

Intersection Types and Counting*

Paweł Parys

University of Warsaw, Poland

parys@mimuw.edu.pl

We present a new approach to the following meta-problem: given a quantitative property of trees, design a type system such that the desired property for the tree generated by an infinitary ground λ -term corresponds to some property of a derivation of a type for this λ -term, in this type system.

Our approach is presented in the particular case of the language finiteness problem for nondeterministic higher-order recursion schemes (HORSes): given a nondeterministic HORS, decide whether the set of all finite trees generated by this HORS is finite. We give a type system such that the HORS can generate a tree of an arbitrarily large finite size if and only if in the type system we can obtain derivations that are arbitrarily large, in an appropriate sense; the latter condition can be easily decided.

1 Introduction

In this paper we consider λY -calculus, which is an extension of the simply typed λ -calculus by a fixed-point operator Y . A term P of λY -calculus that is of sort¹ o can be used to generate an infinite tree $BT(P)$, called the Böhm tree of P . Trees generated by terms of λY -calculus can be used to faithfully represent the control flow of programs in languages with higher-order functions. Traditionally, Higher Order Recursive Schemes (HORSes) are used for this purpose [8, 12, 17, 16]; this formalism is equivalent to λY -calculus, and the translation between them is rather straightforward [21]. Collapsible Pushdown Systems [10] and Ordered Tree-Pushdown Systems [7] are other equivalent formalisms.

Intersection type systems were intensively used in the context of HORSes, for several purposes like model-checking [13, 16, 5, 20], pumping [14], transformations of HORSes [15, 6], etc. Interestingly, constructions very similar to intersection types were used also on the side of collapsible pushdown systems; they were alternating stack automata [4], and types of stacks [19, 11].

In this paper we show how intersection types can be used for deciding quantitative properties of trees generated by λY -terms. We concentrate on the language finiteness problem for nondeterministic HORSes: given a nondeterministic HORS, decide whether the set of all finite trees generated by this HORS is finite.

This problem can be restated in the world of λY -terms (or standard, deterministic HORSes), generating a single infinite tree. Here, instead of resolving nondeterministic choices during the generation process, we leave them in the resulting tree. Those nondeterministic choices are denoted by a distinguished br (“branch”) symbol, below which we put options that could be chosen. Then to obtain a finite tree generated by the original HORS we just need to recursively choose in every br -labeled node which of the two subtrees we want to consider. Thus, in this setting, the language finiteness problem asks whether the set of all finite trees obtained this way is finite.

The difficulty of this problem lies in the fact that sometimes the same finite tree may be found in infinitely many different places of $BT(P)$ (i.e., generated by a nondeterministic HORS in many ways);

*Work supported by the National Science Center (decision DEC-2012/07/D/ST6/02443).

¹We use the word “sort” instead of the usual “type” to avoid confusion with intersection types introduced in this paper.

thus the actual property to decide is whether there is a common bound on the size of each of these trees. This makes the problem inaccessible for standard methods used for analyzing HORSes, as they usually concern only regular properties of the Böhm tree, while boundedness is a problem of different kind. The same difficulty was observed in [14], where they prove a pumping lemma for deterministic HORSes, while admitting (Remark 2.2) that their method is too weak to reason about nondeterministic HORSes.

In order to solve the language finiteness problem, we present an appropriate intersection type system, where derivations are annotated by flags and markers of multiple kinds. The key property of this type system is that the number of flags in a type derivation for a λY -term P approximates the size of some finite tree obtained by resolving nondeterministic choices in the infinite tree $BT(P)$. In consequence, there are type derivations using arbitrarily many flags if, and only if, the answer to the language finiteness problem is “no”.

The language finiteness problem was first attacked in [1] (for safe HORSes only), but their algorithm turned out to be incorrect [2]. To our knowledge, the only known solution of this problem follows from a recent decidability result for the diagonal problem [9, 6]. This problem asks, given a nondeterministic HORS and a set of letters Σ , whether for every $n \in \mathbb{N}$ the HORS generates a finite tree in which every letter from Σ appears at least n times. Clearly, a nondeterministic HORS generates arbitrarily large trees exactly when for some letter a it generates trees having arbitrarily many a letters, i.e., when the answer to the diagonal problem for $\Sigma = \{a\}$ is “yes”.

Our type system is, to some extent, motivated by the algorithm of [6] solving the diagonal problem. This algorithm works by repeating two kinds of transformations of HORSes. The first of them turns the HORS into a HORS generating trees having only a fixed number of branches, one per each letter from Σ (i.e., one branch in our case of $|\Sigma| = 1$). The branches are chosen nondeterministically out of some tree generated by the original HORS; for every $a \in \Sigma$ there is a choice witnessing that a appeared many times in the original tree. Then such a HORS of the special form is turned into a HORS that is of order lower by one, and generates trees having the same nodes as trees generated by the original HORS, but arranged differently (in particular, the new trees may have again arbitrarily many branches). After finitely many repetitions of this procedure, a HORS of order 0 is obtained, and the diagonal problem becomes easily decidable. In some sense we want to do the same, but instead of applying all these transformations one by one, we simulate all of them simultaneously in a single type derivation. In this derivation, for each order n , we allow to place arbitrarily one marker “of order n ”; this corresponds to the nondeterministic choice of one branch in the n -th step of the previous algorithm. We also place some flags “of order n ”, in places that correspond to nodes remaining after the n -th step of the previous algorithm.

The idea of using intersection types for counting is not completely new. Paper [18] presents a type system that, essentially, allows to estimate the size of the β -normal form of a λ -term just by looking at (the number of some flags in) a derivation of a type for this term. A similar idea, but for higher-order pushdown automata, is present in [19], where we can estimate the number of \sharp symbols appearing on a particular, deterministically chosen branch of the generated tree. This previous approach also uses intersection types, where the derivations are marked with just one kind of flags, denoting “productive” places of a λ -term (oppositely to our approach, where we have different flags for different orders, and we also have markers). The trouble with the “one-flag” approach is that it works well only in a completely deterministic setting, where looking independently at each node of the Böhm tree we know how it contributes to the result; the method stops working (or at least we do not know how to prove that it works) in our situation, where we first nondeterministically perform some guesses in the Böhm tree, and only after that we want to count something that depends on the chosen values.

Acknowledgements. I would like to thank Szymon Toruńczyk for stimulating discussions, and anonymous reviewers for useful comments.

2 Preliminaries

Trees. Let Σ be a *ranked alphabet*, i.e., a set of symbols together with a rank function assigning a nonnegative integer to each of the symbols. We assume that Σ contains a distinguished symbol br of rank 2, used to denote nondeterministic choices. A Σ -*labeled tree* is a tree that is rooted (there is a distinguished root node), node-labeled (every node has a label from Σ), ranked (a node with label of rank n has exactly n children), and ordered (children of a node of rank n are numbered from 1 to n).

When t is a Σ -labeled tree t , by $\mathcal{L}(t)$ we denote the set of all finite trees that can be obtained by choosing in every br -labeled node of t which of the two subtrees we want to consider. More formally, we consider the following relation \rightarrow_{br} : we have $t \rightarrow_{\text{br}} u$ if u can be obtained from t by choosing in t a br -labeled node x and its child y , and replacing the subtree starting in x by the subtree starting in y (which removes x and the other subtree of x). Let $\rightarrow_{\text{br}}^*$ be the reflexive transitive closure of \rightarrow_{br} . Then $\mathcal{L}(t)$ contains all trees u that do not use the br label, are finite, and such that $t \rightarrow_{\text{br}}^* u$.

Infinitary λ -calculus. The set of *sorts* (a.k.a. simple types), constructed from a unique basic sort o using a binary operation \rightarrow , is defined as usual. The order of a sort is defined by: $\text{ord}(o) = 0$, and $\text{ord}(\alpha \rightarrow \beta) = \max(1 + \text{ord}(\alpha), \text{ord}(\beta))$.

We consider infinitary, sorted λ -calculus. *Infinitary λ -terms* (or just λ -terms) are defined by coinduction, according to the following rules:

- if $a \in \Sigma$ is a symbol of rank r , and P_1^o, \dots, P_r^o are λ -terms, then $(aP_1^o \dots P_r^o)^o$ is a λ -term,
- for every sort α there are infinitely many variables $x^\alpha, y^\alpha, z^\alpha, \dots$; each of them is a λ -term,
- if $P^{\alpha \rightarrow \beta}$ and Q^α are λ -terms, then $(P^{\alpha \rightarrow \beta} Q^\alpha)^\beta$ is a λ -term, and
- if P^β is a λ -term and x^α is a variable, then $(\lambda x^\alpha. P^\beta)^{\alpha \rightarrow \beta}$ is a λ -term.

We naturally identify λ -terms differing only in names of bound variables. We often omit the sort annotations of λ -terms, but we keep in mind that every λ -term (and every variable) has a particular sort. A λ -term P is *closed* if it has no free variables. Notice that, for technical convenience, a symbol of positive rank is not a λ -term itself, but always comes with arguments. This is not a restriction, since e.g. instead of a unary symbol a one may use the term $\lambda x. ax$.

The order of a λ -term is just the order of its sort. The *complexity* of a λ -term P is the smallest number m such that the order of every subterm of P is at most m . We restrict ourselves to λ -terms that have finite complexity.

A β -reduction is defined as usual. We say that a β -reduction $P \rightarrow_\beta Q$ is of *order* n if it concerns a redex $(\lambda x. R)S$ such that $\text{ord}(\lambda x. R) = n$. In this situation the order of x is at most $n - 1$, but may be smaller (when other arguments of R are of order $n - 1$).

Böhm Trees. We consider Böhm trees only for closed λ -terms of sort o . For such a term P , its *Böhm tree* $BT(P)$ is constructed by coinduction, as follows: if there is a sequence of β -reductions from P to a λ -term of the form $aP_1 \dots P_r$ (where a is a symbol), then the root of the tree t has label a and r children, and the subtree starting in the i -th child is $BT(P_i)$. If there is no sequence of β -reductions from P to a

λ -term of the above form, then $BT(P)$ is the full binary tree with all nodes labeled by br .² By $\mathcal{L}(P)$ we denote $\mathcal{L}(BT(P))$.

λY -calculus. The syntax of λY -calculus is the same as that of finite λ -calculus, extended by symbols $Y^{(\alpha \rightarrow \alpha) \rightarrow \alpha}$, for each sort α . A term of λY -calculus is seen as a term of infinitary λ -calculus if we replace each symbol $Y^{(\alpha \rightarrow \alpha) \rightarrow \alpha}$ by the unique infinite λ -term Z such that Z is syntactically the same as $\lambda x^{\alpha \rightarrow \alpha}.x(Zx)$. In this way, we view λY -calculus as a fragment of infinitary λ -calculus.

It is standard to convert a nondeterministic HORS \mathcal{G} into a closed λY -term P^o such that $\mathcal{L}(P)$ is exactly the set of all finite trees generated by \mathcal{G} . The following theorem, which is our main result, states that the *language finiteness problem* is decidable.

Theorem 1. *Given a closed λY -term P of sort o , one can decide whether $\mathcal{L}(P)$ is finite.*

3 Intersection Type System

In this section we introduce a type system that allows to determine the desired property: whether in $\mathcal{L}(P)$ there is an arbitrarily large tree.

Intuitions. The main novelty of our type system is in using flags and markers, which may label nodes of derivation trees. To every flag and marker we assign a number, called an order. While deriving a type for a λ -term of complexity m , we may place in every derivation tree at most one marker of each order $n \in \{0, \dots, m-1\}$, and arbitrarily many flags of each order $n \in \{0, \dots, m\}$.

Consider first a λ -term M_0 of complexity 0. Such a term actually equals its Böhm tree. Our aim is to describe some finite tree t in $\mathcal{L}(M_0)$, i.e., obtained from M_0 by resolving nondeterministic choices in some way. We thus just put flags of order 0 in all those (appearances of) symbols in M_0 that contribute to this tree t ; the type system ensures that indeed all symbols of some finite tree in $\mathcal{L}(M_0)$ are labeled by a flag. Then clearly we have the desired property that there is a derivation with arbitrarily many flags if, and only if, there are arbitrarily large trees in $\mathcal{L}(M_0)$.

Next, consider a λ -term M_1 that is of complexity 1, and reduces to M_0 . Of course every finite tree from $\mathcal{L}(M_0)$ is composed of symbols appearing already in M_1 ; we can thus already in M_1 label (by order-0 flags) all symbols that contribute to some tree $t \in \mathcal{L}(M_0)$ (and an intersection type system can easily check correctness of such labeling). There is, however, one problem: a single appearance of a symbol in M_1 may result in many appearances in M_0 (since a function may use its argument many times). Due to this, the number of order-0 flags in M_1 does not correspond to the size of t . We rescue ourselves in the following way. In t we choose one leaf, we label it by an order-0 marker, and on the path leading from the root to this marker we place order-1 flags. On the one hand, $\mathcal{L}(M_0)$ contains arbitrarily large trees if, and only if, it contains trees with arbitrarily long paths, i.e., trees with arbitrarily many order-1 flags. On the other hand, we can perform the whole labeling (and the type system can check its correctness) already in M_1 , and the number of order-1 flags in M_1 will be precisely the same as it would be in M_0 . Indeed, in M_1 we have only order-1 functions, i.e., functions that take trees and use them as subtrees of larger trees; although a tree coming as an argument may be duplicated, the order-0 marker can be placed in at most one copy. This means that, while reducing M_1 to M_0 , every symbol of M_1 can result in at most

²Usually one uses a special label \perp of rank 0 for this purpose, but from the perspective of our problem both definitions are equivalent.

one symbol of M_0 lying on the selected path to the order-0 marker (beside of arbitrarily many symbols outside of this path).

This procedure can be repeated for M_2 of complexity 2 that reduces to M_1 via β -reductions of order 2 (and so on for higher orders). We now place a marker of order 1 in some leaf of M_1 ; afterwards, we place an order-2 flag in every node that is on the path to the marked leaf, and that has a child outside of this path whose some descendant is labeled by an order-1 flag. In effect, for some choice of a leaf to be marked, the number of order-2 flags approximates the number of order-1 flags, up to logarithm. Moreover, the whole labeling can be done in M_2 instead of in M_1 , without changing the number of order-2 flags.

In this intuitive description we have talked about labeling “nodes of a λ -term”, but formally we label nodes of a derivation tree that derives a type for the term, in our type system. Every such node contains a type judgment for some subterm of the term.

Type Judgments. For every sort α we define the set \mathcal{T}^α of *types* of sort α , and the set \mathcal{F}^α of *full types* of sort α . This is done as follows, where \mathcal{P} denotes the powerset:

$$\begin{aligned} \mathcal{T}^{\alpha \rightarrow \beta} &= \mathcal{P}(\mathcal{F}_{ord(\alpha \rightarrow \beta)}^\alpha) \times \mathcal{T}^\beta, & \mathcal{T}^o &= o, \\ \mathcal{F}_k^\alpha &= \{(k, F, M, \tau) \mid F, M \subseteq \{0, \dots, k-1\}, F \cap M = \emptyset, \tau \in \mathcal{T}^\alpha\}, & \mathcal{F}^\alpha &= \bigcup_{k \in \mathbb{N}} \mathcal{F}_k^\alpha. \end{aligned}$$

Notice that the sets \mathcal{T}^α and \mathcal{F}_k^α are finite (unlike \mathcal{F}^α). A type $(T, \tau) \in \mathcal{T}^{\alpha \rightarrow \beta}$ is denoted as $T \rightarrow \tau$. A full type $\hat{\tau} = (k, F, M, \tau) \in \mathcal{F}_k^\alpha$ consists of its order k , a set F of flag orders, a set M of marker orders, and a type τ ; we write $ord(\hat{\tau}) = k$. In order to distinguish types from full types, the latter are denoted by letters with a hat, like $\hat{\tau}$.

A *type judgment* is of the form $\Gamma \vdash P : \hat{\tau} \triangleright c$, where Γ , called a *type environment*, is a function that maps every variable x^α to a subset of \mathcal{F}^α , P is a λ -term, $\hat{\tau}$ is a full type of the same sort as P (i.e., $\hat{\tau} \in \mathcal{F}^\beta$ when P is of sort β), and $c \in \mathbb{N}$.

As usual for intersection types, the intuitive meaning of a type $T \rightarrow \tau$ is that a λ -term having this type can return a λ -term having type τ , while taking an argument for which we can derive all full types from T . Moreover, in \mathcal{T}^o there is just one type o , which can be assigned to every λ -term of sort o . Suppose that we have derived a type judgment $\Gamma \vdash P : \hat{\tau} \triangleright c$ with $\hat{\tau} = (m, F, M, \tau)$. Then

- τ is the type derived for P ;
- Γ contains full types that could be used for free variables of P in the derivation;
- m bounds the order of flags and markers that could be used in the derivation: flags could be of order at most m , and markers of order at most $m-1$;
- $M \subseteq \{0, \dots, m-1\}$ contains the orders of markers used in the derivation, together with those provided by free variables (i.e., we imagine that some derivations, specified by the type environment, are already substituted in our derivation for free variables); we, however, do not include markers provided by arguments of the term (i.e., coming from the sets T_i when $\tau = T_1 \rightarrow \dots \rightarrow T_k \rightarrow o$);
- F contains those numbers $n \in \{0, \dots, m-1\}$ (excluding $n = m$) for which a flag of order n is placed in the derivation itself, or provided by a free variable, or provided by an argument; for technical convenience we, however, remove n from F whenever $n \in M$ (when $n \in M$, the information about order- n flags results in placing an order- $(n+1)$ flag, and need not to be further propagated);
- c , called a *flag counter*, counts the number of order- m flags present in the derivation.

Type System. Before giving rules of the type system, we need a few definitions. We use the symbol \uplus to denote disjoint union. When $A \subseteq \mathbb{N}$ and $n \in \mathbb{N}$, we write $A \upharpoonright_{<n}$ for $\{k \in A \mid k < n\}$, and similarly $A \upharpoonright_{\geq n}$ for $\{k \in A \mid k \geq n\}$. By ε we denote the type environment mapping every variable to \emptyset , and by $\Gamma[x \mapsto T]$ the type environment mapping x to T and every other variable y to $\Gamma(y)$.

Let us now say how a type environment Γ from the conclusion of a rule may be split into type environments $(\Gamma_i)_{i \in I}$ used in premisses of the rule: we say that $\text{Split}(\Gamma \mid (\Gamma_i)_{i \in I})$ holds if and only if for every variable x it holds $\Gamma_i(x) \subseteq \Gamma(x)$ for every $i \in I$, and every full type from $\Gamma(x)$ providing some markers (i.e., (k, F, M, τ) with $M \neq \emptyset$) appears in some $\Gamma_i(x)$. Full types with empty M may be discarded and duplicated freely. This definition forbids to discard full types with nonempty M , and from elsewhere it will follow that they cannot be duplicated. As a special case $\text{Split}(\Gamma \mid \Gamma')$ describes how a type environment can be weakened.

All type derivations are assumed to be finite (although we derive types mostly for infinite λ -terms, each type derivation analyzes only a finite part of a term). Rules of the type system will guarantee that the order m of derived full types will be the same in the whole derivation (although in type environments there may be full types of different orders).

We are ready to give the first three rules of our type system:

$$\frac{\Gamma \vdash P_1 : \hat{\tau} \triangleright c \quad i \in \{1, 2\}}{\Gamma \vdash \text{br } P_1 P_2 : \hat{\tau} \triangleright c} \text{ (BR)} \quad \frac{\text{Split}(\Gamma \mid \varepsilon[x \mapsto \{(k, F, M', \tau)\}]) \quad M \upharpoonright_{<k} = M'}{\Gamma \vdash x : (m, F, M, \tau) \triangleright 0} \text{ (VAR)}$$

$$\frac{\Gamma'[x \mapsto T] \vdash P : (m, F, M, \tau) \triangleright c \quad \text{Split}(\Gamma \mid \Gamma') \quad \Gamma'(x) = \emptyset}{\Gamma \vdash \lambda x. P : (m, F, M \setminus \bigcup_{(k, F', M', \sigma) \in T} M', T \rightarrow \tau) \triangleright c} (\lambda)$$

We see that to derive a type for the nondeterministic choice $\text{br } P_1 P_2$, we need to derive it either for P_1 or for P_2 .

The (VAR) rule allows to have in the resulting set M some numbers that do not come from the set M' assigned to x by the type environment; these are the orders of markers placed in the leaf using this rule. Notice, however, that we allow here only orders not smaller than k (which is the order of the superterm $\lambda x. P$ binding this variable x). This is consistent with the intuitive description of the type system (page 4), which says that a marker of order n can be put in a place that will be a leaf after performing all β -reductions of orders greater than n . Indeed, the variable x remains a leaf after performing β -reductions of orders greater than k , but while performing β -reductions of order k this leaf will be replaced by a subterm substituted for x . Recall also that, by definition of a type judgment, we require that $(k, F, M', \tau) \in \mathcal{F}_k^\alpha$ and $(m, F, M, \tau) \in \mathcal{F}_m^\alpha$, for appropriate sort α ; this introduces a bound on maximal numbers that may appear in the sets F and M .

Example 1. Denoting $\hat{\rho}_1 = (1, \emptyset, \{0\}, o)$ we can derive:

$$\frac{}{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash x : (2, \emptyset, \{0\}, o) \triangleright 0} \text{ (VAR)} \quad \frac{}{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash x : (2, \emptyset, \{0, 1\}, o) \triangleright 0} \text{ (VAR)}$$

In the derivation on the right, the marker of order 1 is placed in the conclusion of the rule.

The (λ) rule allows to use (in a subderivation concerning the λ -term P) the variable x with all full types given in the set T . When the sort of $\lambda x. P$ is $\alpha \rightarrow \beta$, by definition of $\mathcal{T}^{\alpha \rightarrow \beta}$ we have that all full types in T have the same order $k = \text{ord}(\alpha \rightarrow \beta)$ (since $(T \rightarrow \tau) \in \mathcal{T}^{\alpha \rightarrow \beta}$). Recall that we intend to store in the set M the markers contained in the derivation itself and those provided by free variables, but not those provided by arguments. Because of this, in the conclusion of the rule we remove from M the markers provided by x . This operation makes sense only because there is at most one marker of each order, so

markers provided by x cannot be provided by any other free variable nor placed in the derivation itself. The set F , unlike M , stores also flags provided by arguments, so we do not need to remove anything from F .

