenhofer R (eds.), *Automata, Languages, and Programming – 39th International Colloquium, ICALP 2012, Warwick, UK, July 9–13, 2012, Proceedings, Part II*, volume 7392 of *Lecture Notes in Computer Science*. Springer, 2012 pp. 287–298. doi:10.1007/978-3-642-31585-5\_28.



- [41] Vanden Boom M. Weak Cost Monadic Logic Over Infinite Trees. In: Murlak F, Sankowski P (eds.), *Mathematical Foundations of Computer Science 2011 – 36th International Symposium, MFCS 2011, Warsaw, Poland, August 22–26, 2011. Proceedings*, volume 6907 of *Lecture Notes in Computer Science*. Springer, 2011 pp. 580–591. doi:10.1007/978-3-642-22993-0\_52.
- [42] Kuperberg D, Vanden Boom M. Quasi-Weak Cost Automata: A New Variant of Weakness. In: Chakraborty S, Kumar A (eds.), *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2011, December 12–14, 2011, Mumbai, India, volume 13 of LIPIcs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2011 pp. 66–77. doi:10.4230/LIPIcs.FSTTCS.2011.66.
- [43] Grädel E, Thomas W, Wilke T (eds.). *Automata, Logics, and Infinite Games: A Guide to Current Research [outcome of a Dagstuhl seminar, February 2001]*, volume 2500 of *Lecture Notes in Computer Science*. Springer, 2002. ISBN 3-540-00388-6. doi:10.1007/3-540-36387-4.
- [44] Colcombet T, Löding C. The Non-deterministic Mostowski Hierarchy and Distance-Parity Automata. In: Aceto L, Damgård I, Goldberg LA, Halldórsson MM, Ingólfssdóttir A, Walukiewicz I (eds.), *Automata, Languages and Programming, 35th International Colloquium, ICALP 2008, Reykjavik, Iceland, July 7–11, 2008, Proceedings, Part II – Track B: Logic, Semantics, and Theory of Programming & Track C: Security and Cryptography Foundations*, volume 5126 of *Lecture Notes in Computer Science*. Springer, 2008 pp. 398–409. doi:10.1007/978-3-540-70583-3\_33.
- [45] Bojańczyk M. A Bounding Quantifier. In: Marcinkowski J, Tarlecki A (eds.), *Computer Science Logic, 18th International Workshop, CSL 2004, 13th Annual Conference of the EACSL, Karpacz, Poland, September 20–24, 2004, Proceedings*, volume 3210 of *Lecture Notes in Computer Science*. Springer, 2004 pp. 41–55. doi:10.1007/978-3-540-30124-0\_7.
- [46] Parys P. Recursion Schemes, the MSO Logic, and the U Quantifier. *Log. Methods Comput. Sci.*, 2020. **16**(1). doi:10.23638/LMCS-16(1:20)2020.
- [47] Knapik T, Niwiński D, Urzyczyn P. Higher-Order Pushdown Trees Are Easy. In: Nielsen M, Engberg U (eds.), *Foundations of Software Science and Computation Structures, 5th International Conference, FOSACS 2002. Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2002 Grenoble, France, April 8–12, 2002, Proceedings*, volume 2303 of *Lecture Notes in Computer Science*. Springer, 2002 pp. 205–222. doi:10.1007/3-540-45931-6\_15.
- [48] Parys P. On the Expressive Power of Higher-Order Pushdown Systems. *Log. Methods Comput. Sci.*, 2020. **16**(3). doi:10.23638/LMCS-16(3:11)2020.
- [49] Blumensath A, Colcombet T, Kuperberg D, Parys P, Vanden Boom M. Two-Way Cost Automata and Cost Logics Over Infinite Trees. In: Henzinger TA, Miller D (eds.), *Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), CSL-LICS '14, Vienna, Austria, July 14–18, 2014. ACM, 2014 pp. 16:1–16:9. doi:10.1145/2603088.2603104.*
- [50] Finkel A, Goubault-Larrecq J. Forward Analysis for WSTS, Part I: Completions. In: Albers S, Marion J (eds.), *26th International Symposium on Theoretical Aspects of Computer Science, STACS 2009, February 26–28, 2009, Freiburg, Germany, Proceedings*, volume 3 of *LIPIcs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Germany, 2009 pp. 433–444. doi:10.4230/LIPIcs.STACS.2009.1844.
- [51] Barozzini D, Clemente L, Colcombet T, Parys P. Cost Automata, Safe Schemes, and Downward Closures. In: Czumaj A, Dawar A, Merelli E (eds.), *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, July 8–11, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 168 of

- LIPICs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020 pp. 109:1–109:18. doi:10.4230/LIPICs.ICALP.2020.109. (Best paper award for Track B).
- [52] Berarducci A, Dezani-Ciancaglini M. Infinite Lambda-Calculus and Types. *Theor. Comput. Sci.*, 1999. **212**(1-2):29–75. doi:10.1016/S0304-3975(98)00135-2.
  - [53] Kennaway R, Klop JW, Sleep MR, de Vries F. Infinitary Lambda Calculus. *Theor. Comput. Sci.*, 1997. **175**(1):93–125. doi:10.1016/S0304-3975(96)00171-5.
  - [54] Blum W, Ong CL. The Safe Lambda Calculus. *Log. Methods Comput. Sci.*, 2009. **5**(1). doi:10.2168/LMCS-5(1:3)2009.
  - [55] Haddad A. IO vs OI in Higher-Order Recursion Schemes. In: Miller D, Ésik Z (eds.), Proceedings 8th Workshop on Fixed Points in Computer Science, FICS 2012, Tallinn, Estonia, 24th March 2012, volume 77 of *EPTCS*. 2012 pp. 23–30. doi:10.4204/EPTCS.77.4.
  - [56] Colcombet T, Göller S. Games With Bound Guess Actions. In: Grohe M, Koskinen E, Shankar N (eds.), Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5–8, 2016. ACM, 2016 pp. 257–266. doi:10.1145/2933575.2934502.
  - [57] Knapik T, Niwiński D, Urzyczyn P. Deciding Monadic Theories of Hyperalgebraic Trees. In: Abramsky S (ed.), Typed Lambda Calculi and Applications, 5th International Conference, TLCA 2001, Kraków, Poland, May 2–5, 2001, Proceedings, volume 2044 of *Lecture Notes in Computer Science*. Springer, 2001 pp. 253–267. doi:10.1007/3-540-45413-6.21.
  - [58] Comon H, Dauchet M, Gilleron R, Jacquemard F, Lugiez D, Löding C, Tison S, Tommasi M. Tree Automata Techniques and Applications, 2007. URL <http://tata.gforge.inria.fr/>.
  - [59] Colcombet T. Regular Cost Functions, Part I: Logic and Algebra Over Words. *Log. Methods Comput. Sci.*, 2013. **9**(3). doi:10.2168/LMCS-9(3:3)2013.
  - [60] Clemente L, Parys P, Salvati S, Walukiewicz I. The Diagonal Problem for Higher-Order Recursion Schemes is Decidable. *CoRR*, 2016. **abs/1605.00371**.
  - [61] Carayol A, Wöhrle S. The Caucal Hierarchy of Infinite Graphs in Terms of Logic and Higher-Order Pushdown Automata. In: Pandya PK, Radhakrishnan J (eds.), FST TCS 2003: Foundations of Software Technology and Theoretical Computer Science, 23rd Conference, Mumbai, India, December 15–17, 2003, Proceedings, volume 2914 of *Lecture Notes in Computer Science*. Springer, 2003 pp. 112–123. doi:10.1007/978-3-540-24597-1.10.
  - [62] Zetsche G. The Complexity of Downward Closure Comparisons. In: Chatzigiannakis I, Mitzenmacher M, Rabani Y, Sangiorgi D (eds.), 43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11–15, 2016, Rome, Italy, volume 55 of *LIPICs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2016 pp. 123:1–123:14. doi:10.4230/LIPICs.ICALP.2016.123.
  - [63] Parys P. Homogeneity Without Loss of Generality. In: Kirchner H (ed.), 3rd International Conference on Formal Structures for Computation and Deduction, FSCD 2018, July 9–12, 2018, Oxford, UK, volume 108 of *LIPICs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018 pp. 27:1–27:15. doi:10.4230/LIPICs.FSCD.2018.27.
  - [64] Barendregt H. The Lambda Calculus - Its Syntax and Semantics. Elsevier Science Publishers Ltd., 1984.
  - [65] Curry H, Feys R. Combinatory Logic - Volume I. North-Holland Publishing Company, 1958.

## A. Proof of Lemma 3.7

In this appendix, we provide a self-contained proof of Lemma 3.7. The proof in its essence comes from the papers of Knapik et al. [47, 57], up to some minor details.

### A.1. Preparatory steps

A type  $\alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \circ$  is *homogeneous* if  $\text{ord}(\alpha_1) \geq \dots \geq \text{ord}(\alpha_k)$  and all  $\alpha_1, \dots, \alpha_k$  are homogeneous. A recursion scheme  $\mathcal{G} = \langle \mathbb{A}, \mathcal{N}, X_0, \mathcal{R} \rangle$  is *homogeneous* if types of all nonterminals in  $\mathcal{N}$  are homogeneous. Notice that then also the type of every subterm of  $\mathcal{R}(X)$  is homogeneous, for every nonterminal  $X \in \mathcal{N}$ . It is known that every (safe) recursion scheme can be made homogeneous:

**Lemma A.1.** ([63, Theorems 8 and 9])

For every safe recursion scheme  $\mathcal{G}$  one can construct a homogeneous safe recursion scheme  $\mathcal{H}$  of the same order, such that  $\text{BT}(\mathcal{H}) = \text{BT}(\mathcal{G})$ .  $\square$

Thanks to Lemma A.1 we may assume that the recursion scheme  $\mathcal{G}$  given in Lemma 3.7 is homogeneous. We remark that homogeneity of  $\mathcal{G}$  is not at all essential in the remainder of the proof; this assumption is just for technical convenience. Namely, thanks to this assumption, the notion of order-0 arguments coincides with the notion of arguments occurring after the last argument of positive order.

It is also convenient to assume that every nonterminal of positive order takes some parameter of order 0. Again, this can be achieved without loss of generality:

**Lemma A.2.** For every homogeneous safe recursion scheme  $\mathcal{G}$  one can construct a homogeneous safe recursion scheme  $\mathcal{H}$  of the same order, such that  $\text{BT}(\mathcal{H}) = \text{BT}(\mathcal{G})$ , and such that every nonterminal of  $\mathcal{H}$  having positive order takes some parameter of order 0 (i.e., there are no nonterminals of type  $\alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \circ$  with  $k \geq 1$  and  $\text{ord}(\alpha_k) \geq 1$ ).

