Math. Struct. in Comp. Science (2013), *vol.* 23, *pp.* 568–607. © Cambridge University Press 2012
doi:10.1017/S0960129512000412 First published online 9 July 2012

Ticket Entailment is decidable

VINCENT PADOVANI

Equipe Preuves, Programmes et Systèmes,
Université Paris VII – Denis Diderot,
Case 7014,
75205 PARIS Cedex 13, France
Email: padovani@pps.jussieu.fr

Received 19 June 2010; revised 6 March 2012

We prove the decidability of the logic $T_{\to}$ of Ticket Entailment. This issue was first raised by Anderson and Belnap within the framework of relevance logic, and is equivalent to the question of the decidability of type inhabitation in simply typed combinatory logic with the partial basis $\mathsf{BB'IW}$. We solve the equivalent problem of type inhabitation for the restriction of simply typed lambda calculus to hereditarily right-maximal terms.

1. Introduction

The partial bases built using the atomic combinators B, $\mathsf{B'}$, C, I, K, W of combinatory logic are well known for being closely connected with propositional logic. The types of their combinators form the axioms of implicational logic systems that have been studied now for well over 70 years (Trigg *et al.* 1994). The partial basis $\mathsf{BB'IW}$ corresponds, through the types of its combinators, to the system $T_{\to}$ of *Ticket Entailment*, which was introduced and motivated in Anderson and Belnap (1975) and Anderson *et al.* (1990). The system $T_{\to}$ consists of *modus ponens* and four axiom schemes that range over the following types for each atomic combinator:

— $\mathsf{B} : (\chi \to \psi) \to ((\phi \to \chi) \to (\phi \to \psi))$
— $\mathsf{B'} : (\phi \to \chi) \to ((\chi \to \psi) \to (\phi \to \psi))$
— $\mathsf{I} : \phi \to \phi$
— $\mathsf{W} : (\phi \to (\phi \to \chi)) \to (\phi \to \chi)$

The type inhabitation problem for $\mathsf{BB'IW}$ is the problem of deciding for a given type whether there exists within this basis a combinator of this type. This problem is equivalent to the problem of deciding whether a given formula can be derived in $T_{\to}$.

Surprisingly, the question of the decidability of $T_{\to}$ has remained unsolved since it was raised in Anderson and Belnap (1975), though the problem has been thoroughly explored within the framework of relevance logic with proofs of decidability and undecidability for several related systems. For instance, the system $R_{\to}$ of *Relevant Implication* (which corresponds to the basis BCIW) and the system $E_{\to}$ of *Entailment* (Anderson and Belnap 1975) are both decidable (Kripke 1959), whereas the extensions R, E, T of $R_{\to}$, $E_{\to}$, $T_{\to}$ to a larger set of connectives ($\to$, $\wedge$, $\vee$) are undecidable (Urquhart 1984).

In 2004, a partial decidability result for the type inhabitation problem was proposed in Broda *et al.* (2004) for a restricted class of formulas – the class of 1-*unary formulas* in

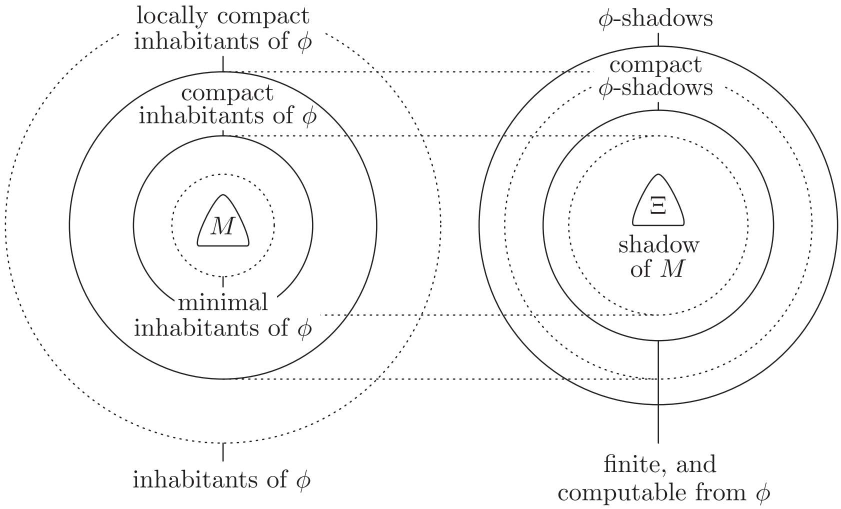

Fig. 1. The principle of our proof of the decidability of type inhabitation for HRM-terms.

which every maximal negative subformula has arity at most 1. Broda, Dams, Finger and Silva e Silva's approach is based on a translation of the problem into a type inhabitation problem for the *hereditary right-maximal* (HRM) terms of lambda calculus (Trigg *et al.* 1994; Bunder 1996; Broda *et al.* 2004). The closed HRM-terms form the closure under β-reduction of all translations of BB′IW-terms, so the type inhabitation problem within the basis BB′IW is equivalent to the type inhabitation problem for HRM-terms.