Example 2. The (λ) rule can be used, e.g., in the following way (where a is a symbol of rank 1):

$$\frac{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash ax : (2, \{1\}, \{0\}, o) \triangleright 0}{\varepsilon \vdash \lambda x. ax : (2, \{1\}, \emptyset, \{\hat{\rho}_1\} \rightarrow o) \triangleright 0} (\lambda) \quad \frac{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash ax : (2, \emptyset, \{0, 1\}, o) \triangleright 1}{\varepsilon \vdash \lambda x. ax : (2, \emptyset, \{1\}, \{\hat{\rho}_1\} \rightarrow o) \triangleright 1} (\lambda)$$

Notice that in the conclusion of the rule, in both examples, we remove 0 from the set of marker orders, because the order-0 marker is provided by x .

The next two rules use a predicate $Comp_m$, saying how flags and markers from premisses contribute to the conclusion. It takes “as input” pairs (F_i, c_i) for $i \in I$; each of them consists of the set of flag orders F_i and of the flag counter c_i from some premiss. Moreover, the predicate takes a set of marker orders M from the current type judgment (it contains orders of markers used in the derivation, including those provided by free variables). The goal is to compute the set of flag orders F and the flag counter c that should be placed in the current type judgment. First, for each $n \in \{1, \dots, m\}$ consecutively, we decide whether a flag of order n should be placed on the current type judgment. We follow here the rules mentioned in the intuitive description. Namely, we place a flag of order n if we are on the path leading to the marker of order $n - 1$ (i.e., if $n - 1 \in M$), and simultaneously we receive an information about a flag of order $n - 1$. By receiving this information we mean that either a flag of order $n - 1$ was placed on the current type judgment, or $n - 1$ belongs to some set F_i . Actually, we place multiple flags of order n : one per each flag of order $n - 1$ placed on the current type judgment, and one per each set F_i containing $n - 1$. Then, we compute F and c . In c we store the number of flags of the maximal order m : we sum all the numbers c_i , and we add the number of order- m flags placed on the current type judgment. In F we keep elements of all F_i , and we add the orders n of flags that were placed on the current type judgment. We, however, remove from F all elements of M . This is because every flag of some order $n - 1$ should result in creating at most one flag of order n , in the closest ancestor that lies on the path leading to the marker of order $n - 1$. If we have created an order- n flag on the current type judgment, i.e., if $n - 1 \in M$, we do not want to do this again in the parent.

Below we give a formal definition, in which f'_n contains the number of order- n flags placed on the current type judgment, while f_n additionally counts the number of premisses for which $n \in F_i$. We say that $Comp_m(M; ((F_i, c_i))_{i \in I}) = (F, c)$ when

$$F = \{n \in \{0, \dots, m-1\} \mid f_n > 0 \wedge n \notin M\}, \quad c = f'_m + \sum_{i \in I} c_i, \quad \text{where, for } n \in \{0, \dots, m\},$$

$$f_n = f'_n + \sum_{i \in I} |F_i \cap \{n\}|, \quad f'_n = \begin{cases} f'_{n-1} & \text{if } n-1 \in M, \\ 0 & \text{otherwise.} \end{cases}$$

We now present a rule for constants other than br:

$$\frac{\begin{array}{l} \Gamma_i \vdash P_i : (m, F_i, M_i, o) \triangleright c_i \text{ for each } i \in \{1, \dots, r\} \quad M = M' \uplus M_1 \uplus \dots \uplus M_r \\ (m=0) \Rightarrow (F' = \emptyset \wedge c' = 1) \quad (m>0) \Rightarrow (F' = \{0\} \wedge c' = 0) \quad (r>0) \Rightarrow (M' = \emptyset) \\ a \neq \text{br} \quad \text{Split}(\Gamma \mid \Gamma_1, \dots, \Gamma_r) \quad Comp_m(M; (F', c'), (F_1, c_1), \dots, (F_r, c_r)) = (F, c) \end{array}}{\Gamma \vdash aP_1 \dots P_r : (m, F, M, o) \triangleright c} (\text{CON})$$

Here, the conditions in the second line say that in a node using the (CON) rule we always place a flag of order 0 (via F' or via c' , depending on m), and that if the node is a leaf (i.e., $r = 0$), then we

are allowed to place markers of arbitrary order (via M'). Then to the $Comp_m$ predicate, beside of pairs (F_i, c_i) coming from premisses, we also pass the information (F', c') about the order-0 flag placed in the current node; this predicate decides whether we should place also some flags of positive orders. Let us emphasize that in this rule (and similarly in the next rule) we have a disjoint union $M' \uplus M_1 \uplus \dots \uplus M_r$, which ensures that a marker of any order may be placed only in one node of a derivation.

Example 3. The (CON) rule may be instantiated in the following way:

$$\frac{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash x : (2, \emptyset, \{0\}, o) \triangleright 0}{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash ax : (2, \{1\}, \{0\}, o) \triangleright 0} \text{ (CON)} \quad \frac{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash x : (2, \emptyset, \{0, 1\}, o) \triangleright 0}{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash ax : (2, \emptyset, \{0, 1\}, o) \triangleright 1} \text{ (CON)}$$

In the left example, flags of order 0 and 1 are placed in the conclusion of the rule (a flag of order 0 is created because we are in a constant; since the marker of order 0 is visible, we do not put 0 into the set of flag orders, but instead we create a flag of order 1). In the right example, a marker of order 1 is visible, which causes that this time flags of order 0, 1, and 2 are placed in the conclusion of the (CON) rule (again, we do not put 0 nor 1 into the set of flag orders, because of 0 and 1 in the set of marker orders).

The next rule describes application:

$$\frac{\begin{array}{l} \Gamma' \vdash P : (m, F', M', \{(ord(P), F_i \upharpoonright_{<ord(P)}, M_i \upharpoonright_{<ord(P)}, \tau_i) \mid i \in I\} \rightarrow \tau) \triangleright c' \\ \Gamma_i \vdash Q : (m, F_i, M_i, \tau_i) \triangleright c_i \text{ for each } i \in I \quad M = M' \uplus \biguplus_{i \in I} M_i \\ ord(P) \leq m \quad Split(\Gamma \mid \Gamma', (\Gamma_i)_{i \in I}) \quad Comp_m(M; (F', c'), ((F_i \upharpoonright_{\geq ord(P)}, c_i))_{i \in I}) = (F, c) \end{array}}{\Gamma \vdash PQ : (m, F, M, \tau) \triangleright c} \text{ (@)}$$

In this rule, it is allowed (but in fact useless) that for two different $i \in I$ the full types (m, F_i, M_i, τ_i) are equal. It is also allowed that $I = \emptyset$, in which case no type needs to be derived for Q . Observe how flags and markers coming from premisses concerning Q are propagated: only flags and markers of order $n < ord(P)$ are visible to P , while only flags of order $n \geq ord(P)$ are passed to the $Comp_m$ predicate. This can be justified if we recall the intuitions staying behind the type system (see page 4). Indeed, while considering flags and markers of order n , we should imagine the λ -term obtained from the current λ -term by performing all β -reductions of all orders greater than n ; the distribution of flags and markers of order n in the current λ -term actually simulates their distribution in this imaginary λ -term. Thus, if $n < ord(P)$, then our application will disappear in this imaginary λ -term, and Q will be already substituted somewhere in P ; for this reason we need to pass the information about flags and markers of order n from Q to P . Conversely, if $n \geq ord(P)$, then in the imaginary λ -term the considered application will be still present, and in consequence the subterm corresponding to P will not see flags and markers of order n placed in the subterm corresponding to Q .

Example 4. Denote by $\hat{\tau}_f$ and $\hat{\tau}_m$ the types derived in Example 2:

$$\hat{\tau}_f = (2, \{1\}, \emptyset, \{\hat{\rho}_1\} \rightarrow o), \quad \text{and} \quad \hat{\tau}_m = (2, \emptyset, \{1\}, \{\hat{\rho}_1\} \rightarrow o).$$

Then, using the (@) rule, we can derive (where e is a symbol of rank 0, and f a variable):

$$\frac{\frac{\varepsilon[f \mapsto \{\hat{\tau}_m\}] \vdash f : \hat{\tau}_m \triangleright 0 \text{ (VAR)}}{\varepsilon[f \mapsto \{\hat{\tau}_f, \hat{\tau}_m\}] \vdash f : \hat{\tau}_f \triangleright 0} \quad \frac{\varepsilon \vdash e : (2, \{1\}, \{0\}, o) \triangleright 0 \text{ (CON)}}{\varepsilon[f \mapsto \{\hat{\tau}_f, \hat{\tau}_m\}] \vdash fe : (2, \emptyset, \{0, 1\}, o) \triangleright 1} \text{ (@)}$$

Recall that $\hat{\rho}_1 = (1, \emptyset, \{0\}, o)$. In the conclusion of the (@) rule the information about a flag of order 1 (from the second premiss) meets the information about the marker of order 1 (from the first premiss), and

thus a flag of order 2 is placed, which increases the flag counter. Notice that we have discarded the full type $\hat{\tau}_f$ assigned to f in the type environment; this is allowed because $\hat{\tau}_f$ provides no markers (equally well $\hat{\tau}_f$ could be assigned to f also in one or two of the premisses, and discarded there). On the other hand, the full type $\hat{\tau}_m$ provides markers, so it cannot be discarded nor duplicated (in particular, we could not pass it to the conclusion of the (CON) rule).

The key property of the type system is described by the following theorem.

Theorem 2. *Let P be a closed λ -term of sort o and complexity m . Then $\mathcal{L}(P)$ is infinite if and only if for arbitrarily large c we can derive $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$, where $\hat{\rho}_m = (m, \emptyset, \{0, \dots, m-1\}, o)$.*

The left-to-right implication of Theorem 2 (completeness of the type system) is shown in Section 4, while the opposite implication (soundness of the type system) in Section 5. In Section 6 we discuss how Theorem 1 follows from Theorem 2. Before all that, we give a few more examples of derivations, illustrating the type system and Theorem 2.

Example 5. In this example we analyze the λ -term $P_1 = R(\lambda x.ax)$, where R is defined by coinduction as $R = (\lambda f.br(fe)(R(\lambda x.f(fx))))$. As previously, a and e are symbols of rank 1 and 0, respectively. In $\mathcal{L}(P_1)$ there are trees that consist of a branch of a symbols ended with an e symbol, but only those where the number of a symbols is 2^k for some $k \in \mathbb{N}$. Notice that the complexity of P_1 is 2.

Continuing Example 4, we derive the full type $\hat{\sigma}_R = (2, \emptyset, \{0\}, \{\hat{\tau}_f, \hat{\tau}_m\} \rightarrow o)$ for R :

$$\frac{\frac{\varepsilon[f \mapsto \{\hat{\tau}_f, \hat{\tau}_m\}] \vdash fe : (2, \emptyset, \{0, 1\}, o) \triangleright 1}{\varepsilon[f \mapsto \{\hat{\tau}_f, \hat{\tau}_m\}] \vdash br(fe)(R(\lambda x.f(fx))) : (2, \emptyset, \{0, 1\}, o) \triangleright 1} \text{ (BR)}}{\varepsilon \vdash R : \hat{\sigma}_R \triangleright 1} \text{ (\lambda)}$$

Next, we derive the same full type for R , but using the second argument of the br symbol; this results in greater values of the flag counter. We start by deriving the full type $\hat{\tau}_f$ for the subterm $\lambda x.f(fx)$:

$$\frac{\frac{\varepsilon[f \mapsto \{\hat{\tau}_f\}] \vdash f : \hat{\tau}_f \triangleright 0 \quad \frac{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash x : (2, \emptyset, \{0\}, o) \triangleright 0}{\varepsilon[f \mapsto \{\hat{\tau}_f\}, x \mapsto \{\hat{\rho}_1\}] \vdash fx : (2, \{1\}, \{0\}, o) \triangleright 0} \text{ (@)}}{\varepsilon[f \mapsto \{\hat{\tau}_f\}] \vdash f : \hat{\tau}_f \triangleright 0} \text{ (@)}}{\varepsilon[f \mapsto \{\hat{\tau}_f\}, x \mapsto \{\hat{\rho}_1\}] \vdash f(fx) : (2, \{1\}, \{0\}, o) \triangleright 0} \text{ (\lambda)}}{\varepsilon[f \mapsto \{\hat{\tau}_f\}] \vdash \lambda x.f(fx) : \hat{\tau}_f \triangleright 0}$$

In the above derivation there are no flags nor markers. Next, we derive $\hat{\tau}_m$ for the same subterm:

$$\frac{\frac{\varepsilon[f \mapsto \{\hat{\tau}_m\}] \vdash f : \hat{\tau}_m \triangleright 0 \quad \frac{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash x : (2, \emptyset, \{0\}, o) \triangleright 0}{\varepsilon[f \mapsto \{\hat{\tau}_m\}, x \mapsto \{\hat{\rho}_1\}] \vdash fx : (2, \emptyset, \{0, 1\}, o) \triangleright 0} \text{ (@)}}{\varepsilon[f \mapsto \{\hat{\tau}_f\}] \vdash f : \hat{\tau}_f \triangleright 0} \text{ (@)}}{\varepsilon[f \mapsto \{\hat{\tau}_f, \hat{\tau}_m\}, x \mapsto \{\hat{\rho}_1\}] \vdash f(fx) : (2, \emptyset, \{0, 1\}, o) \triangleright 1} \text{ (\lambda)}}{\varepsilon[f \mapsto \{\hat{\tau}_f, \hat{\tau}_m\}] \vdash \lambda x.f(fx) : \hat{\tau}_m \triangleright 1}$$

Below the lower (@) rule the information about a flag of order 1 meets the information about the marker of order 1, and thus a flag of order 2 is placed, which increases the flag counter. We continue with the λ -term R :

$$\frac{\frac{\varepsilon \vdash R : \hat{\sigma}_R \triangleright c \quad \varepsilon[f \mapsto \{\hat{\tau}_f\}] \vdash \lambda x.f(fx) : \hat{\tau}_f \triangleright 0 \quad \varepsilon[f \mapsto \{\hat{\tau}_f, \hat{\tau}_m\}] \vdash \lambda x.f(fx) : \hat{\tau}_m \triangleright 1}{\varepsilon[f \mapsto \{\hat{\tau}_f, \hat{\tau}_m\}] \vdash R(\lambda x.f(fx)) : (2, \emptyset, \{0, 1\}, o) \triangleright c+1} \text{ (BR)}}{\varepsilon[f \mapsto \{\hat{\tau}_f, \hat{\tau}_m\}] \vdash br(fe)(R(\lambda x.f(fx))) : (2, \emptyset, \{0, 1\}, o) \triangleright c+1} \text{ (\lambda)}}{\varepsilon \vdash R : \hat{\sigma}_R \triangleright c+1}$$

In this fragment of a derivation no flag nor marker is placed. In particular, there is no order-2 flag in conclusion of the (@) rule, although its second premiss provides a flag of order 1 while the third premiss provides the marker of order 1. We recall from the definition of the (@) rule that the information about flags and markers coming from the arguments is divided into two parts. Numbers smaller than the order of the operator ($\text{ord}(R) = 2$ in our case) are passed to the operator, while only greater numbers (≥ 2 in our case) contribute in creating new flags via the *Comp* predicate.

By composing the above fragments of a derivation, we can derive $\varepsilon \vdash R : \hat{\sigma}_R \triangleright c$ for every $c \geq 1$. Recall that in Examples 1-3 we have derived $\varepsilon \vdash \lambda x.ax : \hat{\tau}_f \triangleright 0$ and $\varepsilon \vdash \lambda x.ax : \hat{\tau}_m \triangleright 1$. Together with the above, this allows to derive for P_1 the full type $\hat{\rho}_2 = (2, \emptyset, \{0, 1\}, o)$ (appearing in Theorem 2):

$$\frac{\varepsilon \vdash R : \hat{\sigma}_R \triangleright c \quad \varepsilon \vdash \lambda x.ax : \hat{\tau}_f \triangleright 0 \quad \varepsilon \vdash \lambda x.ax : \hat{\tau}_m \triangleright 1}{\varepsilon \vdash P_1 : \hat{\rho}_2 \triangleright c + 1} (@)$$

We can notice a correspondence between a derivation with flag counter $c + 1$ and a tree in $\mathcal{L}(P)$ of size $2^{c-1} + 1$. We remark that in every of these derivations only three flags of order 0 and only three flags of order 1 are present, in the three nodes using the (CON) rule.

Example 6. Consider a similar λ -term $P_2 = R(\lambda x.bxx)$, where R is as previously, and b is a symbol of rank 2. In $\mathcal{L}(P_2)$ we have, for every $k \in \mathbb{N}$, a full binary tree in which every branch consist of 2^k symbols b and ends with an e symbol.

This time for the subterm $\lambda x.bxx$ we need to derive three full types:

$$\begin{aligned} \hat{\tau}'_0 &= (2, \{0\}, \emptyset, \{(1, \{0\}, \emptyset, o) \rightarrow o\}), \\ \hat{\tau}'_f &= (2, \{1\}, \emptyset, \{(1, \{0\}, \emptyset, o), \hat{\rho}_1 \} \rightarrow o), \quad \text{and} \\ \hat{\tau}'_m &= (2, \emptyset, \{1\}, \{(1, \{0\}, \emptyset, o), \hat{\rho}_1 \} \rightarrow o). \end{aligned}$$

The last one is derived with flag counter 1. Notice that $\hat{\tau}'_f$ and $\hat{\tau}'_m$ need now two full types for the argument x ; the new one $(1, \{0\}, \emptyset, o)$ describes the subtree that is not on the path to the order-0 marker. We also have a new full type $\hat{\tau}'_0$ that describes the use of $\lambda x.bxx$ outside of the path to the order-0 marker.

Then, similarly as in the previous example, for every $c \geq 1$ we can derive $\varepsilon \vdash R : \hat{\sigma}'_R \triangleright c$, where $\hat{\sigma}'_R = (2, \emptyset, \{0\}, \{\hat{\tau}'_0, \hat{\tau}'_f, \hat{\tau}'_m\} \rightarrow o)$. Again, this allows to derive $\varepsilon \vdash P_2 : \hat{\rho}_2 \triangleright c + 1$. This time a derivation with flag counter $c + 1$ corresponds to a tree in $\mathcal{L}(P)$ of size $2^h - 1$ with $h = 2^{c-1} + 1$.

Example 7. Next, consider the λ -term $P_3 = R(\lambda x.x)$. The only tree in $\mathcal{L}(P_3)$ consists of a single e node. Let us see how the derivation from Example 5 has to be modified. The full type $\hat{\tau}_m$ can still be derived for $\lambda x.x$ (although with flag counter 0 now), but instead of $\hat{\tau}_f$ we have to use $\hat{\tau}''_f = (2, \emptyset, \emptyset, \{\hat{\rho}_1\} \rightarrow o)$ that provides no flag of order 1:

$$\frac{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash x : (2, \emptyset, \{0\}, o) \triangleright 0}{\varepsilon \vdash \lambda x.x : \hat{\tau}''_f \triangleright 0} \begin{matrix} (\text{VAR}) \\ (\lambda) \end{matrix} \quad \frac{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash x : (2, \emptyset, \{0, 1\}, o) \triangleright 0}{\varepsilon \vdash \lambda x.x : \hat{\tau}_m \triangleright 0} \begin{matrix} (\text{VAR}) \\ (\lambda) \end{matrix}$$

Next, for R we want to derive the full type $\hat{\sigma}''_R = (2, \emptyset, \{0\}, \{\hat{\tau}''_f, \hat{\tau}_m\} \rightarrow o)$. We can easily adopt every of the previous derivations for $\varepsilon \vdash R : \hat{\sigma}_R \triangleright c$: we basically replace every $\hat{\tau}_f$ by $\hat{\tau}''_f$. The key point is that while deriving the full type $\hat{\tau}_m$ for the subterm $\lambda x.f(fx)$, previously in the lower (@) rule we have received information about an order-1 flag, and thus we have created an order-2 flag and increased the flag counter; this time there is no information about an order-1 flag, and thus we do not create an order-2 flag and do not increase the flag counter. In consequence, even if this part of the derivation is repeated arbitrarily many times, the value of the flag counter of the whole derivation remains 1.

Example 8. Finally, consider the λ -term $P_4 = (\lambda g.P_3)(\lambda x.a(a(\dots(ax)\dots))$, which β -reduces to P_3 . Notice that we can create the following derivation:

$$\frac{\frac{\frac{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash x : (2, \emptyset, \{0\}, o) \triangleright 0 \quad (\text{VAR})}{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash ax : (2, \{1\}, \{0\}, o) \triangleright 0 \quad (\text{CON})}}{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash a(a(\dots(ax)\dots)) : (2, \{1\}, \{0\}, o) \triangleright 0 \quad (\text{CON})} \quad \vdots \quad \frac{\varepsilon[x \mapsto \{\hat{\rho}_1\}] \vdash a(a(\dots(ax)\dots)) : (2, \{1\}, \{0\}, o) \triangleright 0 \quad (\text{CON})}{\varepsilon \vdash \lambda x.a(a(\dots(ax)\dots)) : \hat{\tau}_f \triangleright 0 \quad (\lambda)}$$

Every (CON) rule used in this derivation places in its conclusion an order-0 flag and an order-1 flag. This derivation can be used as a part of a derivation for P_4 :

$$\frac{\frac{\varepsilon[g \mapsto \{\hat{\tau}_f\}] \vdash P_3 : \hat{\rho}_2 \triangleright 1}{\varepsilon \vdash \lambda g.P_3 : (2, \emptyset, \{0, 1\}, \{\hat{\tau}_f\} \rightarrow o) \triangleright 1 \quad (\lambda)} \quad \varepsilon \vdash \lambda x.a(a(\dots(ax)\dots)) : \hat{\tau}_f \triangleright 0}{\varepsilon \vdash P_4 : \hat{\rho}_2 \triangleright 1 \quad (@)}$$

Because $\hat{\tau}_f$ provides no markers, it can be removed from the type environment and thus for P_3 we can use the derivation from the previous example. We thus obtain a derivation for P_4 in which there are many order-0 and order-1 flags (but only one flag of order 2). This shows that in the flag counter we indeed need to count only the number of flags of the maximal order (not, say, the total number of flags of all orders).