**Proof:**

We say that a type is *bad* if it is of the form  $\alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \circ$  with  $k \geq 1$  and  $\text{ord}(\alpha_k) \geq 1$ . We add one additional order-0 parameter to every lambda-term of bad type. More formally, recall that rules of  $\mathcal{G}$  are of the form  $\mathcal{R}(X) = \lambda x_1. \dots \lambda x_k. K$ , where  $K$  is an applicative term of type  $\circ$ . If the type of  $X$  is bad, we replace this rule by  $\lambda x_1. \dots \lambda x_k. \lambda y. M$  for a fresh variable  $y$  of type  $\circ$ . Note that in  $\Lambda(\mathcal{G})$  this inserts an additional lambda-binder  $\lambda y$  between  $\lambda x$  and  $K$  in every subterm of the form  $\lambda x. K$  with  $\text{ord}(x) \geq 1$  and  $\text{ord}(K) = 0$ . Simultaneously, we replace every application  $K L$  with  $\text{ord}(L) \geq 1$  and  $\text{ord}(K L) = 0$  by  $K L \perp$ : whenever the last argument is applied to a lambda-term having a bad type, we apply an additional order-0 argument, which is chosen to be  $\perp$  (but can be any lambda-term of type  $\circ$ ). This changes the types of lambda-terms as follows: every type  $\alpha = (\alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \circ)$  changes

- to  $\alpha'_1 \rightarrow \dots \rightarrow \alpha'_k \rightarrow \circ \rightarrow \circ$  if  $\alpha$  was bad, and
- to  $\alpha'_1 \rightarrow \dots \rightarrow \alpha'_k \rightarrow \circ$  otherwise,

where  $\alpha'_1, \dots, \alpha'_k$  are obtained by the same transformation applied to the types  $\alpha_1, \dots, \alpha_k$ . It is tedious but straightforward to formally check that this way we obtain a valid recursion scheme  $\mathcal{H}$ , and that it generates the same tree as  $\mathcal{G}$ .  $\square$

## A.2. Reification: Defining $\mathcal{G}^\bullet$

Fix some normalizing homogeneous safe recursion scheme  $\mathcal{G} = \langle \mathbb{A}, \mathcal{N}, X_0, \mathcal{R} \rangle$ , where every nonterminal of positive order takes some parameter of order 0. Let  $\mathcal{X}$  be the (finite) set of order-0 variables used for parameters in  $\mathcal{G}$ .

The *maximal arity* of a type  $\alpha$ , denoted  $\text{mar}(\alpha)$  is defined by induction:

$$\text{mar}(\alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \mathbf{o}) = \max \left( \{k\} \cup \{\text{mar}(\alpha_i) \mid 1 \leq i \leq k\} \right).$$

The *maximal arity* of a lambda-term  $M$ , denoted  $\text{mar}(M)$ , equals

$$\text{mar}(M) = \sup \{ \text{mar}(\alpha) \mid \alpha \text{ is a type of a subterm of } M \}.$$

Finally, the *maximal arity* of a recursion scheme  $\mathcal{G} = \langle \mathbb{A}, \mathcal{N}, X_0, \mathcal{R} \rangle$ , denoted  $\text{mar}(\mathcal{G})$ , equals

$$\text{mar}(\mathcal{G}) = \max \{ \text{mar}(X) \mid X \in \mathcal{N} \}.$$

Observe that  $\text{mar}(\mathcal{R}(X)) \leq \text{mar}(\mathcal{G})$  for every nonterminal  $X$  of  $\mathcal{G}$  (because the only variables occurring in  $\mathcal{R}(X)$  are nonterminals of  $\mathcal{G}$  and parameters of  $X$ ). It follows that  $\text{mar}(\Lambda(\mathcal{G})) \leq \text{mar}(\mathcal{G})$ .

We say that  $M$  is an *input lambda-term* if

- $M$  uses letters from the alphabet  $\mathbb{A}$ ;
- all order-0 variables used in  $M$ , other than nonterminals from  $\mathcal{N}$ , belong to  $\mathcal{X}$ ;
- nonterminals from  $\mathcal{N}$  are not used in lambda-binders in  $M$ ;
- types of all subterms of  $M$  are homogeneous;
- for every lambda-abstraction subterm  $\lambda x_1. \dots \lambda x_k. K$  of  $M$ , where  $k \geq 1$  and  $K$  is not a lambda-abstraction, we have  $\text{ord}(x_k) = \text{ord}(K) = 0$ ;
- $\text{mar}(M) \leq \text{mar}(\mathcal{G})$ ;
- no subterm of  $M$  is an *infinite application*  $\dots M_3 M_2 M_1$ .

Note that the above conditions are satisfied by  $M = \mathcal{R}(X)$  for all nonterminals  $X \in \mathcal{N}$ , as well as by  $M = \Lambda(\mathcal{G})$ . Moreover, every subterm of an input lambda-term is an input lambda-term.

We additionally require that in a first-order input lambda-term  $M$  every free variable of  $M$  belongs to  $\mathcal{X}$  (i.e., no nonterminals, even of order 0, may occur in  $M$ ).

We define a function  $(\cdot)^\bullet$ , called *reification*; it maps an input lambda-term  $M$  of a homogeneous type  $\alpha$  to a corresponding lambda-term  $M^\bullet$  of a homogeneous type  $\alpha^\bullet$ , using letters from the alphabet  $\mathbb{A}_{\mathcal{X}}$ , defined in Section 3. We also say that the lambda-term  $M^\bullet$  *represents* the lambda-term  $M$ . Moreover, if  $M$  is first-order, then  $M^\bullet$  is in fact a lambda-tree, that is, it does not contain variables nor lambda-binders (cf. Lemma A.3).

For every homogeneous type  $\alpha$ , the type  $\alpha^\bullet$  is defined by induction on the structure of  $\alpha$ : if

$$\alpha = \alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \mathbf{o} \rightarrow \dots \rightarrow \mathbf{o} \rightarrow \mathbf{o},$$

where  $k = 0$  or  $\alpha_k \neq \mathbf{o}$ , then we take

$$\alpha^\bullet = \alpha_1^\bullet \rightarrow \dots \rightarrow \alpha_k^\bullet \rightarrow \mathbf{o}.$$

In other words, order-0 arguments are discarded and the transformation is applied recursively to higher-order arguments. For instance,  $\mathbf{o}^\bullet = \mathbf{o}$ ,  $(\mathbf{o} \rightarrow \mathbf{o})^\bullet = \mathbf{o}$ , and  $((\mathbf{o} \rightarrow \mathbf{o}) \rightarrow \mathbf{o} \rightarrow \mathbf{o})^\bullet = \mathbf{o} \rightarrow \mathbf{o}$ . It is easy to see (by induction on the structure of  $\alpha$ ) that  $\text{ord}(\alpha^\bullet) = \max(0, \text{ord}(\alpha) - 1)$ .

We now define reification of an input lambda-term  $M$ . First, to every nonterminal  $X \in \mathcal{N}$  of type  $\alpha$  we assign a unique nonterminal  $X^\bullet$  of type  $\alpha^\bullet$ . Likewise, to every variable  $x \notin (\mathcal{X} \cup \mathcal{N})$  of type  $\alpha$  we assign a unique variable  $x^\bullet$  of type  $\alpha^\bullet$ . Next, we proceed by coinduction on the structure of  $M$ :

1.  $(a)^\bullet = \bar{a}$ ;
2.  $(X)^\bullet = X^\bullet$  if  $X \in \mathcal{N}$  (i.e., the result of the  $(\_)^\bullet$  operation for a nonterminal  $X$  is the nonterminal denoted  $X^\bullet$ );
3.  $(x)^\bullet = \bar{x}$  if  $x \notin \mathcal{N}$  and  $\text{ord}(x) = 0$  (i.e., if  $x \in \mathcal{X}$ );
4.  $(x)^\bullet = x^\bullet$  if  $x \notin \mathcal{N}$  and  $\text{ord}(x) > 0$ ;
5.  $(\lambda x.K)^\bullet = \bar{\lambda x} K^\bullet$  if  $\text{ord}(x) = 0$  (i.e., if  $x \in \mathcal{X}$ );
6.  $(\lambda x.K)^\bullet = \lambda x^\bullet.K^\bullet$  if  $\text{ord}(x) > 0$ ;
7.  $(K L)^\bullet = @ K^\bullet L^\bullet$  if  $\text{ord}(L) = 0$ ;
8.  $(K L)^\bullet = K^\bullet L^\bullet$  if  $\text{ord}(L) > 0$ .

Observe (by coinduction) that if  $M$  has type  $\alpha$  then  $M^\bullet$  is a lambda-term of type  $\alpha^\bullet$ . This is immediate in Cases 2, 3, and 4. In Case 1, a letter  $a$  has type of the form  $\alpha = (\mathbf{o} \rightarrow \dots \rightarrow \mathbf{o} \rightarrow \mathbf{o})$ , while  $\bar{a}$  has type  $\alpha^\bullet = \mathbf{o}$ . In Case 5 we use the assumption that the type  $\mathbf{o} \rightarrow \beta$  of  $\lambda x.K$  is homogeneous, which implies  $\text{ord}(K) = \text{ord}(\beta) \leq 1$ , that is,  $\text{ord}(K^\bullet) = 0$ . Likewise in Case 7 we use the assumption that the type  $\mathbf{o} \rightarrow \beta$  of  $K$  is homogeneous, which implies  $\text{ord}(K) = \text{ord}(\mathbf{o} \rightarrow \beta) \leq 1$ , that is,  $\text{ord}(K^\bullet) = 0$ . This is necessary, because the lambda-terms  $\bar{\lambda x} K^\bullet$  and  $@ K^\bullet L^\bullet$  make sense only if  $\text{ord}(K^\bullet) = 0$ . In Cases 6 and 8 we observe that  $(\beta \rightarrow \gamma)^\bullet = \beta^\bullet \rightarrow \gamma^\bullet$  if  $\text{ord}(\beta) > 0$ .

There is one delicate point of the definition above. Namely, lambda-terms are usually identified up to renaming bound variables (alpha-conversion). The result of the reification operation  $(\_)^\bullet$ , however, depends on particular names given to bound order-0 variables (these names become written explicitly in the letters (constants)  $\bar{x}$  and  $\bar{\lambda x}$ ). Thus, it is understood that no implicit renaming of bound order-0 variables is performed for lambda-terms to which the  $(\_)^\bullet$  operation is going to be applied.

When starting from a lambda-term that is first-order (defined on Page 5), we can see that Cases 2, 4, 6, and 8 can never occur. In such a circumstance, reification produces a lambda-tree.

**Lemma A.3.** If an input lambda-term  $M$  is first-order then  $M^\bullet$  is a lambda-tree. □

Using the reification operation  $(\_)^\bullet$  for lambda-terms, we can define the resulting recursion scheme  $\mathcal{G}^\bullet$ : we take

$$\mathcal{G}^\bullet = \langle \mathbb{A}_{\mathcal{X}}, \mathcal{N}^\bullet, X_0^\bullet, \mathcal{R}^\bullet \rangle, \quad (11)$$

where  $\mathcal{N}^\bullet = \{X^\bullet \mid X \in \mathcal{N}\}$  and  $\mathcal{R}^\bullet(X^\bullet) = (\mathcal{R}(X))^\bullet$  for all  $X \in \mathcal{N}$ .