In this paper we use the same approach as used by Broda, Dams, Finger and Silva e Silva, and prove that the type inhabitation problem for HRM-terms is decidable, and thus conclude that the logic $T_{\rightarrow}$ is decidable[†].

1.1. *Organisation of the paper*

In Section 2, we recall the definition of hereditarily right-maximal terms and the equivalence between the decidability of type inhabitation for BB′IW and the decidability of type inhabitation for HRM-terms. The principle of our proof is shown in Figure 1.

In Sections 3 and 4, we provide for each formula ϕ a partial characterisation of the inhabitants of ϕ in normal form and of minimal size. We show that all of these inhabitants belong to two larger sets of terms, *viz*. the set of *compact* and *locally compact* inhabitants of ϕ.

In Section 5, we show how to associate with each locally compact inhabitant M of a formula ϕ, a labelled tree with the same tree structure as M – we call this tree the *shadow* of M. We then define for shadows the analogue of compactness for terms, and prove that the shadow of a compact term is itself compact.

† In the course of the publication of this article we heard of work in progress by Katalin Bimbò and Michael Dunn towards a solution that seems to be based on a different approach.

Finally, in Section 6, we prove that for each formula ϕ the set of all compact shadows of inhabitants of ϕ is a finite set (hence the set of compact inhabitants of ϕ is also a finite set), and then prove that this set is effectively computable from ϕ. The proof appeals to the Higman and Kruskal Theorems – more precisely, to Melliès' Axiomatic Kruskal Theorem.

The decidability of the type inhabitation problem for HRM-terms and the decidability of $T_{\rightarrow}$ follow from a final key result: given an arbitrary formula ϕ, this formula is inhabited if and only if there exists a compact shadow with the same tree structure as an inhabitant of ϕ, and our key lemma proves that the existence of such a shadow is decidable.

1.2. *Preliminaries*

Section 2 assumes some familiarity with pure and simply typed lambda calculus and the usual notions of α-conversion, β-reduction and β-normal form (Barendregt 1984; Krivine 1993). These last three notions are not essential to our discussion, as we later focus exclusively on a particular set of simply typed terms in β-normal form. We shall now briefly recall the definitions and results used in Section 2.

The set of *terms of pure lambda calculus* (*λ-terms*) is defined inductively as follows:

— Every variable x is a λ-term.
— If M is a λ-term and x is a variable, then $(\lambda x M)$ is a λ-term.
— If M, N are λ-terms, then (MN) is a λ-term.

The terms given by the second and third rules are called *abstractions* and *applications*, respectively. The parentheses surrounding applications and abstractions are often omitted if no ambiguity arises. We write

$$\lambda x_1 \ldots x_n.MN_1 \ldots N_p$$

to abbreviate

$$(\lambda x_1(\ldots(\lambda x_n(((MN_1)\ldots)N_p))\ldots)).$$

For instance, $\lambda xy.x(xy)z$ stands for $(\lambda x(\lambda y((x(xy))z)))$.

The *bound variables* of M are all x such that λx occurs in M. A variable x is *free* in M if and only if any of the following holds:

— $M = x$.
— $M = \lambda y.N$, $y \neq x$ and x is free in N.
— $M = NP$ and x is free in N or free in P.

A *closed* term is a term with no free variables. The *raw substitution of N for x in M*, written $M\langle x \leftarrow N\rangle$, is the term obtained by substituting N for every free occurrence of x in M (that is, every occurrence of x that is not in the scope of a λx). We require this substitution to avoid variable capture (for all y free in N, no free occurrence of x in M is in the scope of a λy):

— If $y = x$, then $y\langle x \leftarrow N\rangle$ is equal to N, otherwise it is equal to y.
— $(\lambda x.M)\langle x \leftarrow N\rangle = \lambda x.M$.