4 Completeness

The proof of the left-to-right implication of Theorem 2 is divided into the following three lemmata. Recall that a β -reduction $P \rightarrow_\beta Q$ is of order n if it concerns a redex $(\lambda x.R)S$ such that $\text{ord}(\lambda x.R) = n$. The number of nodes of a tree t is denoted $|t|$. As in Theorem 2, we denote $\hat{\rho}_m = (m, \emptyset, \{0, \dots, m-1\}, o)$.

Lemma 3. *Let P be a closed λ -term of sort o and complexity m , and let $t \in \mathcal{L}(P)$. Then there exist λ -terms Q_m, Q_{m-1}, \dots, Q_0 such that $P = Q_m$, and for every $k \in \{1, \dots, m\}$ the term Q_{k-1} can be reached from Q_k using only β -reductions of order k , and we can derive $\varepsilon \vdash Q_0 : \hat{\rho}_0 \triangleright |t|$.*

Lemma 4. *Suppose that we can derive $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$. Then we can also derive $\varepsilon \vdash P : \hat{\rho}_{m+1} \triangleright c'$ for some $c' \geq \log_2 c$.*

Lemma 5. *Suppose that $P \rightarrow_\beta Q$ is a β -reduction of order m , and we can derive $\Gamma \vdash Q : \hat{\tau} \triangleright c$ with $\text{ord}(\hat{\tau}) = m$. Then we can also derive $\Gamma \vdash P : \hat{\tau} \triangleright c$.*

Now the left-to-right implication of Theorem 2 easily follows. Indeed, take a closed λ -term P of sort o and complexity m such that $\mathcal{L}(P)$ is infinite, and take any $c \in \mathbb{N}$. By \log_2^k we denote the k -fold application of the logarithm: $\log_2^0 x = x$ and $\log_2^{k+1} x = \log_2(\log_2^k x)$. Since $\mathcal{L}(P)$ is infinite, it contains a tree t so big that $\log_2^m |t| \geq c$. We apply Lemma 3 to this tree, obtaining λ -terms Q_m, Q_{m-1}, \dots, Q_0 and a derivation of $\varepsilon \vdash Q_0 : \hat{\rho}_0 \triangleright |t|$. Then repeatedly for every $k \in \{1, \dots, m\}$ we apply Lemma 4, obtaining a derivation of $\varepsilon \vdash Q_{k-1} : \hat{\rho}_k \triangleright c_k$ for some $c_k \geq \log_2^k |t|$, and Lemma 5 for every β -reduction (of order k) between Q_k and Q_{k-1} , obtaining a derivation of $\varepsilon \vdash Q_k : \hat{\rho}_k \triangleright c_k$. We end with a derivation of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c_m$, where $c_m \geq \log_2^m |t| \geq c$, as needed. In the remaining part of this section we prove the three lemmata.

Proof of Lemma 3 (sketch). Recall that $t \in \mathcal{L}(P)$ is a finite tree, thus it can be found in some finite prefix of the Böhm tree of P . By definition, this prefix will be already expanded after performing some finite number of β -reductions from P . We need to observe that these β -reductions can be rearranged, so that those of higher order are performed first. The key point is to observe that when we perform a β -reduction of some order k , then no new β -redexes of higher order appear in the term. Indeed, suppose that $(\lambda x.R)S$ is changed into $R[S/x]$ somewhere in a term, where $\text{ord}(\lambda x.R) = k$. One new redex that may appear is when R starts with a λ , and to the whole $R[S/x]$ some argument is applied; this redex is of order $\text{ord}(R) \leq k$. Some other redexes may appear when S starts with a λ , and is substituted for such appearance of x to which some argument is applied; but this redex is of order $\text{ord}(S) < k$.

We can thus find a sequence of β -reductions in which β -reductions are arranged according to their order, that leads from P to some Q_0 such that t can be found in the prefix of Q_0 that is already expanded to a tree. It is now a routine to use the rules of our type system and derive $\varepsilon \vdash Q_0 : \hat{\rho}_0 \triangleright |t|$: in every br-labeled node we choose the subtree in which t continues, and this effects in counting the number of nodes of t in the flag counter. \square

Proof of Lemma 4. Consider some derivation of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$. In this derivation we choose a leaf in which we will put the order- m marker, as follows. Starting from the root of the derivation, we repeatedly go to this premiss in which the flag counter is the greatest (arbitrarily in the case of a tie). In every node that is not on the path to the selected leaf, we replace the current type judgment $\Gamma \vdash Q : (m, F, M, \tau) \triangleright d$ by $\Gamma \vdash Q : (m+1, F', M, \tau) \triangleright 0$, where $F' = F \cup \{m\}$ if $d > 0$, and $F' = F$ otherwise. In the selected leaf and all its ancestors, we change the order from m to $m+1$, we add m to the set of marker orders, and we recalculate the flag counter.

Let us see how such transformation changes the flag counter on the path to the selected leaf. We will prove (by induction) that the previous value d and the new value d' of the flag counter in every node on this path satisfy $d' \geq \log_2 d$. In the selected leaf itself, the flag counter (being either 0 or 1) remains unchanged; we have $d' = d \geq \log_2 d$. Next, consider any proper ancestor of the selected node. Let k be the number of those of its children in which the flag counter was positive, plus the number of order- m flags placed in the considered node itself. Let also d_{\max} and d'_{\max} be the previous value and the new value of the flag counter in this child that is in the direction of the selected leaf. By construction, the flag counter in this child was maximal, which implies $k \cdot d_{\max} \geq d$, while by the induction assumption $d'_{\max} \geq \log_2 d_{\max}$. To d' we take the flag counter only from the special child, while for other children with positive flag counter we add 1, i.e., $d' = k - 1 + d'_{\max}$. Altogether we obtain $d' = k - 1 + d'_{\max} \geq k - 1 + \log_2 d_{\max} \geq \log_2(k \cdot d_{\max}) \geq \log_2 d$, as required. \square

Proof of Lemma 5. We consider the base case when $P = (\lambda x.R)S$ and $Q = R[S/x]$; the general situation (redex being deeper in P) is easily reduced to this one. In the derivation of $\Gamma \vdash Q : \hat{\tau} \triangleright c$ we identify the set I of places (nodes) where we derive a type for S substituted for x . For $i \in I$, let $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i$ be the type judgment in i . We change the nodes in I into leaves, where we instead derive $\varepsilon[x \mapsto \{\hat{\sigma}_i\}] \vdash x : \hat{\sigma}_i \triangleright 0$. It should be clear that we can repair the rest of the derivation, by changing type environments, replacing S by x in λ -terms, and decreasing flag counters. In this way we obtain derivations of $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i$ for every $i \in I$, and a derivation of $\Sigma' \vdash R : \hat{\tau} \triangleright d$, where $\Sigma' = \Sigma[x \mapsto \{\hat{\sigma}_i \mid i \in I\}]$ with $\Sigma(x) = \emptyset$, and $\text{Split}(\Gamma \mid \Sigma, (\Sigma_i)_{i \in I})$, and $c = d + \sum_{i \in I} d_i$. To the latter type judgment we apply the (λ) rule, and then we merge it with the type judgments for S using the $(@)$ rule, which results in a derivation for $\Gamma \vdash P : \hat{\tau} \triangleright c$. We remark that different $i \in I$ may give identical type judgments for S (as long as the set of markers in $\hat{\sigma}_i$ is empty); this is not a problem. The $(@)$ rule requires that $\text{ord}(\hat{\sigma}_i) = \text{ord}(\lambda x.R)$; we have that $\text{ord}(\hat{\sigma}_i) = \text{ord}(\hat{\tau})$, and $\text{ord}(\hat{\tau}) = m = \text{ord}(\lambda x.R)$ by assumption. \square

5 Soundness

In this section we sketch the proof of the right-to-left implication of Theorem 2. We, basically, need to reverse the proof from the previous section. The following new fact is now needed.

Lemma 6. *If we can derive $\Gamma \vdash P : (m, F, M, \tau) \triangleright c$ with $m - 1 \notin M$ and $\text{ord}(P) \leq m - 1$, then $c = 0$.*

A simple inductive proof is based on the following idea: flags of order m are created only when a marker of order $m - 1$ is visible; the derivation itself (together with free variables) does not provide it ($m - 1 \notin M$), and the arguments, i.e. sets T_1, \dots, T_k in $\tau = T_1 \rightarrow \dots \rightarrow T_k \rightarrow o$, may provide only markers of order at most $\text{ord}(P) - 1 \leq m - 2$ (see the definition of a type), thus no flags of order m can be created.

We say that a λ -term of the form PQ is an application of order n when $\text{ord}(P) = n$, and that an $(@)$ rule is of order n if it derives a type for an application of order n . We can successively remove applications of the maximal order from a type derivation.

Lemma 7. *Suppose that $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ for $m > 0$ is derived by a derivation D in which the $(@)$ rule of order m is used n times. Then there exists Q such that $P \rightarrow_\beta Q$ and $\varepsilon \vdash Q : \hat{\rho}_m \triangleright c$ can be derived by a derivation D' in which the $(@)$ rule of order m is used less than n times.*

Recall from the definition of the type system that the $(@)$ rule of orders higher than m cannot be used while deriving a full type of order m . Thus in D we have type judgments only for subterms of P of order at most m (although P may also have subterms of higher orders), and in type environments we only have variables of order at most $m - 1$. In order to prove Lemma 7 we choose in P a subterm RS with $\text{ord}(R) = m$ such that there is a type judgment for RS in some nodes of D (at least one), but no descendants of those nodes use the $(@)$ rule of order m . Since R is of order m , it cannot be an application (then we would choose it instead of RS) nor a variable; thus $R = \lambda x.R'$. We obtain Q by reducing the redex $(\lambda x.R')S$; the derivation D' is obtained by performing a surgery on D similar to that in the proof of Lemma 5 (but in the opposite direction). Notice that every full type (m, F, M, τ) (derived for S) with nonempty M is used for exactly one appearance of x in the derivation for R' ; full types with empty M may be used many times, or not used at all, but thanks to Lemma 6 duplicating or removing the corresponding derivations for S does not change the flag counter. In the derivations for $R'[S/x]$ no $(@)$ rule of order m may appear, and the application RS disappears, so the total number of $(@)$ rules of order m decreases.

When all $(@)$ rules of order m are eliminated, we can decrease m .

Lemma 8. *Suppose that $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ for $m > 0$ is derived by a derivation D in which the $(@)$ rule of order m is not used. Then we can also derive $\varepsilon \vdash P : \hat{\rho}_{m-1} \triangleright c'$ for some $c' \geq c$.*

The proof is easy; we simply decrease the order m of all derived full types by 1, and we ignore flags of order m and markers of order $m - 1$. To obtain the inequality $c' \geq c$ we observe that when no $(@)$ rule of order m is used, the information about flags of order $m - 1$ goes only from descendants to ancestors, and thus every flag of order m is created because of a different flag of order $m - 1$.

By repeatedly applying the two above lemmata, out of a derivation of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ we obtain a derivation of $\varepsilon \vdash Q : \hat{\rho}_0 \triangleright c'$, where $P \rightarrow_\beta^* Q$ and $c' \geq c$. Since $\hat{\rho}_0$ is of order 0, using the latter derivation it is easy to find in the already expanded part of Q (and thus in $\mathcal{L}(Q) = \mathcal{L}(P)$) a tree t such that $|t| = c' \geq c$.

6 Effectiveness

Finally, we show how Theorem 1 follows from Theorem 2, i.e., how given a λY -term P of complexity m we can check whether $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ can be derived for arbitrarily large c . We say that two type judgments are equivalent if they differ only in the value of the flag counter. Let us consider a set \mathcal{D} of all derivations

of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ in which on each branch (i.e., each root-leaf path) there are at most three type judgments from every equivalence class, and among premisses of each (@) rule there is at most one type judgment from every equivalence class. These derivations use only type judgments $\Gamma \vdash Q : \hat{\tau} \triangleright d$ with Q being a subterm of P and with $\Gamma(x) \neq \emptyset$ only for variables x appearing in P . Since a finite λY -term, even when seen as an infinitary λ -term, has only finitely many subterms, this introduces a common bound on the height of all derivations in \mathcal{D} , and on their degree (i.e., on the maximal number of premisses of a rule). It follows that there are only finitely many derivations in \mathcal{D} , and thus we can compute all of them.

We claim that $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ can be derived for arbitrarily large c if and only if in \mathcal{D} there is a derivation in which on some branch there are two equivalent type judgments with different values of the flag counter (and the latter condition can be easily checked). Indeed, having such a derivation, we can repeat its fragment between the two equivalent type judgments, obtaining derivations of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ with arbitrarily large c . We use here an additivity property of our type system: if out of $\Gamma \vdash Q : \hat{\tau} \triangleright d$ we can derive $\Gamma' \vdash Q' : \hat{\tau}' \triangleright d'$, then out of $\Gamma \vdash Q : \hat{\tau} \triangleright d + k$ we can derive $\Gamma' \vdash Q' : \hat{\tau}' \triangleright d' + k$, for every $k \geq -d$. Conversely, take a derivation of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ for some large enough c . Suppose that some of its (@) rules uses two equivalent premisses. These premisses concern the argument subterm, which is of smaller order than the operator subterm, and thus of order at most $m - 1$. The set of marker orders in these premisses has to be empty, as the sets of marker orders from all premisses have to be disjoint. Thus, by Lemma 6, the flag counter in our two premisses is 0. In consequence, we can remove one of the premisses, without changing anything in the remaining part of the derivation, even the flag counters. In this way we clean the whole derivation, so that at the end among premisses of each (@) rule there is at most one type judgment from every equivalence class. The degree is now bounded, and at each node the flag counter grows only by a constant above the sum of flag counters from the children. Thus, if c is large enough, we can find on some branch two equivalent type judgments with different values of the flag counter. Then, for some pairs of equivalent type judgments, we remove the part of the derivation between these type judgments (and we adopt appropriately the flag counters in the remaining part). It is not difficult to perform this cleaning so that the resulting derivation will be in \mathcal{D} , and simultaneously on some branch there will remain two equivalent type judgments with different values of the flag counter.

7 Conclusions

In this paper, we have shown an approach for expressing quantitative properties of Böhm trees using an intersection type system, on the example of the finiteness problem. It is an ongoing work to apply this approach to the diagonal problem, which should give a better complexity than that of the algorithm from [6]. Another ongoing work is to obtain an algorithm for model checking Böhm trees with respect to the Weak MSO+U logic [3]. This logic extends Weak MSO by a new quantifier U, expressing that a subformula holds for arbitrarily large finite sets. Furthermore, it seems feasible that our methods may help in proving a pumping lemma for nondeterministic HORSEs.

References

- [1] Achim Blumensath (2008): *On the Structure of Graphs in the Caucal Hierarchy*. *Theor. Comput. Sci.* 400(1-3), pp. 19–45, doi:10.1016/j.tcs.2008.01.053.
- [2] Achim Blumensath (2013): *Erratum to "On the Structure of Graphs in the Caucal Hierarchy" [Theoret. Comput. Sci. 400 (2008) 19-45]*. *Theor. Comput. Sci.* 475, pp. 126–127, doi:10.1016/j.tcs.2012.12.044.

- [3] Mikołaj Bojańczyk & Szymon Toruńczyk (2012): *Weak MSO+U over Infinite Trees*. In Christoph Dürr & Thomas Wilke, editors: *29th International Symposium on Theoretical Aspects of Computer Science, STACS 2012, February 29th - March 3rd, 2012, Paris, France, LIPIcs 14*, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, pp. 648–660, doi:10.4230/LIPIcs.STACS.2012.648.
- [4] Christopher H. Broadbent, Arnaud Carayol, Matthew Hague & Olivier Serre (2012): *A Saturation Method for Collapsible Pushdown Systems*. In Artur Czumaj, Kurt Mehlhorn, Andrew M. Pitts & Roger Wattenhofer, editors: *Automata, Languages, and Programming - 39th International Colloquium, ICALP 2012, Warwick, UK, July 9-13, 2012, Proceedings, Part II, Lecture Notes in Computer Science 7392*, Springer, pp. 165–176, doi:10.1007/978-3-642-31585-5_18.
- [5] Christopher H. Broadbent & Naoki Kobayashi (2013): *Saturation-Based Model Checking of Higher-Order Recursion Schemes*. In Simona Ronchi Della Rocca, editor: *Computer Science Logic 2013 (CSL 2013), CSL 2013, September 2-5, 2013, Torino, Italy, LIPIcs 23*, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, pp. 129–148, doi:10.4230/LIPIcs.CSL.2013.129.
- [6] Lorenzo Clemente, Paweł Parys, Sylvain Salvati & Igor Walukiewicz (2016): *The Diagonal Problem for Higher-Order Recursion Schemes is Decidable*. In Martin Grohe, Eric Koskinen & Natarajan Shankar, editors: *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016, ACM*, pp. 96–105, doi:10.1145/2933575.2934527.
- [7] Lorenzo Clemente, Paweł Parys, Sylvain Salvati & Igor Walukiewicz (2015): *Ordered Tree-Pushdown Systems*. In Prahladh Harsha & G. Ramalingam, editors: *35th IARCS Annual Conference on Foundation of Software Technology and Theoretical Computer Science, FSTTCS 2015, December 16-18, 2015, Bangalore, India, LIPIcs 45*, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, pp. 163–177, doi:10.4230/LIPIcs.FSTTCS.2015.163.
- [8] Werner Damm (1982): *The IO- and OI-Hierarchies*. *Theor. Comput. Sci.* 20, pp. 95–207, doi:10.1016/0304-3975(82)90009-3.
- [9] Matthew Hague, Jonathan Kochems & C.-H. Luke Ong (2016): *Unboundedness and Downward Closures of Higher-Order Pushdown Automata*. In Rastislav Bodík & Rupak Majumdar, editors: *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*, pp. 151–163, doi:10.1145/2837614.2837627.
- [10] Matthew Hague, Andrzej S. Murawski, C.-H. Luke Ong & Olivier Serre (2008): *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*, pp. 452–461, doi:10.1109/LICS.2008.34.
- [11] Alexander Kartzow & Paweł Parys (2012): *Strictness of the Collapsible Pushdown Hierarchy*. In Branislav Rovan, Vladimiro Sassone & Peter Widmayer, editors: *Mathematical Foundations of Computer Science 2012 - 37th International Symposium, MFCS 2012, Bratislava, Slovakia, August 27-31, 2012. Proceedings, Lecture Notes in Computer Science 7464*, Springer, pp. 566–577, doi:10.1007/978-3-642-32589-2_50.
- [12] Teodor Knapik, Damian Niwiński & Paweł Urzyczyn (2002): *Higher-Order Pushdown Trees Are Easy*. In Mogens Nielsen & Uffe Engberg, editors: *Foundations of Software Science and Computation Structures, 5th International Conference, FOSSACS 2002. Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2002 Grenoble, France, April 8-12, 2002, Proceedings, Lecture Notes in Computer Science 2303*, Springer, pp. 205–222, doi:10.1007/3-540-45931-6_15.
- [13] Naoki Kobayashi (2009): *Types and Higher-Order Recursion Schemes for Verification of Higher-Order Programs*. In Zhong Shao & Benjamin C. Pierce, editors: *Proceedings of the 36th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2009, Savannah, GA, USA, January 21-23, 2009, ACM*, pp. 416–428, doi:10.1145/1480881.1480933.
- [14] Naoki Kobayashi (2013): *Pumping by Typing*. In: *28th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2013, New Orleans, LA, USA, June 25-28, 2013, IEEE Computer Society*, pp. 398–407, doi:10.1109/LICS.2013.46.

- [15] Naoki Kobayashi, Kazuhiro Inaba & Takeshi Tsukada (2014): *Unsafe Order-2 Tree Languages Are Context-Sensitive*. In Anca Muscholl, editor: *Foundations of Software Science and Computation Structures - 17th International Conference, FOSSACS 2014, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2014, Grenoble, France, April 5-13, 2014, Proceedings, Lecture Notes in Computer Science* 8412, Springer, pp. 149–163, doi:10.1007/978-3-642-54830-7_10.
- [16] Naoki Kobayashi & C.-H. Luke Ong (2009): *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, pp. 179–188, doi:10.1109/LICS.2009.29.
- [17] C.-H. Luke Ong (2006): *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, pp. 81–90, doi:10.1109/LICS.2006.38.
- [18] Paweł Parys (2016): *A Characterization of Lambda-Terms Transforming Numerals*. *J. Funct. Program.* 26, p. e12, doi:10.1017/S0956796816000113.
- [19] Paweł Parys (2012): *On the Significance of the Collapse Operation*. In: *Proceedings of the 27th Annual IEEE Symposium on Logic in Computer Science, LICS 2012, Dubrovnik, Croatia, June 25-28, 2012*, IEEE Computer Society, pp. 521–530, doi:10.1109/LICS.2012.62.
- [20] Steven J. Ramsay, Robin P. Neatherway & C.-H. Luke Ong (2014): *A Type-Directed Abstraction Refinement Approach to Higher-Order Model Checking*. In Suresh Jagannathan & Peter Sewell, editors: *The 41st Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL '14, San Diego, CA, USA, January 20-21, 2014*, ACM, pp. 61–72, doi:10.1145/2535838.2535873.
- [21] Sylvain Salvati & Igor Walukiewicz (2016): *Simply Typed Fixpoint Calculus and Collapsible Pushdown Automata*. *Mathematical Structures in Computer Science* 26(7), pp. 1304–1350, doi:10.1017/S0960129514000590.

A Proof of Lemma 3

Let us write $P \approx_n P'$ if the λ -terms P and P' agree up to depth $n \in \mathbb{N}$. Formally, \approx_n is defined by induction on n as the smallest equivalence relation such that:

- $P \approx_0 Q$ for all λ -terms P, Q ,
- $aP_1 \dots P_r \approx_n aP'_1 \dots P'_r$ if $P_i \approx_{n-1} P'_i$ for all $i \in \{1, \dots, r\}$,
- $PQ \approx_n P'Q'$ if $P \approx_{n-1} P'$ and $Q \approx_{n-1} Q'$, and
- $\lambda x.P \approx_n \lambda x.P'$ if $P \approx_{n-1} P'$.

Observe that $P \approx_n P'$ implies $P \approx_k P'$ for $k < n$.

We split the proof of Lemma 3 into several lemmata.