It is easy to see that  $\mathcal{G}^\bullet$  is of order  $m - 1$  if  $\mathcal{G}$  was of order  $m \geq 1$ : the order of every nonterminal, if positive, drops by one. Let us now observe that  $\mathcal{G}^\bullet$  is safe:

**Lemma A.4.** If  $\mathcal{G}$  is safe, then  $\mathcal{G}^\bullet$  is safe.

**Proof:**

Recall that, by definition,  $\mathcal{G}$  is safe when the lambda-term  $\Lambda(\mathcal{G})$  is safe; likewise for  $\mathcal{G}^\bullet$  and  $\Lambda(\mathcal{G}^\bullet)$ . First, it is easy to see that  $\Lambda(\mathcal{G}^\bullet) = (\Lambda(\mathcal{G}))^\bullet$ . In order to ensure that  $\mathcal{G}^\bullet$  is safe, we thus need to ensure that every subterm of  $(\Lambda(\mathcal{G}))^\bullet$  occurring in argument position of some application is superficially safe. Subterms occurring in argument position of an application in  $(\Lambda(\mathcal{G}))^\bullet$  are

- $K^\bullet$  in  $\bar{\lambda}x K^\bullet$ ,
- $K^\bullet$  and  $L^\bullet$  in  $@ K^\bullet L^\bullet$ , and
- $L^\bullet$  in  $K^\bullet L^\bullet$ .

In the first two cases, the subterms are of order 0, so they are automatically superficially safe. In the last case,  $L$  occurs in argument position of the application  $K L$  in  $\Lambda(\mathcal{G})$ , which means that  $L$  is superficially safe; we have  $\text{ord}(x) \geq \text{ord}(L)$  for every free variable  $x$  of  $L$ . Every free variable of  $L^\bullet$  is of the form  $x^\bullet$  for  $x$  being a free variable of  $L$ ; we then have  $\text{ord}(x^\bullet) = \max(0, \text{ord}(x) - 1) \geq \max(0, \text{ord}(L) - 1) = \text{ord}(L^\bullet)$ , as required.  $\square$

The relation between  $\mathcal{G}^\bullet$  and  $\mathcal{G}$  is described by the following lemma:

**Lemma A.5.** There exists a closed first-order input lambda-term  $M$  of type  $\mathbf{o}$  such that

$$\text{BT}(\mathcal{G}^\bullet) = M^\bullet \quad \text{and} \quad \text{BT}(M) = \text{BT}(\mathcal{G}).$$

Notice that there is at most one lambda-term  $M$  such that  $\text{BT}(\mathcal{G}^\bullet) = M^\bullet$ . So the lambda-term  $M$  in the lemma above is in fact unique. The remaining part of this subsection is devoted to the proof of Lemma A.5.

First, let us see that safety is preserved by beta-reductions:<sup>4</sup>

**Lemma A.6.** If  $M$  is safe and  $M \rightarrow_\beta N$ , then  $N$  is safe.

**Proof:**

Before starting, let us state two inductive properties of safety, following directly from its definition:

Inductive Property 1:  $\lambda x.P$  is safe if, and only if,  $P$  is safe;

Inductive Property 2:  $P Q$  is safe if, and only if,  $P$  and  $Q$  are safe, and  $Q$  is superficially safe.

Next, let us prove an auxiliary claim concerning substitution:

**Claim A.7.** If  $K$  and  $L$  are safe, and  $L$  is superficially safe, then  $K[L/x]$  is safe.

We prove this claim by structural coinduction. When  $x$  is not free in  $K$ , or when  $K = x$ , then  $K[L/x]$  equals  $K$  or  $L$ , respectively, and the thesis holds by assumption. When  $K = \lambda y.P$ , the thesis is an immediate consequence of the coinduction hypothesis and Inductive Property 1. The only remaining case is that  $K = P Q$ . By the coinduction hypothesis we obtain that  $P[L/x]$  and  $Q[L/x]$  are safe. To conclude, we also need to know that  $Q[L/x]$  is superficially safe (cf. Inductive Property 2). If  $x$  is not free in  $Q$ , then this is immediate:  $Q[L/x] = Q$  and the latter is superficially safe by

<sup>4</sup>Blum and Ong [54] write that safety is not preserved by arbitrary beta-reductions, only by beta-reductions of a special kind. Note, however, that they consider a slightly different definition of safe lambda-terms (leading to the same definition of safe recursion schemes).

assumption. Otherwise, every free variable  $y$  of  $Q[L/x]$  is free either in  $Q$  or in  $L$ . In the former case we simply have that  $\text{ord}(y) \geq \text{ord}(Q) = \text{ord}(Q[L/x])$ , because  $Q$  is superficially safe; in the latter case we have  $\text{ord}(y) \geq \text{ord}(L) = \text{ord}(x) \geq \text{ord}(Q) = \text{ord}(Q[L/x])$ , because  $L$  and  $Q$  are superficially safe and  $x$  is free in  $Q$ . It follows that  $Q[L/x]$  is superficially safe and thus  $K[L/x]$  is safe, as required.

We can now come back to the proof of Lemma A.6, which we perform by induction on the depth of the considered redex. The base case, when  $M = (\lambda x.K) L$  and  $N = K[L/x]$ , is provided directly by Claim A.7 (note that  $L$  occurs in argument position in  $M$ , so it is superficially safe by safety of  $M$ ). For the induction step, we have three cases:

1.  $M = \lambda x.P$  and  $N = \lambda x.P'$ , where  $P \rightarrow_\beta P'$ ;
2.  $M = P Q$  and  $N = P' Q$ , where  $P \rightarrow_\beta P'$ ;
3.  $M = P Q$  and  $N = P Q'$ , where  $Q \rightarrow_\beta Q'$ .

In the first two cases, we simply use the induction hypothesis for  $P \rightarrow_\beta P'$ . In the last case, we also need to observe that  $Q'$  is superficially safe, which holds because  $Q$  is superficially safe, and every free variable of  $Q'$  is free already in  $Q$ .  $\square$

**Beta-reductions of positive order.** It is useful to consider *beta-reductions of positive order*, denoted “ $\rightarrow_{\beta+}$ ”: We have  $M \rightarrow_{\beta+} N$  if  $N$  is obtained from  $M$  by replacing some subterm  $(\lambda x.K) L$  thereof with  $K[L/x]$ , where we additionally require  $\text{ord}(x) \geq 1$ .

We use the “ $\rightarrow_{\beta+}$ ” relation only for safe lambda-terms, and when writing  $M \rightarrow_{\beta+} N$  we implicitly assume that names of bound order-0 variables do not change. Note that if  $M$  is safe, then the argument  $L$  of the redex is superficially safe. It follows that every free variable  $y$  of  $L$  satisfies  $\text{ord}(y) \geq \text{ord}(L) = \text{ord}(x) \geq 1$ . In other words,  $L$  has no free variables of order 0. Thus there is no danger that these variables will conflict with bound order-0 variables in  $K$ ; there is never the need to rename bound order-0 variables.

Recall that the  $(\_)^\bullet$  operation is defined only for input lambda-terms, as defined at the beginning of the subsection. With the above assumption in hand, we have that if  $M$  is a safe input lambda-term and  $M \rightarrow_{\beta+} N$ , then  $N$  is also an input lambda-term (most importantly, all order-0 variables used in  $N$ , other than nonterminals, belong to  $\mathcal{X}$ ); in particular, it makes sense to write  $N^\bullet$ .

Our next lemma connects the “ $\rightarrow_{\beta+}$ ” relation with the “ $\rightarrow_\beta$ ” relation and reification:

**Lemma A.8.** Let  $M$  be a safe input lambda-term.

1. If  $M \rightarrow_{\beta+} N$ , then  $M^\bullet \rightarrow_\beta N^\bullet$ .
2. If  $M^\bullet \rightarrow_\beta O$ , then  $O = N^\bullet$  for a lambda-term  $N$  such that  $M \rightarrow_{\beta+} N$ .

In order to prove Lemma A.8, we first need to see that higher-order substitution commutes with reification:

**Lemma A.9.** For every input lambda-term of the form  $K[L/x]$ , where  $\text{ord}(x) \geq 1$ , we have

$$(K[L/x])^\bullet = K^\bullet[L^\bullet/x^\bullet].$$

**Proof:**

Follows directly from the definition of reification.  $\square$

In Lemma A.9 we implicitly assume that the substitution  $K[L/x]$  does not change names of bound order-0 variables in  $K$ . As already said, it is never needed to rename them if  $L$  does not have free order-0 variables, that is, when  $(\lambda x.K) L$  is a subterm of a safe lambda-term. Note also that Lemma A.9 does not make sense when  $x$  has order zero, because in that case there is no variable  $x^\bullet$  (the variable  $x$  is reified to  $\bar{x}$ , which is a letter).

**Proof of Lemma A.8:**

For the first item, suppose that  $N$  is obtained from  $M$  by replacing a redex  $(\lambda x.K) L$  with  $K[L/x]$ , where  $\text{ord}(x) = \text{ord}(L) \geq 1$ . Then in  $M^\bullet$  we have a redex

$$((\lambda x.K) L)^\bullet = (\lambda x^\bullet.K^\bullet) L^\bullet,$$

which beta-reduces to  $K^\bullet[L^\bullet/x^\bullet] = (K[L/x])^\bullet$  (equality by Lemma A.9). We thus have  $M^\bullet \rightarrow_\beta N^\bullet$ .

For the second item, observe that the definition of reification produces a lambda-binder only in Case 6, and an application whose operator is not a letter only in Case 8. Thus the redex of  $M^\bullet$  reduced in  $M^\bullet \rightarrow_\beta O$  is necessarily of the form

$$((\lambda x.K) L)^\bullet = (\lambda x^\bullet.K^\bullet) L^\bullet \rightarrow_\beta K^\bullet[L^\bullet/x^\bullet] = (K[L/x])^\bullet, \quad (12)$$

where  $\text{ord}(x) = \text{ord}(L) \geq 1$ , and where the second equality follows from Lemma A.9. Let  $N$  be obtained from  $M$  by reducing (the corresponding occurrence of)  $(\lambda x.K) L$  to  $K[L/x]$ , and thus  $M \rightarrow_{\beta+} N$ . A structural induction on the subterms of  $O$  (the base case being provided by Formula (12)) shows  $O = N^\bullet$ , as required.  $\square$

One of consequences of Lemma A.8 is the Church-Rosser property for  $\rightarrow_{\beta+}$ :

**Lemma A.10.** If  $M \rightarrow_{\beta+}^* N_1$  and  $M \rightarrow_{\beta+}^* N_2$  for a safe input lambda-term  $M$ , then  $N_1 \rightarrow_{\beta+}^* P$  and  $N_2 \rightarrow_{\beta+}^* P$  for some lambda-term  $P$ .