Lemma 9. *Let P be a closed λ -term of sort o , and let $t \in \mathcal{L}(P)$. Then there exists a number $n \in \mathbb{N}$ and a λ -term Q such that $P \rightarrow_\beta^* Q$, and whenever $Q \approx_n Q'$ and $Q' \rightarrow_\beta^* Q''$ for some λ -terms Q', Q'' , then one can derive $\varepsilon \vdash Q'' : \hat{\rho}_0 \triangleright |t|$.*

Proof. Let $\rightarrow_{\text{br}}^k$ be the relation obtained by composing \rightarrow_{br} with itself k times. By definition of $\mathcal{L}(P)$ we have that $BT(P) \rightarrow_{\text{br}}^* t$, and thus $BT(P) \rightarrow_{\text{br}}^k t$ for some $k \in \mathbb{N}$. We prove the lemma by induction on $|t| + k$, where k is the smallest number such that $BT(P) \rightarrow_{\text{br}}^k t$. Because $\mathcal{L}(P) \neq \emptyset$, we have that $P \rightarrow_\beta^* aP_1 \dots P_r$ (where possibly $a = \text{br}$). Then $BT(P)$ has a in its root, and the subtrees starting in root's children are $BT(P_1), \dots, BT(P_r)$. We have two cases.

Suppose first that $a \neq \text{br}$ (this case serves as the induction base when $r = 0$). Then also t has a in its root, and the subtrees t_1, \dots, t_r starting in root's children are such that $BT(P_i) \rightarrow_{\text{br}}^{k_i} t_i$ for all $i \in \{1, \dots, r\}$, where $k_1 + \dots + k_r = k$. We have $|t_i| + k_i < |t| + k$, since $|t_i| < |t|$. By the induction assumption, for every $i \in \{1, \dots, r\}$ we obtain a number n_i and a λ -term Q_i such that $P_i \rightarrow_\beta^* Q_i$, and whenever $Q_i \approx_{n_i} Q'_i$ and $Q'_i \rightarrow_\beta^* Q''_i$ for some λ -terms Q'_i, Q''_i , then one can derive $\varepsilon \vdash Q''_i : \hat{\rho}_0 \triangleright |t_i|$. Taking $Q = aQ_1 \dots Q_r$ we have $P \rightarrow_\beta^* aP_1 \dots P_r \rightarrow_\beta^* aQ_1 \dots Q_r = Q$. As n we take $1 + n_1 + \dots + n_r$. Let now Q' and Q'' be such that $Q \approx_n Q'$ and $Q' \rightarrow_\beta^* Q''$. Then $Q' = aQ'_1 \dots Q'_r$ and $Q'' = aQ''_1 \dots Q''_r$, where $Q_i \approx_{n-1} Q'_i$ (thus also $Q_i \approx_{n_i} Q'_i$) and $Q'_i \rightarrow_\beta^* Q''_i$ for $i \in \{1, \dots, r\}$. From the induction assumption we obtain derivations of $\varepsilon \vdash Q''_i : \hat{\rho}_0 \triangleright |t_i|$. Recall that $\hat{\rho}_0 = (0, \emptyset, \emptyset, o)$. We apply the (CON) rule to these derivations. Since $\text{ord}(\hat{\rho}) = 0$, the pair (F', c') appearing in the rule's definition equals $(\emptyset, 1)$. The Comp_0 predicate simply adds flag counters from its arguments, and hence the resulting flag counter is $1 + |t_1| + \dots + |t_r|$, which equals $|t|$. Thus the resulting type judgment is $\varepsilon \vdash Q'' : \hat{\rho}_0 \triangleright |t|$, as required.

Next, suppose that $a = \text{br}$ (and hence $r = 2$). It should be clear that in the shortest reduction sequence $BT(P) \rightarrow_{\text{br}}^k t$ we can rearrange reductions (without increasing their number) so that we first eliminate the br symbol from the root of $BT(P)$. In other words, we have $BT(P) \rightarrow_{\text{br}} BT(P_i) \rightarrow_{\text{br}}^{k-1} t$, for some $i \in \{1, 2\}$. Let us focus our attention on the case of $i = 1$; the case of $i = 2$ is completely symmetric. Since $|t| + k - 1 < |t| + k$, from the induction assumption we obtain a number n_1 and a λ -term Q_1 such that $P_1 \rightarrow_\beta^* Q_1$, and whenever $Q_1 \approx_{n_1} Q'_1$ and $Q'_1 \rightarrow_\beta^* Q''_1$ for some λ -terms Q'_1, Q''_1 , then one can derive $\varepsilon \vdash Q''_1 : \hat{\rho}_0 \triangleright |t|$. Taking $Q = \text{br} Q_1 P_2$ we have $P \rightarrow_\beta^* \text{br} P_1 P_2 \rightarrow_\beta^* \text{br} Q_1 P_2 = Q$. As n we take $1 + n_1$. Let now Q', Q'' be such that $Q \approx_n Q'$ and $Q' \rightarrow_\beta^* Q''$. Then $Q' = \text{br} Q'_1 Q'_2$ and $Q'' = \text{br} Q''_1 Q''_2$, where $Q_1 \approx_{n-1} Q'_1$ and $Q'_1 \rightarrow_\beta^* Q''_1$. By the induction assumption we can derive $\varepsilon \vdash Q''_1 : \hat{\rho}_0 \triangleright |t|$, which after applying the (BR) rule gives $\varepsilon \vdash Q'' : \hat{\rho}_0 \triangleright |t|$. \square

Lemma 10. *If $P \approx_n P'$ and $Q \approx_n Q'$ for some $n \in \mathbb{N}$, then also $P[Q/x] \approx_n P'[Q'/x]$.*

Proof. Induction on n . For $n = 0$ the lemma is obvious: \approx_0 always holds. When $n > 0$ and $P = RS$, then $P' = R'S'$ with $R \approx_{n-1} R'$ and $S \approx_{n-1} S'$. By the induction assumption we have $R[Q/x] \approx_{n-1} R'[Q'/x]$ and $S[Q/x] \approx_{n-1} S'[Q'/x]$, and thus $P[Q/x] \approx_n P'[Q'/x]$. The cases when $P = aP_1 \dots P_r$ or $P = \lambda y.Q$ are similar. Finally, when $P = P'$ is a variable, the thesis follows immediately from $Q \approx_n Q'$. \square

Lemma 11. *If $P \approx_{n+2} P'$ and $P \rightarrow_\beta Q$, then for some Q' we have $P' \rightarrow_\beta^* Q'$ and $Q \approx_n Q'$.*

Proof. Induction on n . If $n = 0$, the thesis holds for $Q' = P'$. Suppose that $n > 0$ and $P = (\lambda x.R)S$ and $Q = R[S/x]$. Then $P' = (\lambda x.R')S'$, where $R \approx_n R'$ and $S \approx_{n+1} S'$. Taking $Q' = R'[S'/x]$ we have $P' \rightarrow_\beta Q'$, and by Lemma 10 $Q \approx_n Q'$. The remaining case is that $n > 0$ and the redex involved in the β -reduction $P \rightarrow_\beta Q$ is not located on the front of P . Then the thesis follows from the induction assumption. Let us consider only a representative example: suppose that $P = RS$, and $Q = TS$, and $R \rightarrow_\beta T$. In this case $P' = R'S'$ with $R \approx_{n+1} R'$ and $S \approx_{n+1} S'$. The induction assumption gives us T' such that $R' \rightarrow_\beta^* T'$ and $T \approx_{n-1} T'$. Thus for $Q' = T'S'$ we have $P' \rightarrow_\beta^* Q'$ and $Q \approx_n Q'$. \square

Lemma 12. *For every $n \in \mathbb{N}$, we can represent every λ -term P as $P = P'[S_1/x_1, \dots, S_s/x_s]$ so that P' is finite and $P \approx_n P'$.*

Proof. Induction on n . For $n = 0$ we represent $P = x[S/x]$, and clearly $x \approx_0 P$. For $n > 0$ we consider the representative case of $P = QR$; for other forms of P the proof is similar. The induction assumption gives us representations $Q = Q'[S_1/x_1, \dots, S_r/x_r]$ and $R = R'[S_{r+1}/x_{r+1}, \dots, S_s/x_s]$ with Q', R' finite and such that $Q \approx_{n-1} Q'$ and $R \approx_{n-1} R'$. W.l.o.g. we can assume that the fresh variables x_1, \dots, x_r are not free in R' , and the fresh variables x_{r+1}, \dots, x_s are not free in Q' . Then, for $P' = Q'R'$ we have $P \approx_n P'$ and $P = P'[S_1/x_1, \dots, S_s/x_s]$. \square

Corollary 13. *Let $n \in \mathbb{N}$, and let P, Q be λ -terms such that $P \rightarrow_\beta^* Q$. Then we can represent P as $P = P'[S_1/x_1, \dots, S_s/x_s]$ so that P' is finite, and for some λ -term Q' it holds $P' \rightarrow_\beta^* Q'$ and $Q \approx_n Q'$.*

Proof. Let us write $P = P_0 \rightarrow_\beta P_1 \rightarrow_\beta \dots \rightarrow_\beta P_l = Q$. By Lemma 12 we can write $P = P'[S_1/x_1, \dots, S_s/x_s]$ so that P' is finite and $P \approx_{n+2l} P'$. Take $P'_0 = P'$. Consecutively for all $i \in \{1, \dots, l\}$ Lemma 11 gives us a λ -term P'_i such that $P'_{i-1} \rightarrow_\beta^* P'_i$ and $P_i \approx_{n+2(l-i)} P'_i$. At the end, for $Q' = P'_l$ we have $P' \rightarrow_\beta^* Q'$ and $Q \approx_n Q'$. \square

Below, a λ -term $(\lambda x.P)Q$ is called a β -redex of order k if $k = \text{ord}(\lambda x.P)$.

Lemma 14. *Let $k \in \mathbb{N}$, and let P be a λ -term without β -redexes of orders higher than k (as subterms). If $P \rightarrow_\beta Q$ is a β -reduction of order k , then also in Q there are no β -redexes of order higher than k .*

Proof. This lemma was already justified on page 12, but let us repeat. Let P' be obtained from P by replacing by y the β -redex $(\lambda x.R)S$ involved in the β -reduction $P \rightarrow_\beta Q$. Then we have $P = P'[(\lambda x.R)S/y]$ and $Q = P'[R[S/x]/y]$. Suppose that Q has a β -redex of order higher than k , i.e., a subterm $(\lambda z.T)U$ with $\text{ord}(\lambda z.T) > k$; we want to prove that this is impossible. The subterm $(\lambda z.T)U$, like every subterm of Q , can be found in one of the following three places:

- Possibly $(\lambda z.T)U$ is a subterm of S . This is impossible, because by assumption P (and thus also its subterm S) contains no β -redexes of orders higher than k .

- Possibly $(\lambda z.T)U = V[S/x]$ for a subterm V of R , where $V \neq x$. Then V has to be an application $V = WX$, with $\text{ord}(W) = \text{ord}(\lambda z.T) > k$. We have $W \neq x$, because $\text{ord}(x) < \text{ord}(\lambda x.R) = k$. Thus W is a λ -abstraction, with V being itself a β -redex of order higher than k , which is again impossible by assumption.
- Otherwise $(\lambda z.T)U = V[R[S/x]/y]$ for a subterm V of P' , where $V \neq y$. Again, V has to be an application $V = WX$, with $\text{ord}(W) = \text{ord}(\lambda z.T) > k$. We have $W \neq y$, because $\text{ord}(y) = \text{ord}(R) \leq \text{ord}(\lambda x.R) = k$. Thus W is a λ -abstraction, with V being itself a β -redex of order higher than k ; this is impossible, since $V[(\lambda x.R)S/y]$ is a subterm of P and also a β -redex of order higher than k . \square

Lemma 15. *Let P' be a finite λ -term of complexity at most m . Then there exist λ -terms $Q'_m, Q'_{m-1}, \dots, Q'_0$ such that $P' = Q'_m$, and for every $k \in \{1, \dots, m\}$ the term Q'_{k-1} can be reached from Q'_k using only β -reductions of order k , and Q'_0 is in β -normal form.*

Proof. Take $Q'_m = P'$. Then, for $k = m, \dots, 1$, consecutively, out of Q'_k we perform β -reductions of order k as long as possible, and as Q'_{k-1} we take the resulting λ -term, from which no more β -reductions of order k are possible. Every such sequence of β -reductions finally ends, because P' is finite. For every $k \in \{1, \dots, m\}$ Lemma 14 ensures that in Q'_0 there are no β -redexes of order k , because all β -reductions between Q'_{k-1} and Q'_0 were of orders smaller than k . Moreover, because the complexity of a λ -term cannot grow during β -reductions, the complexity of Q'_0 is at most m , and hence it has no β -redexes of order higher than m . Thus Q'_0 is in β -normal form. \square

Proof of Lemma 3. Recall that in this lemma we are given a closed λ -term P of sort o and complexity m , and some $t \in \mathcal{L}(P)$, and our goal is to exhibit λ -terms Q_m, Q_{m-1}, \dots, Q_0 such that $P = Q_m$, and for every $k \in \{1, \dots, m\}$ the term Q_{k-1} can be reached from Q_k using only β -reductions of order k , and we can derive $\varepsilon \vdash Q_0 : \hat{\rho}_0 \triangleright |t|$.

We first apply Lemma 9 to P and t , obtaining a number n and a λ -term Q such that $P \rightarrow_\beta^* Q$. Then, we apply to them Corollary 13, obtaining a representation $P = P'[S_1/x_1, \dots, S_s/x_s]$ for finite P' , and a λ -term Q' . Notice that the complexity of P' cannot be higher than that of P , as for every subterm R of P' , the term $R[S_1/x_1, \dots, S_s/x_s]$ is a subterm of P , and has the same order as R . We then apply Lemma 15 to P' , obtaining λ -terms $Q'_m, Q'_{m-1}, \dots, Q'_0$. As Q_k we take $Q'_k[S_1/x_1, \dots, S_s/x_s]$, for $k \in \{0, \dots, m\}$. We have $Q_m = P$ since $Q'_m = P'$. For every $k \in \{1, \dots, m\}$ Lemma 15 gives us a sequence of β -reduction of order k from Q'_k to Q'_{k-1} . After performing the substitution $[S_1/x_1, \dots, S_s/x_s]$ to every λ -term in this sequence, all β -reductions in this sequence remain correct β -reductions of order k , and now the sequence leads from Q_k to Q_{k-1} .

Finally, by Corollary 13 we have $P' \rightarrow_\beta^* Q'$ and $Q \approx_n Q'$. Since, by Lemma 15, the β -normal form of P' is Q'_0 , we also have $Q' \rightarrow_\beta^* Q'_0$. Thus by Lemma 9 (where we take Q'_0 as Q'') we obtain a derivation of $\varepsilon \vdash Q'_0 : \hat{\rho}_0 \triangleright |t|$. To every λ -term in this derivation we apply the substitution $[S_1/x_1, \dots, S_s/x_s]$, obtaining a derivation of $\varepsilon \vdash Q_0 : \hat{\rho}_0 \triangleright |t|$. The derivation remains correct: the problem could only appear in a (VAR) rule used for some of the variables x_1, \dots, x_s , but since the type environment of the resulting type judgment is empty, the derivation uses the (VAR) rule only for bound variables of Q'_0 , not for x_1, \dots, x_s . \square

B Proof of Lemma 4

Lemma 16. *If $\text{Comp}_m(M; ((F_i, c_i))_{i \in I}) = (F, c)$, where M and F_i are subsets of $\{0, \dots, m-1\}$, then it holds that $\text{Comp}_{m+1}(M; ((G_i, 0))_{i \in I}) = (G, 0)$, where $G = F \cup \{m \mid c > 0\}$ and $G_i = F_i \cup \{m \mid c_i > 0\}$.*

In the above, by $\{m \mid c > 0\}$ we mean the set $\{m\}$ when $c > 0$, and \emptyset otherwise.

Proof. In the definition of the Comp predicate some numbers f_n and f'_n are computed. Denote by $f_{n,m}$ and $f'_{n,m}$ the values taken by f_n and f'_n in the above instantiation of the Comp_m predicate, and by $f_{n,m+1}$ and $f'_{n,m+1}$ their values in the above instantiation of the Comp_{m+1} predicate. Since the sets F_i and G_i are the same when restricted to numbers smaller than m , we have $f_{n,m} = f_{n,m+1}$ for $n < m$, and $f'_{n,m} = f'_{n,m+1}$ for $n \leq m$. In particular, for $n < m$ the definition of Comp says that $n \in F \Leftrightarrow n \in G$. We also have $f'_{m+1,m+1} = 0$, since $m \notin M$; thus the flag counter computed by the Comp_{m+1} predicate (as the sum of $f'_{m+1,m+1} = 0$ and of the zeroes given as arguments to Comp_{m+1}) is actually 0. Finally, we recall that $c = f'_{m,m} + \sum_{i \in I} c_i$; thus $c > 0$ if and only if $f'_{m,m} > 0$ or $c_i > 0$ for some $i \in I$. On the other hand, we have $f_{m,m+1} = f'_{m,m+1} + \sum_{i \in I} |G_i \cap \{m\}|$; thus $f_{m,m+1} > 0$ if and only if $f'_{m,m+1} > 0$ or $m \in G_i$ for some $i \in I$. Since $f'_{m,m} = f'_{m,m+1}$ and $c_i > 0 \Leftrightarrow m \in G_i$, we obtain $c > 0 \Leftrightarrow f_{m,m+1} > 0$. The Comp_{m+1} predicate wants to put m to the set G if only if $f_{m,m+1} > 0$ (since $m \notin M$), thus if and only if $c > 0$; this agrees with our definition of G . \square

Lemma 17. *If we can derive $\Gamma \vdash R : (m, F, M, \tau) \triangleright c$, then we can also derive $\Gamma \vdash R : (m+1, G, M, \tau) \triangleright 0$, where $G = F \cup \{m \mid c > 0\}$.*

Proof. Denote $\hat{\tau} = (m, F, M, \tau)$ and $\hat{\tau}' = (m+1, G, M, \tau)$. The proof is by induction on the structure of the derivation of $\Gamma \vdash R : \hat{\tau} \triangleright c$. We have several cases, depending on the shape of R .

Suppose that $R = \text{br } P_1 P_2$. Then the last rule of the derivation is (BR) with premiss $\Gamma \vdash P_i : \hat{\tau} \triangleright c$ for some $i \in \{1, 2\}$. The induction assumption gives us a derivation of $\Gamma \vdash P_i : \hat{\tau}' \triangleright 0$. Applying back the (BR) rule we derive $\Gamma \vdash R : \hat{\tau}' \triangleright 0$.

Next, suppose that $R = x$ is a variable. Then the (VAR) rule requires that $c = 0$, and thus $G = F$. We just need to change in this rule m to $m+1$ (which is not problematic at all), and we obtain a derivation of $\Gamma \vdash R : \hat{\tau}' \triangleright 0$.

Next, suppose that $R = \lambda x. P$. Then the derivation ends with the (λ) rule, whose premiss is $\Gamma'[x \mapsto T] \vdash P : (m, F, M_\lambda, \tau_\lambda) \triangleright c$, where $\tau = T \rightarrow \tau_\lambda$, and $M = M_\lambda \setminus \bigcup_{(k, F', M', \sigma) \in T} M'$, and $\text{Split}(\Gamma \mid \Gamma')$, and $\Gamma'(x) = \emptyset$. The induction assumption gives us a derivation of $\Gamma'[x \mapsto T] \vdash P : (m+1, G, M_\lambda, \tau_\lambda) \triangleright 0$. We can apply back the (λ) rule, and derive $\Gamma \vdash R : \hat{\tau}' \triangleright 0$.

Next, suppose that $R = a P_1 \dots P_r$ for $a \neq \text{br}$. We have $\tau = o$. Let $\Gamma_i \vdash P_i : (m, F_i, M_i, o) \triangleright c_i$ for $i \in \{1, \dots, r\}$ be the premisses of the final (CON) rule. Using the induction assumption we derive $\Gamma_i \vdash P_i : (m+1, G_i, M_i, o) \triangleright 0$, where $G_i = F_i \cup \{m \mid c_i > 0\}$. We want to apply back the (CON) rule. Some of the conditions of the (CON) rule remain unchanged: $M = M' \uplus M_1 \uplus \dots \uplus M_r$, and $M' = \emptyset$ if $r > 0$, and $\text{Split}(\Gamma \mid \Gamma_1, \dots, \Gamma_r)$. We also know that $(F, c) = \text{Comp}_m(M; (F', c'), (F_1, c_1), \dots, (F_r, c_r))$, where $F' = \emptyset$ and $c' = 1$ if $m = 0$, and $F' = \{0\}$ and $c' = 0$ if $m > 0$. Notice that $F' \cup \{m \mid c' > 0\} = \{0\}$. Thus from Lemma 16 we obtain that $\text{Comp}_{m+1}(M; (\{0\}, 0), (G_1, 0), \dots, (G_r, 0)) = (G, 0)$, as required.

Finally, suppose that $R = PQ$. Let $\Gamma' \vdash P : (m, F', M', T \rightarrow \tau) \triangleright c'$ and $\Gamma_i \vdash Q : (m, F_i, M_i, \tau_i) \triangleright c_i$ for each $i \in I$ be the premisses of the final (@) rule, where $T = \{(ord(P), F_i \upharpoonright_{<ord(P)}, M_i \upharpoonright_{<ord(P)}, \tau_i) \mid i \in I\}$. Using the induction assumption we derive $\Gamma' \vdash P : (m+1, G', M', T \rightarrow \tau) \triangleright 0$, where $G' = F' \cup \{m \mid c' > 0\}$, and $\Gamma_i \vdash Q : (m+1, G_i, M_i, \tau_i) \triangleright c_i$, where $G_i = F_i \cup \{m \mid c_i > 0\}$, for each $i \in I$. We want to apply back the (@) rule. The conditions $M = M' \uplus \biguplus_{i \in I} M_i$ and $\text{Split}(\Gamma \mid \Gamma', (\Gamma_i)_{i \in I})$ required by the (@) rule remain unchanged, while $ord(P) \leq m+1$ holds because previously we had $ord(P) \leq m$.

Also due to $\text{ord}(P) \leq m$ we have $G_i \upharpoonright_{<\text{ord}(P)} = F_i \upharpoonright_{<\text{ord}(P)}$, and thus the type needed now for P , that is $\{(\text{ord}(P), G_i \upharpoonright_{<\text{ord}(P)}, M_i \upharpoonright_{<\text{ord}(P)}, \tau_i) \mid i \in I\} \rightarrow o$, is actually equal $T \rightarrow o$. The condition $\text{Comp}_m(M; (F', c'), ((F_i \upharpoonright_{\geq \text{ord}(P)}, c_i))_{i \in I}) = (F, c)$ implies that $\text{Comp}_{m+1}(M; (G', 0), ((G_i \upharpoonright_{\geq \text{ord}(P)}, 0))_{i \in I}) = (G, 0)$ by Lemma 16 (notice that $G_i \upharpoonright_{\geq \text{ord}(P)} = F_i \upharpoonright_{\geq \text{ord}(P)} \cup \{m \mid c_i > 0\}$ because $m \geq \text{ord}(P)$). \square

Lemma 18. *Suppose that $\text{Comp}_m(M; ((F_i, c_i))_{i \in I}) = (F, c)$, where M and F_i are subsets of $\{0, \dots, m-1\}$. If $s \in I$ is such that $c_i \leq c_s$ for all $i \in I$, and $d_s \geq \log_2 c_s$, and $G_i = F_i \cup \{m \mid c_i > 0\}$ for $i \in I$, then $\text{Comp}_{m+1}(M \cup \{m\}; (F_s, d_s), ((G_i, 0))_{i \in I \setminus \{s\}}) = (F, d)$ for some d such that $d \geq \log_2 c$.*

Proof. As in the proof of Lemma 16, denote by $f_{n,m}$ and $f'_{n,m}$ the values taken by the variables f_n and f'_n in the above instantiation of the Comp_m predicate, and by $f_{n,m+1}$ and $f'_{n,m+1}$ their values in the above instantiation of the Comp_{m+1} predicate. As previously we have $f_{n,m} = f_{n,m+1}$ for $n < m$, and $f'_{n,m} = f'_{n,m+1}$ for $n \leq m$, and thus Comp_{m+1} correctly says which numbers smaller than m should belong to F . Moreover, it says that $m \notin F$, since $m \in M \cup \{m\}$, which is also correct. It remains to check that the flag counter d computed by Comp_{m+1} actually satisfies $d \geq \log_2 c$. Because $m \in M \cup \{m\}$ we have $f'_{m+1,m+1} = f_{m,m+1}$, and thus

$$\begin{aligned} d &= f'_{m+1,m+1} + d_s = f_{m,m+1} + d_s = f'_{m,m+1} + |F_s \cap \{m\}| + \sum_{i \in I \setminus \{s\}} |G_i \cap \{m\}| + d_s = \\ &= f'_{m,m} + 0 + |\{i \in I \setminus \{s\} \mid c_i > 0\}| + d_s. \end{aligned} \quad (1)$$

On the other hand, we have

$$c = f'_{m,m} + \sum_{i \in I} c_i. \quad (2)$$

In the degenerate case of $c_s = 0$ we have that $c_i = 0$ for all $i \in I$, and by the above $d = c$, thus even more $d \geq \log_2 c$.

Next, suppose that $c_s > 0$. Denote $k = f'_{m,m} + |\{i \in I \mid c_i > 0\}|$. Then (1) gives $d = k - 1 + d_s$, while continuing (2) we have $c \leq f'_{m,m} \cdot c_s + |\{i \in I \mid c_i > 0\}| \cdot c_s = k \cdot c_s$. Altogether we obtain as required:

$$d = k - 1 + d_s \geq k - 1 + \log_2 c_s \geq \log_2 (k \cdot c_s) \geq \log_2 c. \quad \square$$

Corollary 19. *Suppose that $\text{Comp}_m(M; (F', c')) = (F, c)$, where $M \subseteq \{0, \dots, m-1\}$, and $F' = \emptyset \wedge c' = 1$ if $m = 0$, and $F' = \{0\} \wedge c' = 0$ if $m > 0$. Then $\text{Comp}_{m+1}(M \cup \{m\}; (\{0\}, 0)) = (F, d)$ for some d such that $d \geq \log_2 c$.*

Proof. For $m > 0$ this is a direct consequence of Lemma 18 (where the assumption $d_s \geq \log_2 c_s$ is instantiated as $0 \geq \log_2 0$). For $m = 0$ we have $M = \emptyset$, and $\text{Comp}_0(M; (\emptyset, 1)) = (\emptyset, 1)$, while $\text{Comp}_1(M \cup \{0\}; (\{0\}, 0)) = (\emptyset, 1)$; this is fine, since $d = 1 \geq \log_2 1 = \log_2 c$ (for $m = 0$ we could not apply Lemma 18, because the set $F' = \emptyset$ in Comp_m changes into $\{0\}$ in Comp_{m+1}). \square

Corollary 20. *Suppose that $\text{Comp}_m(M; (F', c'), ((F_i, c_i))_{i \in I}) = (F, c)$, where M and F_i are subsets of $\{0, \dots, m-1\}$, and $F' = \emptyset \wedge c' = 1$ if $m = 0$, and $F' = \{0\} \wedge c' = 0$ if $m > 0$. If $s \in I$ is such that $c_i \leq c_s$ for all $i \in I$, and $d_s \geq \log_2 c_s$, and $G_i = F_i \cup \{m \mid c_i > 0\}$ for $i \in I$, then $\text{Comp}_{m+1}(M \cup \{m\}; (\{0\}, 0), (F_s, d_s), ((G_i, 0))_{i \in I \setminus \{s\}}) = (F, d)$ for some d such that $d \geq \log_2 c$.*

Proof. For $m > 0$ this is a direct consequence of Lemma 18 (we notice that $F' \cup \{m \mid c' > 0\} = \{0\}$ and $c' = 0 \leq c_s$). If $m = 0$ and $c_s \geq 1$ we can use Lemma 18 as well. Suppose that $m = 0$ and $c_s = 0$. Then $M = \emptyset$, and $F_i = G_i = \emptyset \wedge c_i = 0$ for all $i \in I$, thus $\text{Comp}_0(M; (\emptyset, 1), ((F_i, c_i))_{i \in I}) = (\emptyset, 1)$, while $\text{Comp}_1(M \cup \{0\}; (\{0\}, 0), (F_s, d_s), ((G_i, 0))_{i \in I \setminus \{s\}}) = (\emptyset, 1 + d_s)$; this is fine, since $d = 1 + d_s \geq \log_2 1 = \log_2 c$. \square

Lemma 21. *Suppose that we can derive $\Gamma \vdash R : (m, F, M, \tau) \triangleright c$, where $\text{ord}(R) \leq m$ and that every full type assigned by Γ to a variable is of order at most m . Then we can also derive $\Gamma \vdash R : (m+1, F, M \cup \{m\}, \tau) \triangleright d$ for some d such that $d \geq \log_2 c$.*

Proof. Denote $\hat{\tau} = (m, F, M, \tau)$ and $\hat{\tau}' = (m+1, F, M \cup \{m\}, \tau)$. The proof is by induction on the structure of the derivation of $\Gamma \vdash R : \hat{\tau} \triangleright c$. We have several cases, depending on the shape of R .

The case of $R = \text{br } P_1 P_2$ follows immediately from the induction assumption, as in Lemma 17.

Suppose that $R = x$ is a variable. The (VAR) rule used in the derivation ensures that $c = 0$ and that $\Gamma(x)$ contains a full type (k, F, M', τ) with $M \upharpoonright_{<k} = M'$. By assumptions of the lemma, $k \leq m$. Then $(M \cup \{m\}) \upharpoonright_{<k} = M'$, and hence the (VAR) rule can equally well derive $\Gamma \vdash R : \hat{\tau}' \triangleright 0$; we have $0 \geq \log_2 0$.

Next, suppose that $R = \lambda x. P$. Then the derivation ends with the (λ) rule, whose premiss is $\Gamma'[x \mapsto T] \vdash P : (m, F, M_\lambda, \tau_\lambda) \triangleright c$, where $\tau = T \rightarrow \tau_\lambda$, and $M = M_\lambda \setminus \bigcup_{(k, F', M', \sigma) \in T} M'$, and $\text{Split}(\Gamma \mid \Gamma')$, and $\Gamma'(x) = \emptyset$. Because τ is a type, the definition of a type ensures that all full types in T are of order $\text{ord}(R) \leq m$. Additionally $\text{ord}(P) \leq \text{ord}(R) \leq m$, so assumptions of the lemma are satisfied for the premiss; the induction assumption gives us a derivation of $\Gamma'[x \mapsto T] \vdash P : (m+1, F, M_\lambda \cup \{m\}, \tau_\lambda) \triangleright d$, where $d \geq \log_2 c$. Again because all full types in T are of order $\text{ord}(R) \leq m$ (and hence $M' \subseteq \{0, \dots, m-1\}$ for all $(k, F', M', \sigma) \in T$) we have $M \cup \{m\} = M_\lambda \cup \{m\} \setminus \bigcup_{(k, F', M', \sigma) \in T} M'$. Thus after applying back the (λ) rule we obtain a derivation of $\Gamma \vdash R : \hat{\tau}' \triangleright d$.

Next, suppose that $R = a$ (where a is a symbol of rank 0). We have $\tau = o$. The conditions of the (CON) rule are $\text{Split}(\Gamma \mid \varepsilon)$ and $\text{Comp}_m(M; (F', c')) = (F, c)$, where $F' = \emptyset$ and $c' = 1$ if $m = 0$, and $F' = \{0\}$ and $c' = 0$ if $m > 0$. By Corollary 19, $\text{Comp}_{m+1}(M \cup \{m\}; (\{0\}, 0)) = (F, d)$ for some d such that $d \geq \log_2 c$. We can use the (CON) rule to derive $\Gamma \vdash R : \hat{\tau}' \triangleright d$.

Next, suppose that $R = a P_1 \dots P_r$ for $a \neq \text{br}$ and $r > 0$. We have $\tau = o$. Denote $I = \{1, \dots, r\}$. Let $\Gamma_i \vdash P_i : (m, F_i, M_i, o) \triangleright c_i$ for $i \in I$ be the premisses of the final (CON) rule, and let $s \in I$ be such that $c_s \geq c_i$ for all $i \in I$. Using the induction assumption we derive $\Gamma_s \vdash P_s : (m+1, F_s, M_s \cup \{m\}, o) \triangleright d_s$ for some d_s such that $d_s \geq \log_2 c_s$, while for $i \in I \setminus \{s\}$ we use Lemma 17 to derive $\Gamma_i \vdash P_i : (m+1, G_i, M_i, o) \triangleright 0$, where $G_i = F_i \cup \{m \mid c_i > 0\}$. We want to apply back the (CON) rule. The condition $M = M' \uplus M_1 \uplus \dots \uplus M_r$ updates accordingly: m is added to M and to M_s . The conditions $(r > 0) \Rightarrow (M' = \emptyset)$ and $\text{Split}(\Gamma \mid \Gamma_1, \dots, \Gamma_r)$ remain unchanged. We know that $\text{Comp}_m(M; (F', c'), ((F_i, c_i))_{i \in I}) = (F, c)$, where $F' = \emptyset$ and $c' = 1$ if $m = 0$, and $F' = \{0\}$ and $c' = 0$ if $m > 0$; by Corollary 20 this implies $\text{Comp}_{m+1}(M \cup \{m\}; (\{0\}, 0), (F_s, d_s), ((G_i, 0))_{i \in I \setminus \{s\}}) = (F, d)$ for some d such that $d \geq \log_2 c$, as required.

Finally, suppose that $R = P Q$. Let $\Gamma' \vdash P : (m, F', M', T \rightarrow \tau) \triangleright c'$ and $\Gamma_i \vdash Q : (m, F_i, M_i, \tau_i) \triangleright c_i$ for each $i \in I$ be the premisses of the final (@) rule, where $T = \{(\text{ord}(P), F_i \upharpoonright_{<\text{ord}(P)}, M_i \upharpoonright_{<\text{ord}(P)}, \tau_i) \mid i \in I\}$. A condition of the (@) rule implies that $\text{ord}(Q) < \text{ord}(P) \leq m$, so we can use the induction assumption for these premisses. We know that $\text{Comp}_m(M; (F', c'), ((F_i \upharpoonright_{\geq \text{ord}(P)}, c_i))_{i \in I}) = (F, c)$. Denote $G' = F' \cup \{m \mid c' > 0\}$ and $G_i = F_i \cup \{m \mid c_i > 0\}$ for $i \in I$. Notice that $G_i \upharpoonright_{\geq \text{ord}(P)} = F_i \upharpoonright_{\geq \text{ord}(P)} \cup \{m \mid c_i > 0\}$. We have two subcases. If $c' \geq c_i$ for all $i \in I$, then using the induction assumption we derive $\Gamma' \vdash P : (m+1, F', M' \cup \{m\}, T \rightarrow \tau) \triangleright d'$ for some d' such that $d' \geq \log_2 c'$, and using Lemma 17 we derive $\Gamma_i \vdash Q : (m+1, G_i, M_i, \tau_i) \triangleright 0$ for each $i \in I$. Lemma 18 then implies that $\text{Comp}_{m+1}(M \cup \{m\}; (F', d'), ((G_i \upharpoonright_{\geq \text{ord}(P)}, 0))_{i \in I}) = (F, d)$ for some d such that $d \geq \log_2 c$. Otherwise, we choose $s \in I$

such that $c_s \geq c_i$ for all $i \in I$ (and $c_s \geq c'$); using the induction assumption we derive $\Gamma_s \vdash Q : (m+1, F_s, M_s \cup \{m\}, \tau_s) \triangleright d_s$ for some d_s such that $d_s \geq \log_2 c_s$, and using Lemma 17 we derive $\Gamma' \vdash P : (m+1, G', M', T \rightarrow \tau) \triangleright 0$ and $\Gamma_i \vdash Q : (m+1, G_i, M_i, \tau_i) \triangleright 0$ for each $i \in I \setminus \{s\}$. Lemma 18 then implies that $\text{Comp}_{m+1}(M \cup \{m\}; (F_s \upharpoonright_{\geq \text{ord}(P)}, d_s), (G', 0), ((G_i \upharpoonright_{\geq \text{ord}(P)}, 0))_{i \in I \setminus \{s\}}) = (F, d)$ for some d such that $d \geq \log_2 c$. In both cases we apply back the $(@)$ rule to the obtained type judgments. The condition $M = M' \uplus \biguplus_{i \in I} M_i$ is updated accordingly: m is added to M and either to M' or to M_s . The condition $\text{Split}(\Gamma \mid \Gamma', (\Gamma_i)_{i \in I})$ remains unchanged, and $\text{ord}(P) \leq m+1$ holds since we even have $\text{ord}(P) \leq m$. We also have that $G_i \upharpoonright_{< \text{ord}(P)} = F_i \upharpoonright_{< \text{ord}(P)}$ and $(M_s \cup \{m\}) \upharpoonright_{< \text{ord}(P)} = M_s$, so the type required in a premiss for P is indeed $T \rightarrow o$. \square

Lemma 4 says that if we can derive $\varepsilon \vdash R : (m, \emptyset, \{0, \dots, m-1\}, o) \triangleright c$, then we can also derive $\varepsilon \vdash R : (m+1, \emptyset, \{0, \dots, m\}, o) \triangleright d$ for some d such that $d \geq \log_2 c$. This is just a special case of Lemma 21.

C Proof of Lemma 5

Below, by $\text{Mk}(\hat{\tau})$ we denote the set of marker orders of the full type $\hat{\tau}$, i.e., $\text{Mk}(\hat{\tau}) = M$ if $\hat{\tau} = (m, F, M, \tau)$. We extend this notation to sets of full types: $\text{Mk}(T) = \bigcup_{\hat{\tau} \in T} \text{Mk}(\hat{\tau})$, and to type environments: $\text{Mk}(\Gamma) = \bigcup_x \text{Mk}(\Gamma(x))$, where x ranges over all variables.

Lemma 22. *Suppose that we can derive $\Gamma \vdash R : \hat{\tau} \triangleright c$, and x is not free in R . Then for $\Sigma = \Gamma[x \mapsto \emptyset]$ we can also derive $\Sigma \vdash R : \hat{\tau} \triangleright c$, and $\text{Split}(\Gamma \mid \Sigma)$ holds.*

Proof. Because x is not free in R , all full types assigned to x by Γ are discarded somewhere in the derivation of $\Gamma \vdash R : \hat{\tau} \triangleright c$. On the one hand, this means that the set of marker orders in all these full types is empty, and thus $\text{Split}(\Gamma \mid \Sigma)$ holds. On the other hand, we can remove these full types from type environments in the derivation, which results in a derivation of $\Sigma \vdash R : \hat{\tau} \triangleright c$. \square

Lemma 23. *Suppose that we can derive $\Gamma \vdash P : \hat{\tau} \triangleright c$. If $\text{Split}(\Gamma' \mid \Gamma)$ holds, then we can also derive $\Gamma' \vdash P : \hat{\tau} \triangleright c$.*

Proof. Induction on the structure of a fixed derivation of $\Gamma \vdash P : \hat{\tau} \triangleright c$. If the last rule is (BR), we change Γ to Γ' in the premiss of the rule using the induction assumption, and we obtain a derivation of $\Gamma' \vdash P : \hat{\tau} \triangleright c$. In every other rule we can change Γ to Γ' only in the conclusion. \square

Lemma 24. *Suppose that we can derive $\Gamma \vdash R : \hat{\tau} \triangleright c$. Then $\text{Mk}(\Gamma) \subseteq \text{Mk}(\hat{\tau})$, and $\text{Mk}(\Gamma(x)) \cap \text{Mk}(\Gamma(y)) = \emptyset$ for all variables x, y with $x \neq y$.*

Proof. Fix some derivation of $\Gamma \vdash R : \hat{\tau} \triangleright c$; the proof is by induction on the structure of this derivation. We analyze the shape of R .

Suppose first that $R = x$. The (VAR) rule says that $\text{Split}(\Gamma, \varepsilon[x \mapsto T])$ holds for T such that $\text{Mk}(T) \subseteq \text{Mk}(\hat{\tau})$. Thus $\text{Mk}(\Gamma(x)) \subseteq \text{Mk}(\hat{\tau})$ and $\text{Mk}(\Gamma(y)) = \emptyset$ for all variables y with $y \neq x$.

In the case when $R = \text{br } P_1 P_2$ the thesis follows immediately from the induction assumption applied to the premiss of the final (BR) rule.

Next, suppose that $R = \lambda z. P$. Let $\Gamma'[z \mapsto T] \vdash P : \hat{\tau}' \triangleright c$ with $\Gamma'(z) = \emptyset$ be the premiss of the final (λ) rule. By the conditions of the rule we have $\text{Mk}(\hat{\tau}) = \text{Mk}(\hat{\tau}') \setminus \text{Mk}(T)$ and $\text{Split}(\Gamma \mid \Gamma')$. The latter condition implies that $\text{Mk}(\Gamma(z)) = \emptyset$. For every variable x other than z the induction assumption ensures that $\text{Mk}(\Gamma'[z \mapsto T](x)) \subseteq \text{Mk}(\hat{\tau}')$ and that $\text{Mk}(\Gamma'[z \mapsto T](x)) \cap \text{Mk}(\Gamma'[z \mapsto T](z)) = \emptyset$. Since $\text{Mk}(\Gamma(x)) =$

$\text{Mk}(\Gamma'[z \mapsto T](x))$ and $\text{Mk}(\Gamma'[z \mapsto T](z)) = \text{Mk}(T)$ we obtain $\text{Mk}(\Gamma(x)) \subseteq \text{Mk}(\hat{\tau}') \setminus \text{Mk}(T) = \text{Mk}(\hat{\tau})$. For any two variables x, y with $z \neq x \neq y \neq z$ we have $\text{Mk}(\Gamma'[z \mapsto T](x)) \cap \text{Mk}(\Gamma'[z \mapsto T](y)) = \emptyset$ by the induction assumption, and thus $\text{Mk}(\Gamma(x)) \cap \text{Mk}(\Gamma(y)) = \emptyset$.

Next, suppose that $R = aP_1 \dots P_r$. Let $\Gamma_i \vdash P_i : \hat{\tau}_i \triangleright c_i$ for $i \in \{1, \dots, r\}$ be the premisses of the final (CON) rule. Consider some variable x , and some $k \in \text{Mk}(\Gamma(x))$. Because of the condition $\text{Split}(\Gamma \mid \Gamma_1, \dots, \Gamma_r)$ of the (CON) rule we have $k \in \text{Mk}(\Gamma_i(x))$ for some $i \in \{1, \dots, r\}$; then $k \in \text{Mk}(\hat{\tau}_i)$ by the induction assumption $\text{Mk}(\Gamma_i(x)) \subseteq \text{Mk}(\hat{\tau}_i)$, and thus $k \in \text{Mk}(\hat{\tau})$ by the condition $\text{Mk}(\hat{\tau}) = \text{Mk}(\hat{\tau}_1) \uplus \dots \uplus \text{Mk}(\hat{\tau}_r)$ of the (CON) rule. Suppose that we also have $k \in \text{Mk}(\Gamma(y))$ for some variable y other than x . Then $k \in \text{Mk}(\Gamma_j(y)) \subseteq \text{Mk}(\hat{\tau}_j)$ for some $j \in \{1, \dots, r\}$; we cannot have $j = i$ by the induction assumption ($\text{Mk}(\Gamma_i(x))$ and $\text{Mk}(\Gamma_i(y))$ are disjoint), and we cannot have $j \neq i$ because $\text{Mk}(\hat{\tau}_i)$ and $\text{Mk}(\hat{\tau}_j)$ are disjoint.

The case when $R = PQ$ is completely analogous to the previous one. \square

Lemma 25. *Suppose that we can derive $\Gamma \vdash R[S/x] : \hat{\tau} \triangleright c$. Then, for some finite set I , we can derive $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i$ for every $i \in I$, and $\Lambda^x \vdash R : \hat{\tau} \triangleright e$, where $\Lambda^x = \Lambda[x \mapsto \{\hat{\sigma}_i \mid i \in I\}]$ with $\Lambda(x) = \emptyset$, and $\text{Split}(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ holds, and $c = e + \sum_{i \in I} d_i$, and $\text{Mk}(\hat{\sigma}_i) \cap \text{Mk}(\hat{\sigma}_j) = \emptyset$ for $i, j \in I$ if $i \neq j$.*

Proof. Fix some derivation of $\Gamma \vdash R[S/x] : \hat{\tau} \triangleright c$; the proof is by induction on the structure of this derivation.