**Proof:**

By Item 1 of Lemma A.8 (and using also Lemma A.6 to ensure that lambda-terms under consideration are safe) we have  $M^\bullet \rightarrow_\beta^* N_1^\bullet$  and  $M^\bullet \rightarrow_\beta^* N_2^\bullet$ . The Church-Rosser property for  $\rightarrow_\beta$  gives us a lambda-term  $O$  such that  $N_1^\bullet \rightarrow_\beta^* O$  and  $N_2^\bullet \rightarrow_\beta^* O$ . Then, by Item 2 of Lemma A.8 (and again by Lemma A.6) we obtain lambda-terms  $P_1$  and  $P_2$  such that  $O = P_1^\bullet = P_2^\bullet$ , and  $N_1 \rightarrow_{\beta+}^* P_1$  and  $N_2 \rightarrow_{\beta+}^* P_2$ . Observing that the reification operation  $(\cdot)^\bullet$  is injective, we actually have  $P_1 = P_2$ , so this lambda-term can be taken as  $P$  in the thesis.  $\square$

Let  $M$  be a (possibly infinite) safe input lambda-term of order at most 1 (we mean here the order of the type of  $M$ ; subterms of  $M$  may have higher order) such that all free variables thereof belong to  $\mathcal{X}$ . We define the first-order lambda-term obtained as the *limit* of applying  $\rightarrow_{\beta+}$  reductions to  $M$ , denoted  $\text{BT}^+(M)$ , analogously to how  $\text{BT}(P)$  is defined as the limit of applying the  $\rightarrow_\beta$  reductions to a closed lambda-term  $P$  of type o. The definition is coinductive:

- if  $M \rightarrow_{\beta+}^* a$  (for a letter  $a$ ), then  $\text{BT}^+(M) = a$ ,
- if  $M \rightarrow_{\beta+}^* x$  (for a variable  $x \in \mathcal{X}$ ), then  $\text{BT}^+(M) = x$ ,



- if  $M \rightarrow_{\beta+}^* \lambda x.N$  with  $x \in \mathcal{X}$ , then  $\text{BT}^+(M) = \lambda x.(\text{BT}^+(N))$ , and
- if  $M \rightarrow_{\beta+}^* K L$  with  $\text{ord}(L) = 0$ , then  $\text{BT}^+(M) = (\text{BT}^+(K)) (\text{BT}^+(L))$ .

Clearly  $\text{BT}^+(M)$  is a first-order input lambda-term of the same type as  $M$ .

Observe that the above definition covers all cases. To this end consider all possible forms of  $M$ . If  $M = a$  or  $M = K L$  with  $\text{ord}(L) = 0$ , we have the first or the last case of the definition, respectively. If  $M = x$ , then  $x \in \mathcal{X}$  by the assumption that all free variables of  $M$  belong to  $\mathcal{X}$ ; we have the second case. If  $M = \lambda x.N$ , then  $\text{ord}(x) = 0$  by the assumption that  $\text{ord}(M) \leq 1$ , hence  $x \in \mathcal{X}$  (because  $M$  is an input lambda-term); we have the third case. The only remaining case is that  $M$  is an application with an argument of positive order. Because  $M$  is an input lambda-term, it cannot be an infinite application. Thus,  $M$  can be written as  $H M_1 \dots M_r$ , where  $H$  is not an application,  $r \geq 1$ , and  $\text{ord}(M_r) \geq 1$ . Then  $H$  cannot be a letter (arguments of a letter are all of type  $\circ$ ) nor a variable (all free variables of  $M$  are of type  $\circ$ , because they belong to  $\mathcal{X}$ );  $H$  has to start with a sequence of lambda-binders:  $H = \lambda x_1. \dots \lambda x_k. K$ , where  $k \geq 1$  and  $K$  does not start with a lambda-binder. One of the assumptions for being an input lambda-term implies that  $\text{ord}(x_k) = \text{ord}(K) = 0$ . Then necessarily  $k \geq r$  (each of the provided arguments corresponds to some lambda-binder), and  $\text{ord}(x_r) = \text{ord}(M_r) \geq 1$  implies that  $k > r$ . Moreover, because  $M$  is of order (at most) 1, the variables  $x_{r+1}, \dots, x_k$  are of type  $\circ$ . On the other hand,  $x_1, \dots, x_r$  are of order at least 1, by homogeneity. Thus  $M \rightarrow_{\beta+}^* \lambda x_{r+1}. \dots \lambda x_k. K[M_1/x_1, \dots, M_r/x_r]$ ; we obtain the third case.

Moreover, thanks to Lemma A.10, the resulting lambda-term  $\text{BT}^+(M)$  is uniquely defined.

We now use Lemma A.8 to show a kind of commutativity property between reification and Böhm trees:

**Lemma A.11.** Let  $M$  be a safe input lambda-term of order at most 1, all free variables of which belong to  $\mathcal{X}$ . Then  $\text{BT}(M^\bullet) = (\text{BT}^+(M))^\bullet$ .

**Proof:**

We proceed by coinduction. At every step we use Lemma A.8 (and Lemma A.6 to obtain safety of intermediate lambda-terms) to deduce  $M^\bullet \rightarrow_\beta^* N^\bullet$  from  $M \rightarrow_{\beta+}^* N$ . According to the definition of  $\text{BT}^+(M)$  we have four cases:

- If  $M \rightarrow_{\beta+}^* a$ , then  $M^\bullet \rightarrow_\beta^* a^\bullet = \bar{a}$ , so  $(\text{BT}^+(M))^\bullet = a^\bullet = \bar{a} = \text{BT}(M^\bullet)$ .
- If  $M \rightarrow_{\beta+}^* x$  with  $\text{ord}(x) = 0$ , then  $M^\bullet \rightarrow_\beta^* x^\bullet = \bar{x}$ , so  $(\text{BT}^+(M))^\bullet = x^\bullet = \bar{x} = \text{BT}(M^\bullet)$ .
- If  $M \rightarrow_{\beta+}^* \lambda x.N$  with  $\text{ord}(x) = 0$ , then  $M^\bullet \rightarrow_\beta^* (\lambda x.N)^\bullet = \overline{\lambda x} N$ , so  $(\text{BT}^+(M))^\bullet = (\lambda x.(\text{BT}^+(N)))^\bullet = \overline{\lambda x} (\text{BT}^+(N))^\bullet = \overline{\lambda x} (\text{BT}(N^\bullet)) = \text{BT}(M^\bullet)$ , where the third equality is by the coinductive hypothesis.
- If  $M \rightarrow_{\beta+}^* K L$  with  $\text{ord}(L) = 0$ , then  $M^\bullet \rightarrow_\beta^* (K L)^\bullet = @ K^\bullet L^\bullet$ , so

$$\begin{aligned} (\text{BT}^+(M))^\bullet &= ((\text{BT}^+(K)) (\text{BT}^+(L)))^\bullet = @ (\text{BT}^+(K))^\bullet (\text{BT}^+(L))^\bullet \\ &= @ (\text{BT}(K^\bullet)) (\text{BT}(L^\bullet)) = \text{BT}(M^\bullet), \end{aligned}$$

where the third equality is by the coinductive hypothesis. □

The other important property of  $\text{BT}^+(\cdot)$  is that all higher-order reductions can be performed first, followed by all (necessarily) order-zero reductions. This is formally stated in the next lemma:<sup>5</sup>

**Lemma A.12.**  $\text{BT}(\text{BT}^+(\Lambda(\mathcal{G}))) = \text{BT}(\mathcal{G})$ .

Before proving Lemma A.12, let us see how Lemma A.5 follows from Lemmas A.11 and A.12:

**Proof of Lemma A.5:**

We take  $N = \Lambda(\mathcal{G})$  and  $M = \text{BT}^+(N)$ . It is easy to check that  $M$  is a closed first-order input lambda-term of type  $\circ$ . We have  $\text{BT}(\mathcal{G}^\bullet) = \text{BT}(\Lambda(\mathcal{G}^\bullet)) = \text{BT}(N^\bullet) = (\text{BT}^+(N))^\bullet$  by Lemma A.11, and  $\text{BT}(M) = \text{BT}(\text{BT}^+(N)) = \text{BT}(N) = \text{BT}(\mathcal{G})$  by Lemma A.12.  $\square$

It remains to prove Lemma A.12. Our proof strategy is to show that the two Böhm trees mentioned in the lemma are equal by showing that they agree on every finite prefix. To this end, we have to define finite cuts of a lambda-term.

**Finite cuts.** For every type  $\alpha$  let us fix a fresh variable  $x_\perp^\alpha$  of type  $\alpha$ , not occurring anywhere in  $\Lambda(\mathcal{G})$ , and called a *cut variable*. We say that  $F$  is a *cut* of  $M$  if  $F$  is obtained from  $M$  by replacing some of its subterms with cut variables (of appropriate type). For example,  $\lambda y. x_\perp^{\circ \rightarrow \circ}$  is a cut of  $\lambda y. \lambda z. a \ y \ z$ : we have replaced the subterm  $\lambda z. a \ y \ z$  of type  $\circ \rightarrow \circ$  with the variable  $x_\perp^{\circ \rightarrow \circ}$ . We are particularly interested in finite cuts, that is, cuts that are finite lambda-terms.

We say that a cut  $F$  is an *order-0 cut* if the only cut variable occurring in  $F$  is  $x_\perp^\circ$  (i.e., only subterms of type  $\circ$  are cut off). We have the following nice property of the lambda-term  $\Lambda(\mathcal{G})$ :

**Lemma A.13.** For every finite cut  $F$  of  $\Lambda(\mathcal{G})$  there exists a finite order-0 cut  $F_0$  of  $\Lambda(\mathcal{G})$  such that  $F$  is a cut of  $F_0$ .

**Proof:**

Consider a subterm of  $\Lambda(\mathcal{G})$  that was replaced by  $x_\perp^\alpha$ . It is necessarily of the form  $K[M_1/X_1, \dots, M_k/X_k]$ , where  $K$  is a subterm of  $\mathcal{R}(X)$  for some nonterminal  $X$ , hence  $K$  is finite. Every lambda-term  $M_i$ , substituted for the nonterminal  $X_i$ , is obtained by further substituting lambda-terms in  $\mathcal{R}(X_i)$ , hence it is of the form  $\lambda x_{i,1}. \dots \lambda x_{i,n_i}. K_i$ , where  $K_i$  is of type  $\circ$ . Instead of cutting off the whole  $K[M_1/X_1, \dots, M_k/X_k]$ , we can rather cut off at every occurrence of  $K_i$ . Our cut remains finite, but all cut variables are of type  $\circ$ .  $\square$

The next lemma says that the relation of being a cut is a simulation with respect to “ $\rightarrow_\beta$ ” and “ $\rightarrow_{\beta+}$ ” reductions.