One possibility is that x is not free in R . Then we can take $I = \emptyset$, and $\Lambda^x = \Lambda = \Gamma[x \mapsto \emptyset]$, and $e = c$. Because x is not free in $R = R[S/x]$, by Lemma 22 applied to the original derivation, we know that $\Lambda^x \vdash R : \hat{\tau} \triangleright e$ can be derived, and that $\text{Split}(\Gamma \mid \Lambda)$ holds. Because $I = \emptyset$, the condition concerning disjointness of $\text{Mk}(\hat{\sigma}_i)$ becomes trivial.

In the sequel we assume that x is free in R . We analyze the shape of R .

Suppose first that $R = x$. Then we take $I = \{1\}$, and $(\Sigma_1, \hat{\sigma}_1, d_1) = (\Gamma, \hat{\tau}, c)$, and $\Lambda = \varepsilon$, and $e = 0$. Obviously $\Lambda(x) = \emptyset$, and $\text{Split}(\Gamma \mid \Lambda, \Sigma_1)$ holds, and $c = e + d_1$. We can derive $\Sigma_1 \vdash S : \hat{\sigma}_1 \triangleright d_1$ by assumption, and $\Lambda^x \vdash R : \hat{\tau} \triangleright e$ using the (VAR) rule, where $\Lambda^x = \Lambda[x \mapsto \{\hat{\sigma}_1\}]$.

Next, suppose that $R = \text{br } P_1 P_2$. Then our derivation ends with the (BR) rule, whose premiss is $\Gamma \vdash P_k[S/x] : \hat{\tau} \triangleright c$, for some $k \in \{1, 2\}$. The induction assumption applied to this premiss gives us a derivation of $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i$ for every $i \in I$, and of $\Lambda^x \vdash P_k : \hat{\tau} \triangleright e$, where appropriate conditions hold. By applying back the (BR) rule to the latter type judgment, we obtain a derivation of $\Lambda^x \vdash R : \hat{\tau} \triangleright e$ as required.

Next, suppose that $R = \lambda y.P$. We have $y \neq x$, and, as always during a substitution, we assume (by performing α -conversion) that y is not free in S . The original derivation ends with the (λ) rule, whose premiss is $\Gamma'[y \mapsto T] \vdash P[S/x] : \hat{\tau}' \triangleright c$ with $\Gamma'(y) = \emptyset$. We apply the induction assumption to this premiss, and we obtain a derivation of $\Sigma_i[y \mapsto T_i] \vdash S : \hat{\sigma}_i \triangleright d_i$ for every $i \in I$, and of $\Lambda^x[y \mapsto T'] \vdash P : \hat{\tau}' \triangleright e$, where $\Lambda^x = \Lambda[x \mapsto \{\hat{\sigma}_i \mid i \in I\}]$ with $\Lambda(x) = \Lambda(y) = \Sigma_i(y) = \emptyset$, and $\text{Split}(\Gamma'[y \mapsto T] \mid \Lambda[y \mapsto T'], (\Sigma_i[y \mapsto T_i])_{i \in I})$ holds, and $c = e + \sum_{i \in I} d_i$, and $\text{Mk}(\hat{\sigma}_i) \cap \text{Mk}(\hat{\sigma}_j) = \emptyset$ for $i, j \in I$ if $i \neq j$. Together with $\text{Split}(\Gamma \mid \Gamma')$ the above implies that $\text{Split}(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ holds. Because y is not free in S , Lemma 22 implies that for every $i \in I$ we can derive $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i$ (instead of $\Sigma_i[y \mapsto T_i] \vdash S : \hat{\sigma}_i \triangleright d_i$) and that $\text{Split}(\Sigma_i[y \mapsto T_i] \mid \Sigma_i)$ holds. Thus $\text{Mk}(T_i) = \emptyset$ for all $i \in I$, and hence $\text{Mk}(T \setminus T') = \emptyset$; simultaneously $T' \subseteq T$, which implies that $\text{Split}(\Lambda^x[y \mapsto T] \mid \Lambda^x[y \mapsto T'])$ holds. In consequence, by Lemma 23 applied to $\Lambda^x[y \mapsto T'] \vdash P : \hat{\tau}' \triangleright e$ we can derive $\Lambda^x[y \mapsto T] \vdash P : \hat{\tau}' \triangleright e$. To the latter type judgment we apply again the (λ) rule, which gives $\Lambda^x \vdash R : \hat{\tau} \triangleright e$.

Another possibility is that $R = aP_1 \dots P_r$. Then the original derivation ends with the (CON) rule, whose premisses are $\Gamma_j \vdash P_j[S/x] : \hat{\tau}_j \triangleright c_j$ for $j \in \{1, \dots, r\}$. We apply the induction assumption to these

premisses. Assuming w.l.o.g. that the resulting sets I_j are disjoint, and taking $I = \bigcup_{j=1}^r I_j$, we obtain a derivation of $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i$ for every $i \in I$, and of $\Lambda_j^x \vdash P_j : \hat{\tau}_j \triangleright e_j$ for every $j \in \{1, \dots, r\}$, where, for every $j \in \{1, \dots, r\}$, we have $\Lambda_j^x = \Lambda_j[x \mapsto \{\hat{\sigma}_i \mid i \in I_j\}]$ with $\Lambda_j(x) = \emptyset$, and $\text{Split}(\Gamma_j \mid \Lambda_j, (\Sigma_i)_{i \in I_j})$ holds, and $c_j = e_j + \sum_{i \in I_j} d_i$, and $\text{Mk}(\hat{\sigma}_i) \cap \text{Mk}(\hat{\sigma}_{i'}) = \emptyset$ for $i, i' \in I_j$ if $i \neq i'$. By Lemma 24 we have $\text{Mk}(\hat{\sigma}_i) \subseteq \text{Mk}(\hat{\tau}_j)$ for $i \in I_j$, $j \in J$. Since $\text{Mk}(\hat{\tau}_j) \cap \text{Mk}(\hat{\tau}_{j'}) = \emptyset$ for $j, j' \in J$ with $j \neq j'$ by a side condition of the (CON) rule, this implies that $\text{Mk}(\hat{\sigma}_i) \cap \text{Mk}(\hat{\sigma}_{i'}) = \emptyset$ for all $i, i' \in I$ with $i \neq i'$. Recalling the side condition $\text{Split}(\Gamma \mid (\Gamma_j)_{j \in \{1, \dots, r\}})$ of the (CON) rule, we observe that $\text{Split}(\Gamma \mid (\Lambda_j)_{j \in \{1, \dots, r\}}, (\Sigma_i)_{i \in I})$ holds. Define Λ by taking $\Lambda(z) = \bigcup_{j=1}^r \Lambda_j(z)$ for every variable z . We then have $\text{Split}(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ and $\text{Split}(\Lambda \mid (\Lambda_j)_{j \in \{1, \dots, r\}})$, as well as $\text{Split}(\Lambda^x \mid (\Lambda_j^x)_{j \in \{1, \dots, r\}})$. Another side condition of the (CON) rule says that $\text{Comp}_m(M; (F', c'), (F_1, c_1), \dots, (F_r, c_r)) = (F, c)$ for appropriate arguments M, F, F', c', F_j . Taking $e = c + \sum_{j=1}^r (e_j - c_j)$ we also have that $\text{Comp}_m(M; (F', c'), (F_1, e_1), \dots, (F_r, e_r)) = (F, e)$. Having all this, we can apply the (CON) rule again, deriving $\Lambda^x \vdash R : \hat{\tau} \triangleright e$. Simultaneously we observe that $c = e + \sum_{i \in I} d_i$.

Finally, suppose that $R = PQ$. This case is very similar to the previous one. The original derivation ends with the (@) rule, whose premisses are $\Gamma_0 \vdash P[S/x] : \hat{\tau}_0 \triangleright c_0$ and $\Gamma_j \vdash Q[S/x] : \hat{\tau}_j \triangleright c_j$ for $j \in J$, where we assume that $0 \notin J$. We apply the induction assumption to all these premisses. Assuming w.l.o.g. that the resulting sets I_j are disjoint, and taking $I = \bigcup_{j \in \{0\} \cup J} I_j$, we obtain a derivation of $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i$ for every $i \in I$, and of $\Lambda_0^x \vdash P : \hat{\tau}_0 \triangleright e_0$, and of $\Lambda_j^x \vdash Q : \hat{\tau}_j \triangleright e_j$ for every $j \in J$, where, for every $j \in \{0\} \cup J$, we have $\Lambda_j^x = \Lambda_j[x \mapsto \{\hat{\sigma}_i \mid i \in I_j\}]$ with $\Lambda_j(x) = \emptyset$, and $\text{Split}(\Gamma_j \mid \Lambda_j, (\Sigma_i)_{i \in I_j})$ holds, and $c_j = e_j + \sum_{i \in I_j} d_i$, and $\text{Mk}(\hat{\sigma}_i) \cap \text{Mk}(\hat{\sigma}_{i'}) = \emptyset$ for $i, i' \in I_j$ if $i \neq i'$. By Lemma 24 we have $\text{Mk}(\hat{\sigma}_i) \subseteq \text{Mk}(\hat{\tau}_j)$ for $i \in I_j$, $j \in \{0\} \cup J$. Since $\text{Mk}(\hat{\tau}_j) \cap \text{Mk}(\hat{\tau}_{j'}) = \emptyset$ for $j, j' \in \{0\} \cup J$ with $j \neq j'$ by a side condition of the (CON) rule, this implies that $\text{Mk}(\hat{\sigma}_i) \cap \text{Mk}(\hat{\sigma}_{i'}) = \emptyset$ for all $i, i' \in I$ with $i \neq i'$. Recalling the side condition $\text{Split}(\Gamma \mid (\Gamma_j)_{j \in \{0\} \cup J})$ of the (@) rule, we observe that $\text{Split}(\Gamma \mid (\Lambda_j)_{j \in \{0\} \cup J}, (\Sigma_i)_{i \in I})$ holds. Define Λ by taking $\Lambda(z) = \bigcup_{j \in \{0\} \cup J} \Lambda_j(z)$ for every variable z . We then have $\text{Split}(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ and $\text{Split}(\Lambda \mid (\Lambda_j)_{j \in \{0\} \cup J})$, as well as $\text{Split}(\Lambda^x \mid (\Lambda_j^x)_{j \in \{0\} \cup J})$. Another side condition of the (CON) rule says that $\text{Comp}_m(M; ((G_j, c_j))_{j \in \{0\} \cup J}) = (F, c)$ for appropriate sets M, F, G_j . Taking $e = c + \sum_{j \in \{0\} \cup J} (e_j - c_j)$ we also have that $\text{Comp}_m(M; ((G_j, c_j))_{j \in \{0\} \cup J}) = (F, e)$. Thus we can apply the (@) rule again, deriving $\Lambda^x \vdash R : \hat{\tau} \triangleright e$. Simultaneously we observe that $c = e + \sum_{i \in I} d_i$. \square

Lemma 26. Suppose that $\text{Comp}_m(M; (F, e), ((\emptyset, d_i))_{i \in I}) = (G, c)$, where $F \cap M = \emptyset$ and $F, M \subseteq \{0, \dots, m-1\}$. Then $G = F$ and $c = e + \sum_{i \in I} d_i$.

Proof. Consider the numbers f_n and f'_n appearing in the definition of the Comp_m predicate. Looking at them consecutively for $n = 0, \dots, m$ we notice that $f'_n = 0$ and $f_n = |F \cap \{n\}|$. Indeed, $f'_n = 0$ implies $f_n = |F \cap \{n\}|$, and if $n-1 \notin M$, we have $f'_n = 0$, while if $n-1 \in M$, we have $f'_n = f_{n-1} = |F \cap \{n-1\}| = 0$, because $F \cap M = \emptyset$. Then $G = \{n \in \{0, \dots, m-1\} \mid f_n > 0 \wedge n \notin M\} = F$ (again because $F \cap M = \emptyset$), and $c = f'_m + e + \sum_{i \in I} d_i = e + \sum_{i \in I} d_i$. \square

Proof of Lemma 5. Recall that we are given a derivation of $\Gamma \vdash Q : \hat{\tau} \triangleright c$ with $\text{ord}(\hat{\tau}) = m$, and a β -reduction $P \rightarrow_\beta Q$ that is of order m , and our goal is to derive $\Gamma \vdash P : \hat{\tau} \triangleright c$.

Suppose first that $P = (\lambda x.R)S$ and $Q = R[S/x]$, where $\text{ord}(\lambda x.R) = m$. From Lemma 25 we obtain a derivation of $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i$ for every $i \in I$ (for some set I), and a derivation of $\Lambda[x \mapsto T] \vdash R : \hat{\tau} \triangleright e$, where $T = \{\hat{\sigma}_i \mid i \in I\}$, and $\Lambda(x) = \emptyset$, and $\text{Split}(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ holds, and $c = e + \sum_{i \in I} d_i$, and $\text{Mk}(\hat{\sigma}_i) \cap \text{Mk}(\hat{\sigma}_j) = \emptyset$ for $i, j \in I$ if $i \neq j$. Let us write $\hat{\tau} = (m, F, M, \tau)$, and $\hat{\sigma}_i = (m, F_i, M_i, \sigma_i)$. To the type judgment $\Lambda[x \mapsto T] \vdash R : \hat{\tau} \triangleright e$ we apply the (λ) rule, deriving $\Lambda \vdash \lambda x.R : (m, F, M \setminus \bigcup_{i \in I} M_i, T \rightarrow \tau) \triangleright e$.

To this type judgment, and to $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i$ for $i \in I$, we want to apply the (@) rule. By definition of a full type, the sets F_i and M_i may contain only numbers smaller than m (since $\hat{\sigma}_i$ is a full type).

Recalling that $\text{ord}(\lambda x.R) = m$ we have that the type $\{(\text{ord}(\lambda x.R), F_i \upharpoonright_{<\text{ord}(\lambda x.R)}, M_i \upharpoonright_{<\text{ord}(\lambda x.R)}, \sigma_i) \mid i \in I\} \rightarrow \tau$ that we have to derive for $\lambda x.R$ is indeed $T \rightarrow \tau$. The conditions $M = (M \setminus \bigcup_{i \in I} M_i) \uplus \biguplus_{i \in I} M_i$, and $\text{ord}(\lambda x.R) \leq m$, and $\text{Split}(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ follow from what we have. Notice that the sets $F_i \upharpoonright_{\geq \text{ord}(\lambda x.R)}$ are empty, and that $F \cap M = \emptyset$ by definition of a full type ($\hat{\tau} = (m, F, M, \tau)$ is a full type), and hence $\text{Comp}_m(M; (F, e), ((F_i \upharpoonright_{\geq \text{ord}(\lambda x.R)}, d_i))_{i \in I}) = (F, c)$ by Lemma 26. Thus the (@) rule can be applied; it derives $\Gamma \vdash P : \hat{\tau} \triangleright c$.

It remains to consider the general situation: the redex involved in the β -reduction $P \rightarrow_\beta Q$ is located somewhere deeper in P . Then the proof is by a trivial induction on the depth of this redex. Formally, we have several cases depending on the shape of P , but let us consider only a representative example: suppose that $P = TU$ and $Q = TV$ with $U \rightarrow_\beta V$. In the derivation $\Gamma \vdash Q : \hat{\tau} \triangleright c$ we apply the induction assumption to those premisses of the final (@) rule that concern the subterm V , and we obtain type judgments in which V is replaced by U . We can apply the (@) rule to them, and to the premiss talking about T , and derive $\Gamma \vdash P : \hat{\tau} \triangleright c$. \square

D Proof of Lemma 6

In order to enable an inductive proof of Lemma 6, we need to strengthen slightly its statement, and consider also λ -terms of order m . We say that a full type $(m, F, M, T_1 \rightarrow \dots \rightarrow T_k \rightarrow o)$ is $(m-1)$ -clear if $m-1 \notin M \cup \bigcup_{i=1}^k \text{Mk}(T_i)$.

Lemma 27. *If $\hat{\tau} \in \mathcal{F}_m^\alpha$ for $\text{ord}(\alpha) < m$, and $m-1 \notin \text{Mk}(\hat{\tau})$, then $\hat{\tau}$ is $(m-1)$ -clear.*

Proof. If $\alpha = \alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow o$, we can write $\hat{\tau} = (m, F, M, T_1 \rightarrow \dots \rightarrow T_k \rightarrow o)$. For $i \in \{1, \dots, k\}$ by definition we have $T_i \subseteq \mathcal{F}_{m_i}^{\alpha_i}$ for $m_i = \text{ord}(\alpha_i \rightarrow \dots \rightarrow \alpha_k \rightarrow o)$, and hence $\text{Mk}(T_i) \subseteq \{0, \dots, m_i - 1\}$; because $\text{ord}(\alpha_i \rightarrow \dots \rightarrow \alpha_k \rightarrow o) \leq \text{ord}(\alpha) < m$, we have $m-1 \notin T_i$. Thus $\hat{\tau}$ is $(m-1)$ -clear. \square

Lemma 28. *If we can derive $\Gamma \vdash R : \hat{\tau} \triangleright c$, where $\text{ord}(\hat{\tau}) = m > 0$ and $\hat{\tau}$ is $(m-1)$ -clear, then $c = 0$.*

Proof. Fix some derivation of $\Gamma \vdash R : \hat{\tau} \triangleright c$. The proof is by induction on the structure of this derivation. Let us write $\hat{\tau} = (m, F, M, \tau)$. We have several cases depending on the shape of R .

If $R = x$, the (VAR) rule ensures that the flag counter c is 0.

If $R = \text{br } P_1 P_2$, then above the final (BR) rule we have a premiss $\Gamma \vdash P_i : \hat{\tau} \triangleright c$ for some $i \in \{1, 2\}$; the induction assumption used for this premiss implies $c = 0$.

Suppose that $R = \lambda x.P$. Then above the final (λ) rule we have a premiss $\Gamma' \vdash P : (m, F, M', \tau') \triangleright c$, where $\tau = T \rightarrow \tau'$ and $M = M' \setminus \text{Mk}(T)$. Because $\hat{\tau}$ is $(m-1)$ -clear, we have $m-1 \notin M \cup \text{Mk}(T)$, hence also $m-1 \notin M'$; thus (m, F, M', τ') is $(m-1)$ -clear. The induction assumption applied to our premiss implies $c = 0$.

Next, suppose that $R = aP_1 \dots P_r$. Let $\Gamma_i \vdash P_i : (m, F_i, M_i, o) \triangleright c_i$ for $i \in \{1, \dots, r\}$ be the premisses of the final (CON) rule. For all $i \in \{1, \dots, r\}$ a side condition of this rule says that $M_i \subseteq M$, so the full type (m, F_i, M_i, o) is $(m-1)$ -clear, and hence $c_i = 0$ by the induction assumption. We know that $\text{Comp}_m(M; (\{0\}, 0), (F_1, c_1), \dots, (F_r, c_r)) = (F, c)$ (recall that $m > 0$). The number f'_m considered in the definition of Comp_m is 0 in our case of $m-1 \notin M$, and thus we have $c = f'_m + \sum_{i=1}^r c_i = 0$.

Finally, suppose that $R = PQ$. Let $\Gamma' \vdash P : (m, F', M', T \rightarrow \tau) \triangleright c'$ and $\Gamma_i \vdash Q : (m, F_i, M_i, \tau_i) \triangleright c_i$ for $i \in I$ be the premisses of the final (@) rule. A side condition of the rule says that $M' \subseteq M$ and $M_i \subseteq M$ for $i \in I$, so $m-1$ does not belong to these sets. By definition the marker sets in full types in T are subsets of some M_i , so $m-1 \notin \text{Mk}(T)$, and hence $(m, F', M', T \rightarrow \tau)$ is $(m-1)$ -clear. On the other hand $\text{ord}(Q) < \text{ord}(P) \leq m$, so (m, F, M_i, τ_i) for $i \in I$ are $(m-1)$ -clear by Lemma 27. Thus the induction

assumption can be used for all premisses; it says that $c' = 0$ and $c_i = 0$ for all $i \in I$. Because $\text{Comp}(M; (F', c'), ((F_i, c_i))_{i \in I}) = (F, c)$, and $m - 1 \notin M$, we obtain $c = 0$ (as in the previous case). \square

Proof of Lemma 6. Recall that in this lemma we are given a derivation of $\Gamma \vdash P : \hat{\tau} \triangleright c$, where $m - 1 \notin \text{Mk}(\hat{\tau})$ and $\text{ord}(P) \leq m - 1$ for $m = \text{ord}(\hat{\tau})$, and we have to prove that $c = 0$. Lemma 27 implies that $\hat{\tau}$ is $(m - 1)$ -clear, and thus $c = 0$ by Lemma 28 (we have $m > 0$ because $\text{ord}(P) \leq m - 1$). \square

E Proof of Lemma 7

Let us formalize the notion of counting (@) rules of order m in a derivation. We use here *extended type judgments* of the form $\Gamma \vdash P : \hat{\tau} \triangleright c \succ n$, where $\Gamma \vdash P : \hat{\tau} \triangleright c$ is a type judgment, and $n \in \mathbb{N}$. The number n is called *application counter*. The meaning is that $\Gamma \vdash P : \hat{\tau} \triangleright c$ can be derived by a derivation in which the (@) rule of order $\text{ord}(\hat{\tau})$ is used n times. Formally, we lift our type system so that it can derive extended type judgments, as follows. If, in the original type system, using premisses $\Gamma_i \vdash P_i : \hat{\tau}_i \triangleright c_i$ for $i \in I$ some rule could derive $\Gamma \vdash R : \hat{\tau} \triangleright c$, then

- if $R = PQ$ with $\text{ord}(P) = \text{ord}(\hat{\tau})$, then using premisses $\Gamma_i \vdash P_i : \hat{\tau}_i \triangleright c_i \succ n_i$ for $i \in I$ we can derive $\Gamma \vdash R : \hat{\tau} \triangleright c \succ 1 + \sum_{i \in I} n_i$;
- otherwise (i.e., when $R = PQ$ with $\text{ord}(P) \neq \text{ord}(\hat{\tau})$, or R does not start with an application) using premisses $\Gamma_i \vdash P_i : \hat{\tau}_i \triangleright c_i \succ n_i$ for $i \in I$ we can derive $\Gamma \vdash R : \hat{\tau} \triangleright c \succ \sum_{i \in I} n_i$.