**Lemma A.14.** Let  $F$  be a cut of  $M$ .

<sup>5</sup>We remark that Lemma A.12 can be generalized to say that  $\text{BT}(\text{BT}^+(M)) = \text{BT}(M)$  for any closed input lambda-term  $M$  of order 0, not necessarily for  $M = \Lambda(\mathcal{G})$ . The lemma can even be further generalized to say that  $\text{BT}(N) = \text{BT}(M)$  whenever  $N$  is obtained as an (appropriately defined) limit of applying any finite or infinite sequence of beta-reductions to  $M$ . Nevertheless, we prove only the specific statement written above—in Lemma A.13 we explicitly use the fact that the lambda-term is of the form  $\Lambda(\mathcal{G})$ .

1. If  $F \rightarrow_\beta G$ , then  $G$  is a cut of a lambda-term  $N$  such that  $M \rightarrow_\beta N$ .
2. Likewise, if  $F \rightarrow_{\beta+} G$ , then  $G$  is a cut of a lambda-term  $N$  such that  $M \rightarrow_{\beta+} N$ .

**Proof:**

We just reduce the redex of  $M$  whose cut was reduced in  $F \rightarrow_\beta G$  (in  $F \rightarrow_{\beta+} G$ , respectively). It is easy to check that  $G$  is indeed a cut of the resulting lambda-term  $N$ .  $\square$

We also need to state formally in which sense a lambda-term agrees with a finite prefix of a tree. Let  $n \in \mathbb{N}$ , let  $M$  be a lambda-term, and let  $T$  be a tree. We define when  $M$  agrees with  $T$  up to level  $n$ , by induction on  $n$ :

- every  $M$  agrees with every  $T$  up to level 0;
- $M$  agrees with  $T$  up to level  $n + 1$  if  $M = a M_1 \dots M_r$ ,  $T = a T_1 \dots T_r$ , and  $M_i$  agrees with  $T_i$  up to level  $n$ , for every  $i \in \{1, \dots, r\}$ .

The next lemma says that every finite prefix of  $\text{BT}(M)$  depends only on some finite prefix of  $M$ :

**Lemma A.15.** Let  $M$  be a closed normalizing lambda-term of type o. For every  $n \in \mathbb{N}$  there exists a finite cut  $F$  of  $M$ , and a lambda-term  $G$  such that  $F \rightarrow_\beta^* G$  and  $G$  agrees with  $\text{BT}(M)$  up to level  $n$ .  $\square$

We skip the proof of Lemma A.15, which is a standard fact. A very similar lemma is shown for instance in Parys [25, Lemma 4.2]. In Lemma A.15 it is important that  $M$  is normalizing, so that every node of  $\text{BT}(M)$  is created after finitely many reductions from  $M$ . When  $M = \Lambda(\mathcal{G})$ , we can strengthen Lemma A.15 as follows:

**Lemma A.16.** For every  $n \in \mathbb{N}$  there exists a finite order-0 cut  $F_0$  of  $\Lambda(\mathcal{G})$ , and a lambda-term  $G_0$  such that  $F_0 \rightarrow_\beta^* G_0$  and  $G_0$  agrees with  $\text{BT}(\mathcal{G})$  up to level  $n$ .

**Proof:**

First, from Lemma A.15 we obtain a finite cut  $F$  of  $\Lambda(\mathcal{G})$ , and a lambda-term  $G$  such that  $F \rightarrow_\beta^* G$  and  $G$  agrees with  $\text{BT}(\mathcal{G}) = \text{BT}(\Lambda(\mathcal{G}))$  up to level  $n$ . It is not necessarily an order-0 cut, but by Lemma A.13 we can extend it to a finite order-0 cut  $F_0$  (such that  $F$  is a cut of  $F_0$ ). Then, by Lemma A.14 we know that  $G$  is a cut of some  $G_0$  such that  $F_0 \rightarrow_\beta^* G_0$ . It is easy to see that if  $G$  agrees with some tree (in particular, with  $\text{BT}(\mathcal{G})$ ) up to some level  $n$ , and  $G$  is a cut of  $G_0$ , then also  $G_0$  agrees with this tree up to the same level  $n$ .  $\square$

**Lemma A.17.** Let  $H$  be a cut of two trees,  $T_1$  and  $T_2$ . If  $H$  agrees with  $T_1$  up to some level  $n$ , then both  $H$  and  $T_1$  agree with  $T_2$  up to level  $n$ .

**Proof:**

Straightforward: if  $H$  agrees with  $T_1$  up to some level  $n$ , then cut variables may appear in  $H$  only below this level.  $\square$

Recall that a lambda-term is in beta-normal form if it does not contain any redex.

**Lemma A.18.** Let  $H$  be a finite order-0 cut of a closed lambda-term  $M$  of type  $\mathbf{o}$ . If  $H$  is in beta-normal form, then  $H$  is also a cut of  $\text{BT}(M)$ .

**Proof:**

By induction on the size of  $H$ . Let us write  $H = H_0 H_1 \dots H_r$ , where  $H_0$  is not an application. If  $H_0$  is a letter  $a$ , then  $M = a M_1 \dots M_r$ , where for every  $i \in \{1, \dots, r\}$  the lambda-term  $M_i$  is closed and of type  $\mathbf{o}$ , and  $H_i$  is a finite order-0 cut of  $M_i$ , and is in beta-normal form. By the induction hypothesis, every  $H_i$  is also a cut of  $\text{BT}(M_i)$ , which gives the thesis due to  $\text{BT}(M) = a (\text{BT}(M_1)) \dots (\text{BT}(M_r))$ . If  $H_0$  is a variable, then necessarily  $H_0 = x_\perp^\mathbf{o}$  (because  $M$  is closed) and  $r = 0$ ;  $x_\perp^\mathbf{o}$  is a cut of every lambda-term. Finally, if  $H_0$  is a lambda-abstraction, then necessarily  $r \geq 1$  (because the type of the whole term  $H$  is  $\mathbf{o}$ ), which contradicts the assumption that  $H$  is in beta-normal form.  $\square$

**Lemma A.19.** Let  $M$  be a safe input lambda-term  $M$  of order at most 1 such that all free variables thereof belong to  $\mathcal{X}$ , and let  $G$  be a finite order-0 cut of  $M$ . If no “ $\rightarrow_{\beta^+}$ ” reduction can be executed from  $G$ , then  $G$  is also a cut of  $\text{BT}^+(M)$ .

**Proof:**

By induction on the size of  $G$ . If  $G = x_\perp^\mathbf{o}$ , then it is a cut of every lambda-term. If  $G$  is a variable  $x$  other than  $x_\perp^\mathbf{o}$ , but necessarily from  $\mathcal{X}$  (by assumption), then also  $M = x = \text{BT}^+(M)$ , and the thesis is clear. Likewise, if  $G$  is a letter  $a$ , then also  $M = a = \text{BT}^+(M)$ , and the thesis is clear.

Suppose that  $G = \lambda x.G'$ . We have  $M = \lambda x.M'$ , where  $G'$  is a finite order-0 cut of  $M'$ . Then necessarily  $x \in \mathcal{X}$  (because  $M$  is an input lambda-term and  $\text{ord}(M) \leq 1$ ). The induction hypothesis can be applied to  $G'$  and  $M'$ , implying that  $G'$  is a cut of  $\text{BT}^+(M')$ . Then  $G$  is a cut of  $\text{BT}^+(M) = \lambda x.(\text{BT}^+(M'))$ .

Next, suppose that  $G = G_0 G_1$  with  $\text{ord}(G_1) = 0$ . Then  $M = M_0 M_1$ , where  $G_0$  and  $G_1$  are finite order-0 cuts of  $M_0$  and  $M_1$ , respectively. The induction hypothesis implies that  $G_0$  and  $G_1$  are also cuts of  $\text{BT}^+(M_0)$  and  $\text{BT}^+(M_1)$ , respectively. Then  $G$  is a cut of  $\text{BT}^+(M) = (\text{BT}^+(M_1)) (\text{BT}^+(M_2))$ .

Finally, suppose that  $G = G_0 G_1 \dots G_r$ , where  $G_0$  is not an application,  $r \geq 1$ , and  $\text{ord}(G_r) \geq 1$ . Note that  $G_0$  cannot be a letter nor a variable (of type  $\mathbf{o}$ , by assumption), because they do not take arguments of positive order. So  $G_0$  is a lambda-abstraction. But  $\text{ord}(G_1) \geq \text{ord}(G_r) \geq 1$  by homogeneity, which means that “ $\rightarrow_{\beta^+}$ ” can be applied to the redex  $G_0 G_1$ , contrary to the assumption; thus this case is actually impossible.  $\square$

**Proof of Lemma A.12:**

In order to prove that  $\text{BT}(\text{BT}^+(\Lambda(\mathcal{G}))) = \text{BT}(\mathcal{G})$ , it is enough to prove that  $\text{BT}(\text{BT}^+(\Lambda(\mathcal{G})))$  agrees with  $\text{BT}(\mathcal{G})$  up to every level  $n \in \mathbb{N}$ . Fix some  $n \in \mathbb{N}$ , and consider a finite order-0 cut  $F$  of  $\Lambda(\mathcal{G})$  such that  $F \rightarrow_\beta^* H$  and  $H$  agrees with  $\text{BT}(\mathcal{G})$  up to level  $n$ . The cut  $F$  exists by Lemma A.16. Observe also that if  $H$  agrees with  $\text{BT}(\mathcal{G})$  up to level  $n$ , and  $H \rightarrow_\beta H'$ , then  $H'$  also agrees with  $\text{BT}(\mathcal{G})$  up to level  $n$ . Recall that finite simply-typed lambda-terms are strongly normalizing, which in particular means that no infinite sequence of beta-reductions can start in  $H$ . We can thus assume from this point on, without loss of generality, that  $H$  is in beta-normal form.

Let also  $G$  be a lambda-term such that  $F \rightarrow_{\beta+}^* G$ , but no further “ $\rightarrow_{\beta+}$ ” reductions can be executed from  $G$  (i.e.,  $G$  is in  $\rightarrow_{\beta+}$ -normal form). Using strong normalization again, we have that  $G \rightarrow_{\beta}^* H$ . Recall that  $F$  is a (finite, order-0) cut of  $\Lambda(\mathcal{G})$ . Due to  $F \rightarrow_{\beta}^* H$ , by Lemma A.14 we know that  $H$  is a cut of a lambda-term  $M$  such that  $\Lambda(\mathcal{G}) \rightarrow_{\beta}^* M$ ; then Lemma A.18 implies that  $H$  is also a cut of  $\text{BT}(M) = \text{BT}(\Lambda(\mathcal{G})) = \text{BT}(\mathcal{G})$ . Likewise, due to  $F \rightarrow_{\beta+}^* G$ , by Lemma A.14 we know that  $G$  is a (finite, order-0) cut of a lambda-term  $P$  such that  $\Lambda(\mathcal{G}) \rightarrow_{\beta+}^* P$ ; then Lemma A.19 implies that  $G$  is also a cut of  $\text{BT}^+(P) = \text{BT}^+(\Lambda(\mathcal{G}))$ . Having this, and due to  $G \rightarrow_{\beta}^* H$ , by Lemma A.14 we know that  $H$  is a cut of a lambda-term  $Q$  such that  $\text{BT}^+(\Lambda(\mathcal{G})) \rightarrow_{\beta}^* Q$ ; then Lemma A.18 implies that  $H$  is also a cut of  $\text{BT}(Q) = \text{BT}(\text{BT}^+(\Lambda(\mathcal{G})))$ .

We thus know that  $H$  is a cut of both  $\text{BT}(\mathcal{G})$  and  $\text{BT}(\text{BT}^+(\Lambda(\mathcal{G})))$ , and that it agrees with  $\text{BT}(\mathcal{G})$  up to level  $n$ . In such a situation Lemma A.17 implies that the two trees agree up to level  $n$ , as required.  $\square$

### A.3. From the Böhm tree to the derived tree

We have already defined a safe recursion scheme  $\mathcal{G}^\bullet$ , being of order smaller by one than the order of  $\mathcal{G}$ , and such that

$$\text{BT}(\mathcal{G}^\bullet) = M^\bullet \quad \text{and} \quad \text{BT}(M) = \text{BT}(\mathcal{G}) \quad (13)$$

for some closed first-order input lambda-term  $M$  of type  $\circ$  (cf. Lemma A.5). For Lemma 3.7 we rather need the equality

$$\llbracket \text{BT}(\mathcal{G}^\bullet) \rrbracket_{\mathcal{X}, \text{mar}(\mathcal{G})} = \text{BT}(\mathcal{G}). \quad (14)$$

Because  $\mathcal{G}$  is normalizing,  $M$  is normalizing as well (recall that  $\mathcal{G}$ , resp.,  $M$ , is normalizing if  $\text{BT}(\mathcal{G})$ , resp.,  $\text{BT}(M)$ , does not contain the special letter  $\perp$ ). Thus, Equality (14) follows immediately from Equalities (13) and from the following lemma:

**Lemma A.20.** Let  $M$  be a closed normalizing first-order input lambda-term of type  $\circ$ , and let  $s = \text{mar}(\mathcal{G})$ . Then  $M^\bullet$  is a lambda-tree and moreover

$$\llbracket M^\bullet \rrbracket_{\mathcal{X}, s} = \text{BT}(M).$$

While proving Lemma A.20, we identify a node with a finite sequence of numbers from  $\{1, 2, \dots\}$ , which denote directions when going down the tree (1 for the first child, 2 for the second child, and so on). Thus,  $\varepsilon$  is the root, and the  $i$ -th child of a node  $v \in \{1, 2, \dots\}^*$  is the node  $v \cdot i$ .

It is convenient to consider a more restrictive notion of beta-reduction, namely head beta-reduction. We say that  $M$  *head beta-reduces* to  $N$ , written  $M \rightarrow_{\text{h}\beta} N$ , if  $M$  can be written as

$$M = (\lambda x.K) L L_1 \dots L_j \quad (\text{for some } j \geq 0),$$

and  $N = K[L/x] L_1 \dots L_j$ .<sup>6</sup> When writing  $M \rightarrow_{\text{h}\beta} N$ , we implicitly assume that names of bound variables do not change. Note that when  $M$  as above is closed, then  $L$  is closed as well, and thus

<sup>6</sup>The usual definition of head beta-reduction allows additionally a sequence of lambda-binders outside the two lambda-terms, i.e.,  $M$  is of the form  $\lambda y_1. \dots \lambda y_k. ((\lambda x.K) L L_1 \dots L_j)$ . In our case, we consider head beta-reductions only for lambda-terms of type  $\circ$ , and thus such a sequence of lambda-binders does not exist, i.e.,  $k = 0$ .

indeed there is no need to rename bound variables in  $K$  while performing head beta-reductions. The following is a known fact (c.f. [64, Paragraph 11.4.7, “Standardization theorem”], where it is attributed to Curry and Feys [65]):

**Lemma A.21. (Standardization theorem)**

The Böhm tree can be constructed using only head beta-reductions (instead of arbitrary beta-reductions). In other words, for every closed normalizing lambda-term  $M$  of type  $\mathbf{o}$  we have

$$\text{BT}(M) = a (\text{BT}(M_1)) \dots (\text{BT}(M_r))$$

for some lambda-term  $N = a M_1 \dots M_r$  such that  $M \rightarrow_{\text{h}\beta}^* N$ .

The next lemma states that derived trees are invariant under head beta-reductions:

**Lemma A.22.** If  $M \rightarrow_{\text{h}\beta} N$ , where  $M, N$  are closed first-order input lambda-terms of type  $\mathbf{o}$  and  $N^\bullet$  is normalizing,<sup>7</sup> then  $\llbracket M^\bullet \rrbracket_{\mathcal{X},s} = \llbracket N^\bullet \rrbracket_{\mathcal{X},s}$ .

Before proving Lemma A.22 we show immediately how it is used in the proof of Lemma A.20:

**Proof of Lemma A.20:**

We have already proved in Lemma A.3 that reification for a first-order input lambda-term results in a lambda-tree. In order to prove that  $\llbracket M^\bullet \rrbracket_{\mathcal{X},s} = \text{BT}(M)$ , we proceed by coinduction on the Böhm tree. By Lemma A.21, we have

$$\text{BT}(M) = a (\text{BT}(M_1)) \dots (\text{BT}(M_r))$$

for some lambda-term  $N = a M_1 \dots M_r$  such that  $M \rightarrow_{\text{h}\beta}^* N$ . By the definition of reification, we have

$$N^\bullet = @ (\dots (@ (@ \bar{a} M_1^\bullet) M_2^\bullet) \dots) M_r^\bullet.$$

Let us now compute the derived tree  $\llbracket N^\bullet \rrbracket_{\mathcal{X},s}$ . Following the definition for several successor steps, we arrive at

$$\llbracket N^\bullet \rrbracket_{\mathcal{X},s} = \llbracket N^\bullet, \downarrow, \varepsilon \rrbracket_{\mathcal{X},s} = a \llbracket N^\bullet, \uparrow_1, 1^r \rrbracket_{\mathcal{X},s} \dots \llbracket N^\bullet, \uparrow_r, 1^r \rrbracket_{\mathcal{X},s},$$

where  $1^r$  is the node labeled by  $\bar{a}$  (i.e., the node reached by going  $r$  times left from the root). Performing a few more successor steps from  $(\uparrow_i, 1^r)$  we see, for every  $i \in \{1, \dots, r\}$ , that

$$\llbracket N^\bullet, \uparrow_i, 1^r \rrbracket_{\mathcal{X},s} = \llbracket N^\bullet, \downarrow, 1^{r-i} \cdot 2 \rrbracket_{\mathcal{X},s},$$

where  $1^{r-i} \cdot 2$  is the root of the subtree  $M_i^\bullet$  (i.e., the node reached by going  $r-i$  times left and then one time right from the root). By the coinductive assumption applied to  $M_i$  we have  $\llbracket M_i^\bullet \rrbracket_{\mathcal{X},s} = \text{BT}(M_i)$ .

<sup>7</sup>The lemma holds also when  $N^\bullet$  is not normalizing, but then some additional arguments are needed in the proof. In the following, we need only the version when  $N^\bullet$  is normalizing, for which we provide an easier argument.

By assumption these trees do not contain the special letter  $\perp$  (i.e.,  $M_i$  is normalizing), so the sequence of successors used to define  $\llbracket M_i^\bullet \rrbracket_{\mathcal{X},s}$  never tries to go up from the root of  $M_i^\bullet$ . It follows that

$$\llbracket N^\bullet, \Downarrow, 1^{r-i} \cdot 2 \rrbracket_{\mathcal{X},s} = \llbracket M_i^\bullet \rrbracket_{\mathcal{X},s}.$$

Putting the pieces together yields

$$\llbracket N^\bullet \rrbracket_{\mathcal{X},s} = a \llbracket M_1^\bullet \rrbracket_{\mathcal{X},s} \dots \llbracket M_r^\bullet \rrbracket_{\mathcal{X},s} = a (\text{BT}(M_1)) \dots (\text{BT}(M_r)) = \text{BT}(M).$$

In particular, we now know that the lambda-tree  $N^\bullet$  is normalizing. Recalling that  $M \rightarrow_{\text{h}\beta}^* N$ , we can conclude with the equality  $\llbracket M^\bullet \rrbracket_{\mathcal{X},s} = \llbracket N^\bullet \rrbracket_{\mathcal{X},s}$  obtained by a repeated use of Lemma A.22. More precisely, consider the sequence of head beta-reductions

$$M = N_0 \rightarrow_{\text{h}\beta} N_1 \rightarrow_{\text{h}\beta} \dots \rightarrow_{\text{h}\beta} N_k = N$$

leading from  $M$  to  $N$ . We can prove by induction on  $i \in \{0, \dots, k\}$  that  $\llbracket N_{k-i}^\bullet \rrbracket_{\mathcal{X},s} = \llbracket N^\bullet \rrbracket_{\mathcal{X},s}$ . The base case  $i = 0$  holds trivially. For the inductive case  $i > 0$ , we apply Lemma A.22 to  $N_{k-i}$  and  $N_{k-i+1}$ ; we know that  $N_{k-i+1}^\bullet$  is normalizing due to the induction hypothesis  $\llbracket N_{k-i+1}^\bullet \rrbracket_{\mathcal{X},s} = \llbracket N^\bullet \rrbracket_{\mathcal{X},s}$ . The required equality  $\llbracket M^\bullet \rrbracket_{\mathcal{X},s} = \llbracket N^\bullet \rrbracket_{\mathcal{X},s}$  follows by taking  $i = k$ .  $\square$

Heading towards the proof of Lemma A.22, we introduce some notions. Let  $M$  be a closed first-order input lambda-term of type  $\mathbf{o}$ ,  $d \in \text{Dir}_{\mathcal{X},s}$  a direction, and  $v$  a node in the reified lambda-tree  $M^\bullet$ . We call a triple  $\langle M^\bullet, d, v \rangle$  a *configuration*. For two configurations  $c, b$  let  $c \rightarrow_{\mathcal{X},s} b$  if  $b$  is the  $(\mathcal{X}, s)$ -successor of  $c$  (recall that the successor is unique, if defined).

Consider a configuration  $\langle M^\bullet, d, v \rangle$ . If  $v$  has a child, let  $K$  be such that the subtree of  $M^\bullet$  starting in the node  $v \cdot 1$  equals  $K^\bullet$  (checking the definition of  $M^\bullet$ , where  $M$  is first-order, we see that all subtrees of  $M^\bullet$  are of this form). We say that  $\langle M^\bullet, d, v \rangle$  is *valid* if either

- $d = \Downarrow$ ,
- $d = \Uparrow_x$ ,  $v$  has a child, and  $x$  is free in  $K$ , or
- $d = \Uparrow_i$ ,  $v$  has a child, and  $K$  requires at least  $i$  arguments (i.e.,  $K$  has type  $\mathbf{o}^k \rightarrow \mathbf{o}$  with  $k \geq i$ ).