It should be clear that we can derive $\Gamma \vdash R : \hat{\tau} \triangleright c$ if and only if we can derive $\Gamma \vdash R : \hat{\tau} \triangleright c \succ n$ for some $n \in \mathbb{N}$. We also notice that in proofs of Lemmata 22 and 23 the number of (@) rules of the maximal order (and, more generally, the shape of a derivation) remain unchanged. Thus we can restate these lemmata as follows.

Lemma 29. *Suppose that we can derive $\Gamma \vdash R : \hat{\tau} \triangleright c \succ 0$, and x is not free in R . Then for $\Sigma = \Gamma[x \mapsto \emptyset]$ we can also derive $\Sigma \vdash R : \hat{\tau} \triangleright c \succ 0$, and $\text{Split}(\Gamma \mid \Sigma)$ holds.*

Lemma 30. *Suppose that we can derive $\Gamma \vdash P : \hat{\tau} \triangleright c \succ 0$. If $\text{Split}(\Gamma' \mid \Gamma)$ holds, then we can also derive $\Gamma' \vdash P : \hat{\tau} \triangleright c \succ 0$.*

As in the proof of Lemma 5 we start with a lemma describing substitution.

Lemma 31. *Suppose that we can derive $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i \succ 0$ for every $i \in I$ (for some finite set I), and $\Lambda^x \vdash R : \hat{\tau} \triangleright e \succ 0$, where $\Lambda^x = \Lambda[x \mapsto \{\hat{\sigma}_i \mid i \in I\}]$ with $\Lambda(x) = \emptyset$, and $\text{Split}(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ holds, and $\text{ord}(S) < \text{ord}(\hat{\tau})$, and $\text{ord}(\hat{\sigma}_i) = \text{ord}(\hat{\tau})$ for all $i \in I$. Then we can derive $\Gamma \vdash R[S/x] : \hat{\tau} \triangleright e + \sum_{i \in I} d_i \succ 0$.*

Proof. We start by observing the following property, denoted (\diamond):

If $\text{Mk}(\hat{\sigma}_i) = \emptyset$ for some $i \in I$, then $d_i = 0$ and $\text{Mk}(\Sigma_i) = \emptyset$.

Indeed, $\text{Mk}(\Sigma_i) \subseteq \text{Mk}(\sigma_i) = \emptyset$ follows from Lemma 24, while $d_i = 0$ follows from Lemma 6 if we recall that $\text{ord}(S) < \text{ord}(\hat{\tau}) = \text{ord}(\hat{\sigma}_i)$.

The proof of the lemma is by induction on the structure of some fixed derivation of $\Lambda^x \vdash R : \hat{\tau} \triangleright e \succ 0$.

One possibility is that x is not free in R . In such a situation by Lemma 29 we can derive $\Lambda \vdash R : \hat{\tau} \triangleright e \succ 0$ and $\text{Split}(\Lambda^x \mid \Lambda)$ holds, which means that $\text{Mk}(\hat{\sigma}_i) = \emptyset$ for all $i \in I$. By (\diamond) we have $d_i = 0$ and $\text{Mk}(\Sigma_i) = \emptyset$ for all $i \in I$. Thus $\text{Split}(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ implies $\text{Split}(\Gamma \mid \Lambda)$, and thus by Lemma 30 applied to $\Lambda \vdash R : \hat{\tau} \triangleright e \succ 0$ we can derive $\Gamma \vdash R : \hat{\tau} \triangleright e \succ 0$. This is the desired type judgment since $R[S/x] = R$ and $e + \sum_{i \in I} d_i = e$.

In the sequel we assume that x is free in R . We analyze the shape of R .

Suppose first that $R = x$. Then the derivation for R consists of a single rule, and thus $e = 0$ and for some $s \in I$ we have $Split(\Lambda^x \mid \varepsilon[x \mapsto \{\hat{\sigma}_s\}])$ and $\hat{\tau} = \hat{\sigma}_s$ (this holds because all $\hat{\sigma}_i$ are of the same order as $\hat{\tau}$). It follows that $Mk(\Lambda(y)) = \emptyset$ for every variable y , and that $Mk(\hat{\sigma}_i) = \emptyset$ for every $i \in I \setminus \{s\}$. By (\diamond) we have $d_i = 0$ and $Mk(\Sigma_i) = \emptyset$ for all $i \in I \setminus \{s\}$. It follows that $Split(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ implies $Split(\Gamma \mid \Sigma_s)$; we can thus derive $\Gamma \vdash S : \hat{\sigma}_s \triangleright d_s \succ 0$ by Lemma 30. This is what we need, since $R[S/x] = S$, and $\hat{\tau} = \hat{\sigma}_s$, and $e + \sum_{i \in I} d_i = d_s$.

Next, suppose that $R = br P_1 P_2$. Then our derivation ends with the (BR) rule, whose premiss is $\Lambda^x \vdash P_k : \hat{\tau} \triangleright e \succ 0$, for some $k \in \{1, 2\}$. The induction assumption applied to this premiss gives us a derivation of $\Gamma \vdash P_k[S/x] : \hat{\tau} \triangleright e + \sum_{i \in I} d_i \succ 0$. By applying back the (BR) rule we derive $\Gamma \vdash R[S/x] : \hat{\tau} \triangleright e + \sum_{i \in I} d_i \succ 0$, as required.

Next, suppose that $R = \lambda y.P$. We have $y \neq x$, and, as always during a substitution, we assume (by performing α -conversion) that y is not free in S . The derivation for R ends with the (λ) rule, whose premiss is $\Lambda_\lambda^x[y \mapsto T] \vdash P : \hat{\tau}' \triangleright e \succ 0$, where $\Lambda_\lambda^x(y) = \emptyset$ and $Split(\Lambda^x \mid \Lambda_\lambda^x)$ holds. Denote $\Lambda_\lambda = \Lambda_\lambda^x[x \mapsto \emptyset]$. We then have $\Lambda_\lambda^x = \Lambda_\lambda[x \mapsto \{\hat{\sigma}_i \mid i \in J\}]$ for some $J \subseteq I$. The condition $Split(\Lambda^x \mid \Lambda_\lambda^x)$ implies that $Split(\Lambda \mid \Lambda_\lambda)$ holds, and that $Mk(\hat{\sigma}_i) = \emptyset$ for all $i \in I \setminus J$; then $d_i = 0$ and $Mk(\Sigma_i) = \emptyset$ for $i \in I \setminus J$ by (\diamond) . In the light of $Split(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ this implies that $Split(\Gamma \mid \Lambda_\lambda, (\Sigma_i)_{i \in J})$ holds. Because $\Gamma(y)$ may be nonempty, we have to define $\Gamma_\lambda = \Gamma[y \mapsto \emptyset]$ and $\Sigma_i^\lambda = \Sigma_i[y \mapsto \emptyset]$ for $i \in J$. By assumption y is not free in S , so Lemma 29 says that we can derive $\Sigma_i^\lambda \vdash S : \hat{\sigma}_i \triangleright d_i \succ 0$ and that $Split(\Sigma_i \mid \Sigma_i^\lambda)$ holds for $i \in J$. Recall that $\Lambda_\lambda(y) = \emptyset$. Due to $Split(\Gamma \mid \Lambda_\lambda, (\Sigma_i)_{i \in J})$ we also have $Split(\Gamma \mid \Gamma_\lambda)$ and $Split(\Gamma_\lambda \mid \Lambda_\lambda, (\Sigma_i^\lambda)_{i \in J})$, thus also $Split(\Gamma_\lambda[y \mapsto T] \mid \Lambda_\lambda[y \mapsto T], (\Sigma_i^\lambda)_{i \in J})$. We are ready to apply the induction assumption to our premiss. We obtain a derivation of $\Gamma_\lambda[y \mapsto T] \vdash P[S/x] : \hat{\tau}' \triangleright e + \sum_{i \in J} d_i \succ 0$. Because $Split(\Gamma \mid \Gamma_\lambda)$ holds and $\Gamma_\lambda(y) = \emptyset$, we can apply the (λ) rule, obtaining $\Gamma \vdash R[S/x] : \hat{\tau} \triangleright e + \sum_{i \in J} d_i \succ 0$, where the full type is indeed $\hat{\tau}$, as in the original derivation. Because $d_i = 0$ for $i \in I \setminus J$, this is what we need.

Another possibility is that $R = a P_1 \dots P_r$. Then the derivation for R ends with the (CON) rule, whose premisses are $\Lambda_j^x \vdash P_j : \hat{\tau}_j \triangleright e_j \succ 0$ for $j \in \{1, \dots, r\}$. For $j \in \{1, \dots, r\}$ denote $\Lambda_j = \Lambda_j^x[x \mapsto \emptyset]$. A side condition of the (CON) rule says that $Split(\Lambda^x \mid (\Lambda_j^x)_{j \in \{1, \dots, r\}})$ holds. On the one hand, this implies $Split(\Lambda \mid (\Lambda_j)_{j \in \{1, \dots, r\}})$. On the other hand, for $j \in \{1, \dots, r\}$ we have $\Lambda_j^x = \Lambda_j[x \mapsto \{\hat{\sigma}_i \mid i \in I_j\}]$ for some $I_j \subseteq I$, and for $i \in I \setminus \bigcup_{j=1}^r I_j$ we have $Mk(\hat{\sigma}_i) = \emptyset$, and thus $d_i = 0$ and $Mk(\Sigma_i) = \emptyset$ by (\diamond) . Knowing that $Split(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ holds, we obtain $Split(\Gamma \mid (\Lambda_j)_{j \in \{1, \dots, r\}}, (\Sigma_i)_{i \in I_j, j \in \{1, \dots, r\}})$. For $j \in \{1, \dots, r\}$ we define Γ_j by taking $\Gamma_j(y) = \Lambda_j(y) \cup \bigcup_{i \in I_j} \Sigma_i(y)$ for all variables y . Then we have $Split(\Gamma \mid (\Gamma_j)_{j \in \{1, \dots, r\}})$ and $Split(\Gamma_j \mid \Lambda_j, (\Sigma_i)_{i \in I_j})$ for $j \in \{1, \dots, r\}$. Using the induction assumption for every premiss, we obtain a derivation of $\Gamma_j \vdash P_j[S/x] : \hat{\tau}_j \triangleright e_j + \sum_{i \in I_j} d_i \succ 0$ for $j \in \{1, \dots, r\}$. Because $Split(\Gamma \mid (\Gamma_j)_{j \in \{1, \dots, r\}})$ holds, and the derived full types are the same as in the derivation for R , and the flag counters are higher by $\sum_{i \in I_j} d_i$, we can apply the (CON) rule and derive $\Gamma \vdash R[S/x] : \hat{\tau} \triangleright e + \sum_{j=1}^r \sum_{i \in I_j} d_i \succ 0$. It remains to observe that $\sum_{j=1}^r \sum_{i \in I_j} d_i = \sum_{i \in I} d_i$. As already said, $d_i = 0$ when $i \in I \setminus \bigcup_{j=1}^r I_j$. Consider some $i \in I_j \cap I_{j'}$ for $j, j' \in \{1, \dots, r\}$ with $j \neq j'$. We have $Mk(\hat{\sigma}_i) \subseteq Mk(\Lambda_j^x)$, so Lemma 24 applied to the type judgment $\Lambda_j^x \vdash P_j : \hat{\tau}_j \triangleright e_j$ says that $Mk(\hat{\sigma}_i) \subseteq Mk(\hat{\tau}_j)$, and similarly $Mk(\hat{\sigma}_i) \subseteq Mk(\hat{\tau}_{j'})$. But, by a side condition of the (CON) rule, $Mk(\hat{\tau}_j)$ and $Mk(\hat{\tau}_{j'})$ are disjoint, so $Mk(\hat{\sigma}_i) = \emptyset$, and hence $d_i = 0$ by (\diamond) . In consequence the two sums are indeed equal.

Finally, suppose that $R = PQ$. The proof is almost the same as in the previous case. The derivation for R ends with the (@) rule, whose premisses are $\Lambda_0^x \vdash P : \hat{\tau}_0 \triangleright e_0 \succ 0$ and $\Lambda_j^x \vdash Q : \hat{\tau}_j \triangleright e_j \succ 0$ for $j \in J$, where we assume that $0 \notin J$. For $j \in \{0\} \cup J$ denote $\Lambda_j = \Lambda_j^x[x \mapsto \emptyset]$. A side condition of the (@) rule says that $Split(\Lambda^x \mid (\Lambda_j^x)_{j \in \{0\} \cup J})$ holds. On the one hand, this implies $Split(\Lambda \mid (\Lambda_j)_{j \in \{0\} \cup J})$. On the other hand, for $j \in \{0\} \cup J$ we have $\Lambda_j^x = \Lambda_j[x \mapsto \{\hat{\sigma}_i \mid i \in I_j\}]$ for some $I_j \subseteq I$, and for $i \in I \setminus \bigcup_{j \in \{0\} \cup J} I_j$ we have $Mk(\hat{\sigma}_i) = \emptyset$, and thus $d_i = 0$ and $Mk(\Sigma_i) = \emptyset$ by (\diamond) . Knowing that $Split(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$ holds, we obtain $Split(\Gamma \mid (\Lambda_j)_{j \in \{0\} \cup J}, (\Sigma_i)_{i \in I_j, j \in \{0\} \cup J})$. For $j \in \{0\} \cup J$ we define Γ_j by taking $\Gamma_j(y) = \Lambda_j(y) \cup \bigcup_{i \in I_j} \Sigma_i(y)$

for all variables y . Then we have $Split(\Gamma \mid (\Gamma_j)_{j \in \{0\} \cup J})$ and $Split(\Gamma_j \mid \Lambda_j, (\Sigma_i)_{i \in I_j})$ for $j \in \{0\} \cup J$. Using the induction assumption for every premiss, we obtain a derivation of $\Gamma_0 \vdash P[S/x] : \hat{\tau}_0 \triangleright e_0 + \sum_{i \in I_0} d_i \succ 0$, and of $\Gamma_j \vdash Q[S/x] : \hat{\tau}_j \triangleright e_j + \sum_{i \in I_j} d_i \succ 0$ for all $j \in J$. Because $Split(\Gamma \mid (\Gamma_j)_{j \in \{0\} \cup J})$ holds, and the derived full types are the same as in the derivation for R , and the flag counters are higher by $\sum_{i \in I_j} d_i$, we can apply the (@) rule and derive $\Gamma \vdash R[S/x] : \hat{\tau} \triangleright e + \sum_{j \in \{0\} \cup J} \sum_{i \in I_j} d_i \succ 0$. It remains to observe that $\sum_{j \in \{0\} \cup J} \sum_{i \in I_j} d_i = \sum_{i \in I} d_i$. As already said, $d_i = 0$ when $i \in I \setminus \bigcup_{j \in \{0\} \cup J} I_j$. Consider some $i \in I_j \cap I_{j'}$ for $j, j' \in \{0\} \cup J$ with $j \neq j'$. We have $Mk(\hat{\sigma}_i) \subseteq Mk(\Lambda_j^x)$, so Lemma 24 applied to the type judgment $\Lambda_j^x \vdash T : \hat{\tau}_j \triangleright e_j$ (where $T = P$ or $T = Q$, depending on j) says that $Mk(\hat{\sigma}_i) \subseteq Mk(\hat{\tau}_j)$, and similarly $Mk(\hat{\sigma}_i) \subseteq Mk(\hat{\tau}_{j'})$. But, by a side condition of the (@) rule, $Mk(\hat{\tau}_j)$ and $Mk(\hat{\tau}_{j'})$ are disjoint, so $Mk(\hat{\sigma}_i) = \emptyset$, and hence $d_i = 0$ by (\diamond). In consequence the two sums are indeed equal. \square

Lemma 32. *Suppose that we can derive $\Gamma \vdash RS : \hat{\tau} \triangleright c \succ n$ so that all premisses of the final (@) rule have 0 in the application counter. If $ord(R) = ord(\hat{\tau})$, and $\Gamma(y) \neq \emptyset$ only for variables y of order at most $ord(\hat{\tau}) - 1$, then $R = \lambda x.R'$, and we can derive $\Gamma \vdash R'[S/x] : \hat{\tau} \triangleright c \succ 0$.*

Proof. Let $\Gamma_\lambda \vdash R : \hat{\tau}_\lambda \triangleright e \succ 0$ and $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i \succ 0$ for $i \in I$ be the premisses of the final (@) rule. Let us write $\hat{\tau} = (m, F, M, \tau)$, and $\hat{\tau}_\lambda = (m, F', M_\lambda, T \rightarrow \tau)$, and $\hat{\sigma}_i = (m, F_i, M_i, \sigma_i)$ for $i \in I$. We have $T = \{(ord(R), F_i \upharpoonright_{<ord(R)}, M_i \upharpoonright_{<ord(R)}, \sigma_i) \mid i \in I\} = \{\hat{\sigma}_i \mid i \in I\}$, because $ord(R) = m$ and sets F_i and M_i contain only numbers smaller than m .

We start by determining the shape of R , by looking at the premiss concerning it, i.e., $\Gamma_\lambda \vdash R : \hat{\tau}_\lambda \triangleright e \succ 0$. If R was a variable, then the derivation of this premiss would consist of the (VAR) rule requiring that $\Gamma_\lambda(R) \neq \emptyset$, hence also $\Gamma(R) \neq \emptyset$; this is impossible by assumption since R is of order m . If R is an application, $R = UV$, then the derivation of the premiss concerning R starts with the (@) rule, requiring that $ord(U) \leq m$. However $ord(U) \geq ord(R) = m$, so actually $ord(U) = m$, and hence the application counter in this premiss should be at least 1, violating our assumption. It follows that R cannot be an application. Moreover, R takes an argument, so it cannot start with a symbol. Thus R starts with a λ -abstraction, $R = \lambda x.R'$.

The type judgment concerning R is derived using the (λ) rule out of a premiss $\Lambda[x \mapsto T] \vdash R' : \hat{\tau}' \triangleright e \succ 0$, where $\hat{\tau}' = (m, F', M', \tau)$ and $\Lambda(x) = \emptyset$. The two rules imply that $M_\lambda = M' \setminus Mk(T)$ and $M = M_\lambda \cup Mk(T)$. Lemma 24 applied to the type judgment $\Lambda[x \mapsto T] \vdash R' : \hat{\tau}' \triangleright e$ implies that $Mk(T) \subseteq M'$, and thus we obtain $M = M'$. Next, we notice that the sets $F_i \upharpoonright_{\geq ord(R)}$ are empty, and that $F' \cap M = F' \cap M' = \emptyset$ (because $\hat{\tau}' = (m, F', M', \tau)$ is a full type), and that $Comp_m(M; (F', e), ((F_i \upharpoonright_{\geq ord(R)}, d_i))_{i \in I}) = (F, c)$ by a side condition of the (@) rule. In such a situation Lemma 26 implies that $F = F'$ (thus actually $\hat{\tau}' = \hat{\tau}$) and $c = e + \sum_{i \in I} d_i$. Finally, due to side conditions of the (@) and (λ) rules we have $Split(\Gamma \mid \Gamma_\lambda, (\Sigma_i)_{i \in I})$ and $Split(\Gamma_\lambda \mid \Lambda)$, hence also $Split(\Gamma \mid \Lambda, (\Sigma_i)_{i \in I})$. We also have $ord(S) < ord(R) = m$. Thus we can apply Lemma 31 to type judgments $\Sigma_i \vdash S : \hat{\sigma}_i \triangleright d_i \succ 0$ for $i \in I$ and $\Lambda[x \mapsto T] \vdash R' : \hat{\tau} \triangleright e \succ 0$. We obtain a derivation of $\Gamma \vdash R'[S/x] : \hat{\tau} \triangleright c \succ 0$, as required. \square

We now give a generalization of Lemma 7 suitable for induction. Notice that a subterm of a λ -term may be involved in multiple subtrees of a derivation tree for the whole λ -term. Because of that, we have to handle multiple derivations for the same λ -term at once.

Lemma 33. *Suppose that for a finite set I , a number m , and a λ -term P we can derive $\Gamma_i \vdash P : \hat{\tau}_i \triangleright c_i \succ n_i$ for $i \in I$, where $ord(P) \leq m$, and $\sum_{i \in I} n_i > 0$, and for all $i \in I$ we have $ord(\hat{\tau}_i) = m$ and $\Gamma_i(x) \neq \emptyset$ only for variables x of order at most $m - 1$. Then there is a λ -term Q such that $P \rightarrow_\beta Q$, and we can derive $\Gamma_i \vdash Q : \hat{\tau}_i \triangleright c_i \succ n'_i$ for all $i \in I$, for some numbers n'_i such that $\sum_{i \in I} n'_i < \sum_{i \in I} n_i$.*

Proof. The proof is by induction on the smallest total size of derivations needed to derive $\Gamma_i \vdash P : \hat{t}_i \triangleright c_i \succ n_i$ for all $i \in I$. Because $\sum_{i \in I} n_i > 0$, we have $|I| \geq 1$.

Suppose first that for every $i \in I$ all premisses of the last rule used to derive $\Gamma_i \vdash P : \hat{t}_i \triangleright c_i \succ n_i$ have 0 in the application counter. This is the base case, in which we perform a β -reduction. Because $\sum_{i \in I} n_i > 0$, necessarily $P = RS$ with $\text{ord}(R) = m$, because only in such a situation the application counter in a derived type judgment can be higher than the sum of application counters in premisses. Then, for every $i \in I$ separately we apply Lemma 32 to our type judgment, and we obtain that $R = \lambda x.R'$, and that we can derive $\Gamma_i \vdash R'[S/x] : \hat{t}_i \triangleright c_i \succ 0$. Since $P = (\lambda x.R')S \rightarrow_\beta R'[S/x]$, this gives the thesis.