We have the following lemma:

**Lemma A.23.** All configurations  $\langle M^\bullet, d, v \rangle$  reached while computing  $\llbracket M^\bullet \rrbracket_{\mathcal{X},s}$  are valid.

**Proof:**

By a case-by-case analysis of the definition of  $(\mathcal{X}, s)$ -successor, we immediately see that if  $c \rightarrow_{\mathcal{X},s} b$  and  $c$  is valid, then  $b$  is valid as well. Additionally, if a configuration  $\langle M^\bullet, \Downarrow, v \rangle$  is valid and the node  $v$  is labelled by  $\bar{a}$ , then the configurations  $\langle M^\bullet, \Uparrow_i, v \rangle$  for  $i \in \{1, \dots, r\}$ , where  $r$  is the rank of  $a$ , are valid as well.  $\square$

Let “ $\sqsupseteq$ ” be a binary relation between valid configurations. We say that “ $\sqsupseteq$ ” is a *weak simulation* if, whenever  $c \sqsupseteq b$  holds for two valid configurations  $b, c$ , we then have

1. if  $b \rightarrow_{\mathcal{X},s} b'$ , then there exists a valid configuration  $c'$  such that  $c \rightarrow_{\mathcal{X},s} c'$  and  $c' \sqsupseteq b'$ , and
2. if  $b = \langle N^\bullet, \Downarrow, v \rangle$  and  $v$  has label  $\bar{a}$ , then  $c = \langle M^\bullet, \Downarrow, u \rangle$  and  $u$  has the same label  $\bar{a}$ , and  $\langle N^\bullet, \Uparrow_i, v \rangle \sqsupseteq \langle M^\bullet, \Uparrow_i, u \rangle$  for all  $i \in \{1, \dots, r\}$ , where  $r$  is the rank of  $a$ .

The following lemma shows that weak simulation preserves derived trees:

**Lemma A.24.** If “ $\sqsupseteq$ ” is a weak simulation, and  $\llbracket b \rrbracket_{\mathcal{X},s}$  does not contain the special letter  $\perp$ , then

$$c \sqsupseteq b \quad \text{implies} \quad \llbracket c \rrbracket_{\mathcal{X},s} = \llbracket b \rrbracket_{\mathcal{X},s}.$$

**Proof:**

We proceed by coinduction on derived trees. Let  $c = \langle M^\bullet, d, u \rangle$  and  $b = \langle N^\bullet, e, v \rangle$ . By the definition of the derived tree  $\llbracket N^\bullet, e, v \rrbracket_{\mathcal{X},s}$  there is a maximal sequence of successors

$$\langle N^\bullet, e, v \rangle \rightarrow_{\mathcal{X},s}^* \langle N^\bullet, \Downarrow, v' \rangle,$$

where the node  $v'$  in  $N^\bullet$  is labelled with  $\bar{a}$ , and such that

$$\llbracket N^\bullet, e, v \rrbracket_{\mathcal{X},s} = a \llbracket N^\bullet, \Uparrow_1, v' \rrbracket_{\mathcal{X},s} \cdots \llbracket N^\bullet, \Uparrow_r, v' \rrbracket_{\mathcal{X},s},$$

where  $r$  is the rank of the letter  $a$ . By assumption  $\langle M^\bullet, d, u \rangle \sqsupseteq \langle N^\bullet, e, v \rangle$ . Recall that “ $\sqsupseteq$ ” is a weak simulation, thus a repeated application of the first item in the definition of a weak simulation shows

$$\langle M^\bullet, d, u \rangle \rightarrow_{\mathcal{X},s}^* \langle M^\bullet, d', u' \rangle$$

with  $\langle M^\bullet, d', u' \rangle \sqsupseteq \langle N^\bullet, \Downarrow, v' \rangle$ . By the second item in the definition of a weak simulation we have that  $d' = \Downarrow$  and  $u'$  is labeled by  $\bar{a}$ ; in particular  $\langle M^\bullet, d', u' \rangle$  does not have a successor (cf. the definition of a successor). By the definition of a derived tree we thus have

$$\llbracket M^\bullet, d, u \rrbracket_{\mathcal{X},s} = a \llbracket M^\bullet, \Uparrow_1, u' \rrbracket_{\mathcal{X},s} \cdots \llbracket M^\bullet, \Uparrow_r, u' \rrbracket_{\mathcal{X},s}.$$

This shows that the derived trees of  $\langle M^\bullet, d, u \rangle$  and  $\langle N^\bullet, e, v \rangle$  agree on the label of their root. Moreover, the second item in the definition of a weak simulation also says that

$$\langle M^\bullet, \Uparrow_i, u' \rangle \sqsupseteq \langle N^\bullet, \Uparrow_i, v' \rangle \quad \text{for all } i \in \{1, \dots, r\}.$$

By coinduction on derived trees we thus have

$$\llbracket M^\bullet, \Uparrow_i, u' \rrbracket_{\mathcal{X},s} = \llbracket N^\bullet, \Uparrow_i, v' \rrbracket_{\mathcal{X},s} \quad \text{for all } i \in \{1, \dots, r\}.$$

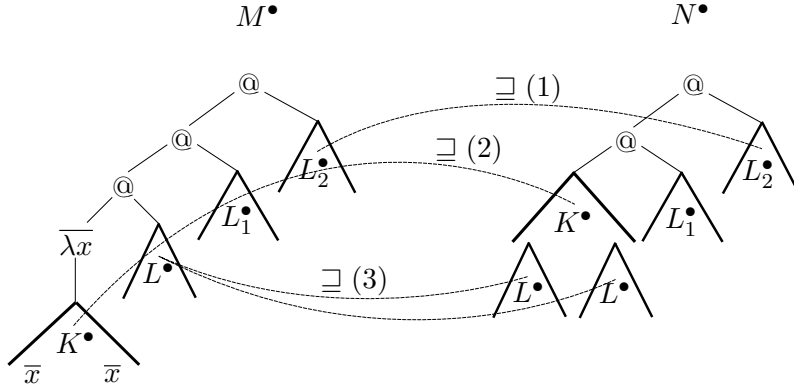
This shows that all relevant subtrees also agree, thus concluding the proof.  $\square$

Recall that our goal is to prove Lemma A.22, saying that derived trees are invariant under head beta-reductions. Fix thus two closed first-order input lambda-terms  $M, N$  of type  $\mathbf{o}$ , such that  $M \rightarrow_{h\beta} N$ . Then

$$M = (\lambda x.K) L \ L_1 \ \dots \ L_j \quad \text{and} \quad N = K[L/x] \ L_1 \ \dots \ L_j.$$

We now define a *concrete* weak simulation, denoted by overloading the same symbol “ $\sqsupseteq$ ”, between valid configurations involving  $M^\bullet$  and  $N^\bullet$ . Define  $\langle M^\bullet, d, u \rangle \sqsupseteq \langle N^\bullet, d, v \rangle$  if either



Figure 5. An illustration of the weak simulation relation “ $\sqsubseteq$ ”

1.  $v$  is not of the form  $1^j \cdot v'$  and  $u = v$  (i.e.,  $v$  is outside of  $(K[L/x])^\bullet$ , and  $u$  is the same node in the other lambda-tree),
2.  $v = 1^j \cdot v'$ , and  $u = 1^{j+2} \cdot v'$ , and  $u$  is not labeled by  $\bar{x}$  in  $M^\bullet$  (i.e.,  $v$  is inside the “ $K^\bullet$  part” of  $(K[L/x])^\bullet$ , and  $u$  is the corresponding node of  $K^\bullet$  in  $@(\bar{\lambda}x K^\bullet) L^\bullet$ ), or
3.  $v$  can be written as  $v = 1^j \cdot v' \cdot v''$ , where  $1^{j+2} \cdot v'$  has label  $\bar{x}$  in  $M^\bullet$ , and  $u = 1^j \cdot 2 \cdot v''$  (i.e.,  $v$  is inside some  $L^\bullet$  in  $(K[L/x])^\bullet$ , and  $u$  is the corresponding node of  $L^\bullet$  in  $@(\bar{\lambda}x K^\bullet) L^\bullet$ ).

See Figure 5 for an illustration of the definition above for  $j = 2$ . As a special case of the last condition, the root of  $L^\bullet$  in  $M^\bullet$  is in relation with the root of some copy of  $L^\bullet$  in  $N^\bullet$ . Note that  $\langle M^\bullet, d, u \rangle \sqsubseteq \langle N^\bullet, d, v \rangle$  holds only for configurations with the same direction  $d$ . Note also that for every node  $v$  of  $N^\bullet$  we can find a (unique) corresponding node  $u$  in  $M^\bullet$ , but it is not the case that for every node  $u$  of  $M^\bullet$  there is a corresponding node  $v$  in  $N^\bullet$ . In particular, in “ $\sqsubseteq$ ” we do not have any pairs with  $u = 1^j$  nor  $u = 1^{j+1}$  (i.e., with  $u$  pointing to the “ $@$ ” or “ $\bar{\lambda}x$ ” at the top of  $@(\bar{\lambda}x K^\bullet) L^\bullet$ ); also, nodes labelled with  $\bar{x}$  in  $K^\bullet$  from  $M^\bullet$  (if any) are not in relation with any node of  $N^\bullet$ ; finally, if  $x$  does not occur in  $K$ , then additionally no node in  $L^\bullet$  is in relation with a node in  $N^\bullet$ .

**Lemma A.25.** The binary relation “ $\sqsubseteq$ ” is a weak simulation.

Before proving Lemma A.25, let us see how Lemma A.22 follows from it:

### Proof of Lemma A.22:

Recall that  $\llbracket N^\bullet \rrbracket_{\mathcal{X},s} = \llbracket b \rrbracket_{\mathcal{X},s}$  with configuration  $b = \langle N^\bullet, \Downarrow, v \rangle$ , where  $v = \varepsilon$  is the root of the lambda-tree  $N^\bullet$ . In order to ensure  $c \sqsubseteq b$ , we have to be a bit careful while choosing the configuration  $c$  for  $M^\bullet$ . We take  $c = \langle M^\bullet, \Downarrow, u \rangle$ , where  $u$  is chosen as follows:

- If  $j \geq 1$ , we can simply take  $u = \varepsilon$ .
- If  $j = 0$  and  $K \neq x$ , we rather take  $u = 1 \cdot 1$  (which is the node where  $K^\bullet$  starts). Note that  $\langle M^\bullet, \Downarrow, \varepsilon \rangle \rightarrow_{\mathcal{X},s}^* c$  in two steps.
- Finally, if  $j = 0$  and  $K = x$ , we take  $u = 2$  (which is the node where  $L^\bullet$  starts). This time we also have  $\langle M^\bullet, \Downarrow, \varepsilon \rangle \rightarrow_{\mathcal{X},s}^* c$ .