Let us now consider the opposite case, when we have a premiss with positive application counter. We have multiple cases depending on the shape of P , but all of them are similar, and boil down to a use of the induction assumption. Suppose, for example, that $P = RS$. Then for every $i \in I$ the type judgment $\Gamma_i \vdash P : \hat{t}_i \triangleright c_i \succ n_i$ is derived by the (@) rule out of premisses $\Gamma'_i \vdash R : \hat{t}'_i \triangleright c'_i \succ n'_i$ and $\Gamma_{i,j} \vdash S : \hat{t}_{i,j} \triangleright c_{i,j} \succ n_{i,j}$ for $j \in J_i$, for some finite set J_i . We have $\text{ord}(S) < \text{ord}(R) \leq m$ by a side condition of the (@) rule. For every variable x of order at least m we have $\Gamma'_i(x) \subseteq \Gamma_i(x) = \emptyset$ for all $i \in I$, and $\Gamma_{i,j}(x) \subseteq \Gamma_i(x) = \emptyset$ for all $i \in I, j \in J_i$. When $n_{i,j} > 0$ for some $i \in I, j \in J_i$, we apply the induction assumption to S and to the collection of all premisses concerning it in all our derivations. We obtain a λ -term S' such that $S \rightarrow_\beta S'$, and derivations of $\Gamma_{i,j} \vdash S' : \hat{t}_{i,j} \triangleright c_{i,j} \succ n'_{i,j}$ for all $i \in I, j \in J_i$, where $\sum_{i \in I} \sum_{j \in J_i} n'_{i,j} < \sum_{i \in I} \sum_{j \in J_i} n_{i,j}$. By applying the (@) rule to these type judgments and to the premisses concerning R , we obtain derivations of $\Gamma_i \vdash Q : \hat{t}_i \triangleright c_i \succ n'_i$ for all $i \in I$, where $Q = RS'$ and $\sum_{i \in I} n'_i = \sum_{i \in I} (n_i + \sum_{j \in J_i} (n'_{i,j} - n_{i,j})) < \sum_{i \in I} n_i$. When $n_{i,j} = 0$ for all $i \in I, j \in J_i$, we necessarily have $n'_i > 0$ for some i (as we are in the case in which some premiss has a positive application counter), and we apply the induction assumption to the premisses concerning R . We proceed similarly when $P = aP_1 \dots P_r$ for $a \neq \text{br}$, when $P = \text{br}P_1P_2$, and when $P = \lambda x.R$ (we cannot have $P = x$, as then the application counter in all the type judgments would be 0). In the case of $P = \lambda x.R$ we use the assumption $\text{ord}(P) \leq m$ to deduce that the full types assigned to x in type environments of premisses are not assigned to a variable of order higher than $m - 1$: we have $\text{ord}(x) < \text{ord}(P) \leq m$. \square

Finally, we observe that Lemma 7 is a special case of Lemma 33, where $|I| = 1$ and the type judgment is of the form $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$.

F Proof of Lemma 8

Lemma 34. *Suppose that $\text{Comp}_m(M; (F_i, c_i)_{i \in I}) = (F, c)$, where $m \geq 1$. If $c'_i \geq c_i + |F_i \cap \{m-1\}|$ for $i \in I$, then $\text{Comp}_{m-1}(M \upharpoonright_{< m-1}; (F_i \upharpoonright_{< m-1}, c'_i)_{i \in I}) = (F \upharpoonright_{< m-1}, c')$ for some $c' \geq c + |F \cap \{m-1\}|$.*

Proof. The definition of the *Comp* predicate specifies variables f_n and f'_n . Let $f_{n,m}$ and $f'_{n,m}$ be values of these variables in the above instantiation of the Comp_m predicate, while $f_{n,m-1}$ and $f'_{n,m-1}$ —in the above instantiation of the Comp_{m-1} predicate. Notice that $f_{n,m-1} = f_{n,m}$ for $n < m-1$, and $f'_{n,m-1} = f'_{n,m}$ for $n \leq m-1$. In consequence the requirements given by Comp_{m-1} on the set F are satisfied, since they are the same as the requirements given by Comp_m .

Next, let us observe that $f_{m-1,m} \geq f'_{m,m} + |F \cap \{m-1\}|$. Indeed, if $m-1 \in M$, we have $f'_{m,m} = f_{m-1,m}$ and $m-1 \notin F$. Conversely, if $m-1 \notin M$, we have $f'_{m,m} = 0$, and if $f_{m-1,m} = 0$ then also $m-1 \notin F$.

Finally, because $f'_{m-1,m-1} = f'_{m-1,m}$, we have

$$\begin{aligned} c' &= f'_{m-1,m-1} + \sum_{i \in I} c'_i = f'_{m-1,m} + \sum_{i \in I} c'_i \geq f'_{m-1,m} + \sum_{i \in I} |F_i \cap \{m-1\}| + \sum_{i \in I} c_i \\ &= f_{m-1,m} + \sum_{i \in I} c_i \geq f'_{m,m} + \sum_{i \in I} c_i + |F \cap \{m-1\}| = c + |F \cap \{m-1\}|. \end{aligned} \quad \square$$

We now generalize Lemma 8 to arbitrary type judgments.

Lemma 35. *Suppose that we can derive $\Gamma \vdash P : (m, F, M, \tau) \triangleright c \succ 0$, where $\text{ord}(P) \leq m-1$, and for every variable x and every $\hat{\eta} \in \Gamma(x)$ we have $\text{ord}(\hat{\eta}) \leq m-1$. Then we can also derive $\Gamma \vdash P : (m-1, F \upharpoonright_{< m-1}, M \upharpoonright_{< m-1}, \tau) \triangleright c'$ with $c' \geq c + |F \cap \{m-1\}|$.*

Proof. Denote $\hat{\tau} = (m, F, M, \tau)$ and $\hat{\sigma} = (m-1, F \upharpoonright_{< m-1}, M \upharpoonright_{< m-1}, \tau)$. The proof is by induction on the structure of some fixed derivation of $\Gamma \vdash P : \hat{\tau} \triangleright c \succ 0$. We have $m \geq 1$, since $\text{ord}(P) \leq m-1$. We distinguish several cases depending on the shape of R .

Suppose first that P is a variable, $P = x$. Then the (VAR) rule used in the derivation implies that $\text{Split}(\Gamma \mid \varepsilon[x \mapsto (k, F, M \upharpoonright_{< k}, \tau)])$ holds, and $c = 0$. By assumption of the lemma we have $k \leq m-1$, so $(M \upharpoonright_{< m-1}) \upharpoonright_{< k} = M \upharpoonright_{< k}$ and $F \upharpoonright_{< m-1} = F$ (because $F \subseteq \{0, \dots, k-1\}$). In consequence, we can use the (VAR) rule to derive $\Gamma \vdash P : \hat{\sigma} \triangleright 0$.

Next, suppose that $P = \text{br } P_1 P_2$. Then the final (BR) rule has a premiss $\Gamma \vdash P_k : \hat{\tau} \triangleright c \succ 0$ for some $k \in \{1, 2\}$. Surely $\text{ord}(P_k) = 0 \leq m-1$. The induction assumption applied to this premiss gives us a derivation of $\Gamma \vdash P_k : \hat{\sigma} \triangleright c'$ with $c' \geq c + |F \cap \{m-1\}|$. We apply back the (BR) rule, obtaining $\Gamma \vdash P : \hat{\sigma} \triangleright c'$.

Next, suppose that $P = \lambda x. Q$. Then the final (λ) rule has a premiss $\Gamma'[x \mapsto T] \vdash Q : (m, F, M', \tau') \triangleright c \succ 0$, where $\tau = T \rightarrow \tau'$, and $M = M' \setminus \text{Mk}(T)$, and $\text{Split}(\Gamma \mid \Gamma')$ holds. Clearly $\text{ord}(Q) \leq \text{ord}(P) \leq m-1$. We also have $T \subseteq \mathcal{F}_{\text{ord}(P)}^\alpha$, where α is the sort of x . In consequence, for every variable y and every $\hat{\eta} \in (\Gamma'[x \mapsto T])(y)$ we have $\text{ord}(\hat{\eta}) \leq m-1$. Using the induction assumption for our premiss we obtain a derivation of $\Gamma'[x \mapsto T] \vdash Q : (m-1, F \upharpoonright_{< m-1}, M' \upharpoonright_{< m-1}, \tau') \triangleright c'$ with $c' \geq c + |F \cap \{m-1\}|$. Because $M' \upharpoonright_{< m-1} \setminus \text{Mk}(T) = (M' \setminus \text{Mk}(T)) \upharpoonright_{< m-1} = M \upharpoonright_{< m-1}$, by applying back the (λ) rule we derive $\Gamma \vdash P : \hat{\sigma} \triangleright c'$.

Next, suppose that $P = a P_1 \dots P_r$ with $a \neq \text{br}$. Then $\tau = o$, and the final (CON) rule has premisses $\Gamma_i \vdash P_i : (m, F_i, M_i, o) \triangleright c_i \succ 0$ for $i \in \{1, \dots, r\}$. For every $i \in \{1, \dots, r\}$ we have $\text{ord}(P_i) = 0 \leq m-1$, and $\Gamma_i(x) \subseteq \Gamma(x)$ for every variable x , so we can use the induction assumption for the i -th premiss and obtain a derivation of $\Gamma_i \vdash P_i : (m-1, F_i \upharpoonright_{< m-1}, M_i \upharpoonright_{< m-1}, o) \triangleright c'_i$ with $c'_i \geq c_i + |F_i \cap \{m-1\}|$. If $r > 0$, we have a side condition $M = \uplus_{i=1}^r M_i$, which implies $M \upharpoonright_{< m-1} = \uplus_{i=1}^r M_i \upharpoonright_{< m-1}$. Another side condition says that $\text{Comp}_m(M; (\{0\}, 0), (F_i, c_i)_{i \in \{1, \dots, r\}}) = (F, c)$, and we need to see that $\text{Comp}_{m-1}(M \upharpoonright_{< m-1}; (F_0, c_0), (F_i \upharpoonright_{< m-1}, c'_i)_{i \in \{1, \dots, r\}}) = (F \upharpoonright_{< m-1}, c')$ for some $c' \geq c + |F \cap \{m-1\}|$, where $(F_0, c_0) = (\{0\}, 0)$ if $m-1 > 0$, and $(F_0, c_0) = (\emptyset, 1)$ if $m-1 = 0$. This follows from Lemma 34, where we notice that $\{0\} \upharpoonright_{< m-1} = F_0$, and $c_0 \geq 0 + |\{0\} \cap \{m-1\}|$. Thus we can apply back the (CON) rule deriving $\Gamma \vdash P : \hat{\sigma} \triangleright c'$.

Finally, suppose that $P = QR$. Then the final (@) rule has premisses $\Gamma' \vdash Q : (m, F', M', T \rightarrow \tau) \triangleright e \succ 0$ and $\Gamma_i \vdash R : (m, F_i, M_i, \tau_i) \triangleright d_i \succ 0$ for $i \in I$, where $T = \{(\text{ord}(Q), F_i \upharpoonright_{< \text{ord}(Q)}, M_i \upharpoonright_{< \text{ord}(Q)}, \tau_i) \mid i \in I\}$. Because the application counter is 0 in the conclusion, we have $\text{ord}(Q) \neq m$, so actually $\text{ord}(Q) \leq m-1$ by a side condition of the (@) rule. Simultaneously $\text{ord}(R) < \text{ord}(Q) \leq m-1$, and the type environments Γ' and $(\Gamma_i)_{i \in I}$ store only full types stored already in Γ . The induction assumption applied to all premisses gives us derivations of $\Gamma' \vdash Q : (m-1, F' \upharpoonright_{< m-1}, M' \upharpoonright_{< m-1}, T \rightarrow \tau) \triangleright e'$ with $e' \geq e + |F' \cap \{m-1\}|$, and of $\Gamma_i \vdash R : (m-1, F_i \upharpoonright_{< m-1}, M_i \upharpoonright_{< m-1}, \tau_i) \triangleright d'_i$ with $d'_i \geq d_i + |F_i \cap \{m-1\}|$ for $i \in I$. The

side condition $M = M' \uplus \uplus_{i \in I} M_i$ implies $M \upharpoonright_{< m-1} = M' \upharpoonright_{< m-1} \uplus \uplus_{i \in I} M_i \upharpoonright_{< m-1}$. Another side condition says that $\text{Comp}_m(M; (F', e), (F_i, d_i)_{i \in I}) = (F, c)$, which by Lemma 34 implies that $\text{Comp}_{m-1}(M \upharpoonright_{< m-1}; (F' \upharpoonright_{< m-1}, e'), (F_i \upharpoonright_{< m-1}, d'_i)_{i \in I}) = (F \upharpoonright_{< m-1}, c')$ for some $c' \geq c + |F \cap \{m-1\}|$. We also have that $T = \{(ord(Q), (F_i \upharpoonright_{< m-1}) \upharpoonright_{ord(Q)}, (M_i \upharpoonright_{< m-1}) \upharpoonright_{ord(Q)}, \tau_i) \mid i \in I\}$, because $ord(Q) \leq m-1$. Having all this, we can apply back the (@) rule, and derive $\Gamma \vdash P : \hat{\sigma} \triangleright c'$. \square

Lemma 8 says that if we can derive $\varepsilon \vdash P : (m, \emptyset, \{0, \dots, m-1\}, o) \triangleright c \succ 0$ with $m > 0$, then we can also derive $\varepsilon \vdash P : (m-1, \emptyset, \{0, \dots, m-2\}, o) \triangleright c'$ for some $c' \geq c$. Here $ord(P) = 0$, so this is just a special case of Lemma 35.

G Remaining proofs for Section 5

In the final part of Section 5 we have implicitly used the following lemma, which we now prove.

Lemma 36. *If we can derive $\varepsilon \vdash P : \hat{\rho}_0 \triangleright c$, then there exists a tree $t \in \mathcal{L}(P)$ such that $|t| = c$.*

Proof. Recall that $\hat{\rho}_0 = (0, \emptyset, \emptyset, o)$. The proof is by induction on the structure of some fixed derivation of $\varepsilon \vdash P : \hat{\rho}_0 \triangleright c$. Let us analyze the shape of P . Because the type environment is empty, P cannot be a variable. The sort of $\hat{\rho}_0$, and hence of P , is o , and thus P cannot start with a λ -abstraction. Moreover, P cannot be an application QR , because the (@) rule requires that $ord(Q) \leq ord(\hat{\rho}_0) = 0$. Thus P starts with a symbol. We have two cases.

Suppose first that $P = aP_1 \dots P_r$ with $a \neq \text{br}$. We notice that $\hat{\rho}_0$ is the only full type in \mathcal{F}_0^o , and that $\text{Split}(\varepsilon \mid \Gamma_1, \dots, \Gamma_r)$ implies $\Gamma_1 = \dots = \Gamma_r = \varepsilon$. Thus the premisses of the final (CON) rule are $\varepsilon \vdash P_i : \hat{\rho}_0 \triangleright c_i$, for $i \in \{1, \dots, r\}$. Because $\text{Comp}_0(\emptyset; (\emptyset, 1), (\emptyset, c_1), \dots, (\emptyset, c_r)) = (\emptyset, c)$, we have $c = 1 + c_1 + \dots + c_r$. The induction assumption gives us, for $i \in \{1, \dots, r\}$, trees t_i such that $|t_i| = c_i$ and $t_i \in \mathcal{L}(P_i)$, which means that $BT(P_i) \rightarrow_{\text{br}}^* t_i$. As t we take the tree having a in its root, and t_1, \dots, t_r as subtrees starting in the root's children. Because $BT(P)$ has a in its root, and $BT(P_1), \dots, BT(P_r)$ as subtrees starting in the root's children, it should be clear that $BT(P) \rightarrow_{\text{br}}^* t$. Moreover, $|t| = 1 + |t_1| + \dots + |t_r| = 1 + c_1 + \dots + c_r = c$.

Another possibility is that $P = \text{br } P_1 P_2$. Then the final (BR) rule has one premiss $\varepsilon \vdash P_i : \hat{\rho}_0 \triangleright c$ for some $i \in \{1, 2\}$. The induction assumption gives us a tree t such that $|t| = c$ and $BT(P_i) \rightarrow_{\text{br}}^* t$. Recalling that $BT(P)$ has br in its root, and $BT(P_i)$ as its subtree starting in the i -th child of the root, we see that $BT(P) \rightarrow_{\text{br}} BT(P_i)$, and thus $t \in \mathcal{L}(P)$. \square

H Proofs for Section 6

Let us recall two definitions from page 13. We say that two type judgments are equivalent if they differ only in the value of the flag counter. Given a λY -term P and a number m , we have also defined a set \mathcal{D} of all derivations of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ in which on each branch there are at most three type judgments from every equivalence class, and among premisses of each (@) rule there is at most one type judgment from every equivalence class.

We complete the proof of effectiveness contained in Section 6 by a formal proof of the following lemma.

Lemma 37. *Suppose that for some λY -term P we can derive $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ for arbitrarily large numbers $c \in \mathbb{N}$. Then in the set \mathcal{D} there is a derivation in which on some branch there are two equivalent type judgments with different values of the flag counter.*

Proof. In this lemma it is convenient to see derivations as trees: type judgments of a derivation constitute nodes of a tree; premisses of a type judgment are located in its children. We consider P and m to be fixed. A derivation is called *narrow* if among premisses of each its (@) rule there is at most one type judgment from every equivalence class. We have already justified on page 14 that if a type judgment $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ can be derived, then it has a narrow derivation. Moreover, we have justified that there are only finitely many equivalence classes of type judgments that can be used in any derivation of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$, for any c ; let E be their number. This gives a bound on the number of premisses of (@) rules in narrow derivations. Rules (VAR), (BR), and (λ) always have at most one premiss. The number of premisses of a (CON) rule is specified by the rank of the symbol involved; when this rule is used in a derivation of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$, this symbol has to appear in P , which gives a bound on its rank. All this gives us a bound D on the degree of nodes appearing in the considered narrow derivations of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ for arbitrarily large c .

Looking at the definition of the $Comp_m$ predicate it is easy to see that if $Comp_m(M; (F_i, c_i)_{i \in I}) = (F, c)$, then $c \leq |I| \cdot m + \sum_{i \in I} c_i$. In consequence, the flag counter in a conclusion of a (@) rule or a (CON) rule, having at most D premisses, can be higher than the sum of flag counters in the premisses at most by $(D+1) \cdot m + 1$ (these “+1” appear here, because in the (CON) rule, beside of pairs (F_i, c_i) coming from premisses, we pass to the $Comp_m$ predicate an additional pair (F', c') with $c' \leq 1$). The conclusion of every (BR) and (λ) rule has the same flag counter as the only premiss, and the conclusion of the (VAR) rule always has 0 in its flag counter. This means that there is constant C such that in any node of any narrow derivation of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ the flag counter can be higher than the sum of flag counters in its premisses at most by C (and simultaneously it cannot be smaller than this sum).

We define a *level* of a node in a derivation, by induction on the depth of the node. Leaves and all nodes with flag counter 0 have level 0. If an internal node has its flag counter positive and equal to the flag counter in some child of this node, then the level of this node is equal to the level of this child (notice that there is at most one such child, and the flag counter in all other children is 0). The level of every other internal node is defined as one plus the maximum of levels of its children.

Next, we prove that for every level i there is a bound C_i on the value of the flag counter among nodes of this level. Indeed, the value of the flag counter in leaves, and thus in all nodes of level 0, is bounded by $C_0 = C$. Take now a node of level $i > 0$ having only children of levels smaller than i . In each of these (at most D) children the flag counter is at most C_{i-1} , so so the flag counter in our node is at most $C_i = C + D \cdot C_{i-1}$. If a node of level i has a child of level i , then their flag counter is equal, thus is also at most C_i (trivial induction on the depth of the node).

Let us now take a narrow derivation of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$ for some $c > C_{E-1}$. It necessarily contains a node of level greater than $E - 1$. Moreover, every node of some level $i > 0$ has a child of level at least $\geq i - 1$. Thus in the considered derivation there exists a branch having nodes of at least $E + 1$ different levels. Among them we can find two nodes with equivalent type judgments and being on different levels (as the type judgments come from at most E equivalence classes). If the two nodes had the same flag counters, then also all nodes on the path between them would have the same flag counter, and thus the two nodes would have the same level, which is not the case. Thus we have found two nodes having equivalent type judgments, different values of the flag counter, and such that one of them is a descendant of the other.

Let x and y be nodes of a derivation of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c$, where y is a descendant of x , and they contain equivalent type judgments. In such a situation, we can *cut out* the fragment between nodes x and y , in the following sense: we decrease by $c_x - c_y$ the flag counter in every ancestor of x , we remove x and all its descendants not being in the subtree starting in y , and we attach y in the place of x . This results in some (correct) narrow derivation of $\varepsilon \vdash P : \hat{\rho}_m \triangleright c'$ for some c' .

Take now the smallest (in the sense of the number of nodes) narrow derivation of a type judgment

$\varepsilon \vdash P : \hat{\rho}_m \triangleright c$, for any c , in which there are two nodes u, v having equivalent type judgments, different values of the flag counter, and such that v is a descendant of u . If this derivation is in \mathcal{D} , we are done. If not, we can find four nodes x_1, x_2, x_3, x_4 with equivalent type judgments, such that x_{i+1} is a descendant of x_i for $i \in \{1, 2, 3\}$. Let c_i be the value of the flag counter in x_i , for $i \in \{1, 2, 3, 4\}$. We have two cases. Suppose first that $c_i > c_{i+1}$ for some $i \in \{1, 2, 3\}$. We then take some $j \in \{1, 2, 3\} \setminus \{i\}$, and we cut off the fragment between x_j and x_{j+1} ; we obtain a smaller narrow derivation of a type judgment $\varepsilon \vdash P : \hat{\rho}_m \triangleright c'$, for some c' , in which there are two nodes, namely x_i and x_{i+1} , having equivalent type judgments, different values of the flag counter, and such that one of them is a descendant of the other. This contradicts minimality of our derivation. Next, suppose that $c_1 = c_2 = c_3 = c_4$. Then we recall that we already have two nodes u, v having equivalent type judgments, different values of the flag counter, and such that v is a descendant of u . For $i \in \{1, 2, 3\}$ consider the set V_i containing x_i and all its descendants not being in the subtree starting in x_{i+1} . These sets are disjoint, so for some $j \in \{1, 2, 3\}$ we have $u \notin V_j$ and $v \notin V_j$. We cut off the fragment between x_j and x_{j+1} . This removes exactly the nodes from V_j , so the nodes u and v are still present in the derivation. Moreover, equality $c_j = c_{j+1}$ implies that we have not changed flag counters in any node, in particular in u and v , so u and v in the new derivation again have different values of the flag counter. Thus also in this case we have obtained a contradiction with minimality of our derivation. \square