Thus, in any case  $\llbracket M^\bullet \rrbracket_{\mathcal{X},s} = \llbracket c \rrbracket_{\mathcal{X},s}$ . Moreover, we have  $c \sqsubseteq b$  by definition. By Lemma A.24 we have  $\llbracket b \rrbracket_{\mathcal{X},s} = \llbracket c \rrbracket_{\mathcal{X},s}$ , as required.  $\square$

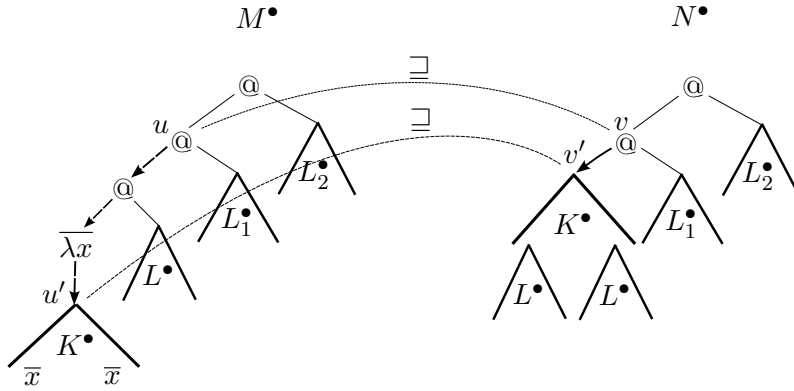


Figure 6. Illustration of Case 1 in the proof of Lemma A.25

What remains is to prove Lemma A.25:

### Proof of Lemma A.25:

Consider two valid configurations  $c = \langle M^\bullet, d, u \rangle$  and  $b = \langle N^\bullet, d, v \rangle$  such that  $c \sqsubseteq b$ . Observe first that then necessarily  $u$  and  $v$  have the same label, and that  $\langle M^\bullet, e, u \rangle \sqsubseteq \langle N^\bullet, e, v \rangle$  for every other direction  $e$  for which the configurations are valid. This immediately implies the second item in the definition of a weak simulation.

Let us check the first item. To this end, consider  $b' = \langle N^\bullet, e, v' \rangle$  such that  $b \rightarrow_{\mathcal{X},s} b'$ ; we have to find  $c'$  such that  $c \rightarrow_{\mathcal{X},s}^* c'$  and  $c' \sqsubseteq b'$ . By the definition of a successor,  $v'$  is either a child of  $v$ , or a parent of  $v$ . A natural candidate for  $c'$  is the unique  $(\mathcal{X}, s)$ -successor of  $c$ . Because  $c$  and  $b$  have the same direction and the same node label, the successor indeed exists, and it is of the form  $\langle M^\bullet, e, u' \rangle$ , where  $u'$  is in the same relation to  $u$  as  $v'$  to  $v$  (i.e.,  $u' = \text{par}(u)$  if  $v' = \text{par}(v)$ , and  $u' = \text{ch}_i(u)$  if  $v' = \text{ch}_i(v)$ ). If  $v$  and  $v'$  are in the same “part” of  $N^\bullet$ , that is, both in  $L^\bullet$ , both in  $(K[L/x])^\bullet$  but outside of  $L^\bullet$ , or both outside of  $(K[L/x])^\bullet$ , then we have  $c' \sqsubseteq b'$ , and we are done.

The situation is more complicated only when  $v$  and  $v'$  are in different parts. Let us first consider the border of  $(K[L/x])^\bullet$ :

1. Suppose that  $v = 1^{j-1}$ ,  $v' = 1^j$  (recall that  $j$  is the number of arguments following the redex, i.e.,  $N = K[L/x] L_1 \dots L_j$ , so  $v'$  is the root of  $(K[L/x])^\bullet$ , and  $v$  its parent). Then  $u = 1^{j-1}$  and  $e = \Downarrow$ . If  $K \not\equiv x$ , in  $M^\bullet$  we can make three successor steps, going through the nodes labeled by  $@$  and  $\lambda x$  to the root of  $K^\bullet$  in  $@(\lambda x K^\bullet) L^\bullet$ ; for  $c' = \langle M^\bullet, \Downarrow, 1^{j+2} \rangle$  we have  $c' \sqsubseteq b'$ ; see Figure 6, where the thick arrow on the right is simulated by the three dashed arrows on the left, denoting successor steps.

If  $K = x$ , we need three more successor steps: from the  $\bar{x}$ -labelled root of  $K^\bullet$  we go up to the  $\lambda x$ -labelled node with direction  $\Uparrow_x$ , then to the  $@$ -labelled node with direction  $\Uparrow_1$ , and finally we go down to the root of  $L^\bullet$  with direction  $\Downarrow$ ; for  $c' = \langle M^\bullet, \Downarrow, 1^j \cdot 2 \rangle$  we have  $c' \sqsubseteq b'$ .

2. Suppose that  $v = 1^j$ ,  $v' = 1^{j-1}$ , and  $e = \Uparrow_i$ . Then  $u = 1^{j+2}$  is the root of  $K^\bullet$  in  $@(\lambda x K^\bullet) L^\bullet$ . Note that  $K[L/x]$  has type  $\circ^j \rightarrow \circ$ , so because  $b'$  is valid, we have  $i \leq j$ . It is important that  $\lambda x.K$  is also a subterm of  $M$  and has type  $\circ^{j+1} \rightarrow \circ$ ; because  $M$  is an input lambda-term, we have  $j + 1 \leq \text{mar}(M) \leq \text{mar}(\mathcal{G}) = s$ . We can thus make three successor steps in  $M^\bullet$

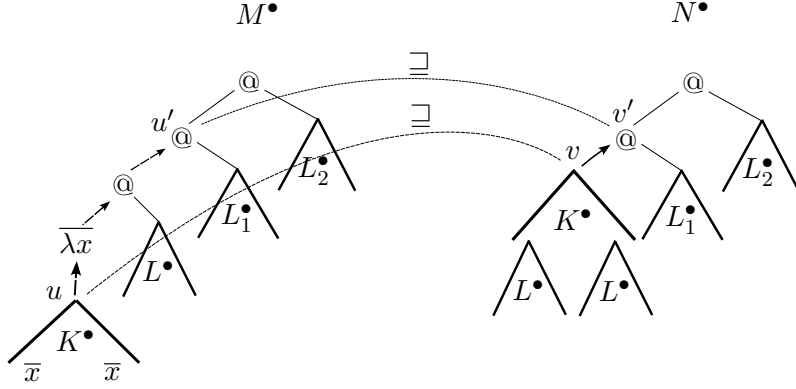


Figure 7. Illustration of Cases 2 and 3 in the proof of Lemma A.25

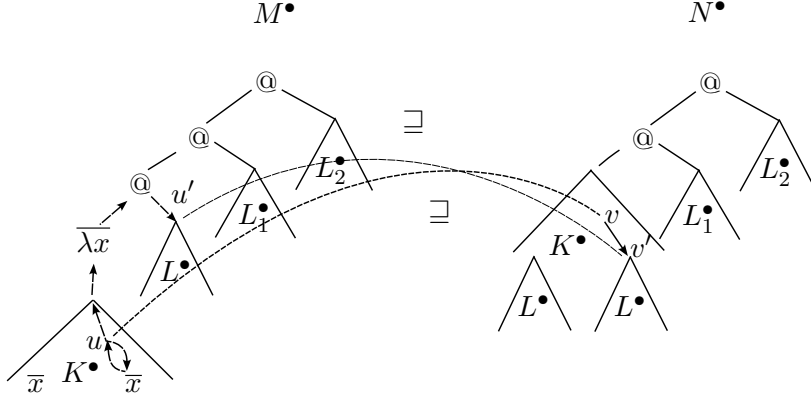


Figure 8. Illustration of the last case in the proof of Lemma A.25

(c.f. Figure 7), going up through the nodes labeled by  $\overline{\lambda x}$  and  $@$ :

$$\langle M^\bullet, d, 1^{j+2} \rangle \rightarrow_{\mathcal{X},s} \langle M^\bullet, \uparrow_i, 1^{j+1} \rangle \rightarrow_{\mathcal{X},s} \langle M^\bullet, \uparrow_{i+1}, 1^j \rangle \rightarrow_{\mathcal{X},s} \langle M^\bullet, \uparrow_i, 1^{j-1} \rangle = c'.$$

3. Finally, suppose that  $v = 1^j$ ,  $v' = 1^{j-1}$ , and  $e$  is not of the form  $\uparrow_i$ . Then necessarily  $e = \uparrow_y$  and again  $u = 1^{j+2}$ . Recall that  $M$  is closed, implying that  $L$  is closed. Thus  $x$  is not free in  $K[L/x]$ , so  $y \neq x$  because  $b'$  is valid. This allows us to make three successor steps in  $M^\bullet$  (c.f. Figure 7), going up through the nodes labeled by  $\overline{\lambda x}$  and  $@$ ; we take  $c' = \langle M^\bullet, \uparrow_y, 1^{j-1} \rangle$ .

Next, we consider the border of  $L^\bullet$ . Recall that  $M$  is a first-order lambda-term, implying that  $L$  is of type o. Thus, because  $b'$  is valid, we can never leave  $L^\bullet$  with direction  $\uparrow_i$ . Likewise, because  $L$  is closed and  $b'$  is valid, we can never leave  $L^\bullet$  with direction  $\uparrow_y$ , for any variable  $y$ . It remains to consider the case when we enter  $L$  from above. This means that  $v' = 1^j \cdot w$  is the root of some copy of  $L^\bullet$  in  $N^\bullet$ , while the node  $1^{j+2} \cdot w$  in  $M^\bullet$  is labelled by  $\overline{x}$ . We then have  $e = \downarrow$ . The case of  $K = x$  is already covered by Item 1 above; we may assume that  $K \neq x$ . Then  $u$  is the parent of the  $\overline{x}$ -labelled node  $1^{j+2} \cdot w$ , and the successor of  $c$  is  $\langle M^\bullet, \downarrow, 1^{j+2} \cdot w \rangle$ . Although  $x$  may occur in some lambda-binders in  $K$ , we know that for the considered occurrence of  $x$  we have substituted  $L$ , so it is not under the scope of  $\lambda x$  inside  $K$  (i.e., no ancestor of the node  $1^{j+2} \cdot w$  inside  $K^\bullet$  is labelled

by  $\overline{\lambda x}$ ). Thus the sequence of successors from  $\langle M^\bullet, \Downarrow, 1^{j+2} \cdot w \rangle$  goes up with direction  $\Uparrow_x$  until it reaches the  $\overline{\lambda x}$ -labelled node  $1^{j+1}$ . The successor of  $\langle M^\bullet, \Uparrow_x, 1^{j+1} \rangle$  is  $\langle M^\bullet, \Uparrow_1, 1^j \rangle$ , and its successor is  $\langle M^\bullet, \Downarrow, 1^j \cdot 2 \rangle$  (whose node is the root of  $L^\bullet$ ; c.f. Figure 8); taking this configuration as  $c'$ , we have  $c' \sqsupseteq b'$ , as required.  $\square$