— If $y \neq x$ and y is free in N, then $(\lambda y.M)\langle x \leftarrow N\rangle$ is undefined.
— If $y \neq x$, y is not free in N and $M\langle x \leftarrow N\rangle = M'$, then $(\lambda y.M)\langle x \leftarrow N\rangle = \lambda y.M'$.
— If $M_1\langle x \leftarrow N\rangle = M_1'$ and $M_2\langle x \leftarrow N\rangle = M_2'$, then $(M_1M_2)\langle x \leftarrow N\rangle = (M_1'M_2')$.

α-conversion is defined as the least binary relation $\equiv_\alpha$ such that:

— $x \equiv_\alpha x$.
— If $M \equiv_\alpha M'$, y is not free in M' and $M'\langle x \leftarrow y\rangle = M''$, then $(\lambda x.M) \equiv_\alpha (\lambda y.M'')$.
— If $M_1 \equiv_\alpha M_1'$ and $M_2 \equiv_\alpha M_2'$, then $(M_1M_2) \equiv_\alpha (M_1'M_2')$.

For instance, $\lambda x.y \equiv_\alpha \lambda z.y \not\equiv_\alpha \lambda y.y$. It is a common practice to consider λ-terms up to α-conversion, but we will not do this in our exposition.

β-reduction is the least binary relation β satisfying:

— If $M \equiv_\alpha (\lambda x.N)P$ and $N\langle x \leftarrow P\rangle = N'$, then $M\beta N'$.
— If $M\beta M'$, then $(\lambda x.M)\beta(\lambda x.M')$, $(MN)\beta(M'N)$ and $(NM)\beta(NM')$.

In the first rule, x is not necessarily free in N, so we may have $N = N'$ – in particular, free variables may disappear in the process of reduction.

We write β^* for the reflexive and transitive closure of β. A term M is *in β-normal form*, or is *β-normal*, if there is no M' such that $M\beta M'$. A term M is *normalising* if there is a normal N, which is called the *normal form* of M, such that $M\beta^*N$. It is *strongly normalising* if there is no infinite sequence $M = M_0\beta M_1\beta M_2\ldots$

It is well known that β-conversion enjoys the *Church–Rosser property*: if $M\beta^*N$ and $M\beta^*N'$, then there exist two α-convertible P, P' such that $N\beta^*P$ and $N'\beta^*P'$. As a consequence, if a term is normalising, its normal form is unique up to α-conversion.

The judgment 'assuming $x_1,\ldots,x_n$ are of types $\psi_1,\ldots\psi_n$, the term M has type ϕ', written $\{x_1 : \psi_1,\ldots,x_n : \psi_n\} \vdash M : \phi$, where $\psi_1,\ldots,\psi_n,\phi$ are formulas of propositional calculus and $x_1,\ldots,x_n$ are distinct variables, is defined by:

— $\Gamma \vdash x : \psi$ for each $x : \psi \in \Gamma$.
— If $\Gamma \cup \{x : \phi\} \vdash M : \psi$, then $\Gamma \vdash \lambda x.M : \phi \to \psi$.
— If $\Gamma \vdash M : \phi \to \psi$ and $\Gamma \vdash N : \phi$, then $\Gamma \vdash (MN) : \psi$

The *simply typable terms* are all M for which there exist Γ, ϕ such that $\Gamma \vdash M : \phi$. Note that Γ contains all variables free in M. The following properties are well known:

— **Strong normalisation:** If $\Gamma \vdash M : \phi$, then M is strongly normalising.
— **Subject reduction:** If $\Gamma \vdash M : \phi$ and $M\beta N$, then $\Gamma \vdash N : \phi$.

2. From BB′IW to simply typed lambda calculus

The aim of this section is to provide a precise characterisation of simply typable terms that are typable with inhabited types in BB′IW so as to transform the problem of type inhabitation in BB′IW into a type inhabitation problem in lambda calculus. The types of atomic combinators in BB′IW are also types for their respective counterparts $\lambda fgx.f(gx)$, $\lambda fgx.g(fx)$, $\lambda x.x$, $\lambda hx.hxx$ in lambda calculus, so to each inhabited type ϕ in BB′IW, there corresponds at least one closed λ-term of type ϕ. Moreover, subject reduction and strong normalisation (see above) also ensure the existence of a closed normal λ-term of type ϕ.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

What we lack is a criterion to distinguish amongst all typed normal forms those that are reducts of translations of combinators within $\mathsf{BB'IW}$.

The material and results of this section are not new (Bunder 1996; Broda *et al.* 2004), and the contents of Sections 2.3 and 2.4, apart from Lemma 2.10, may be skipped entirely to go immediately to the study of stable parts and blueprints in Section 3.

The definition of hereditarily right-maximal terms is an adaptation of the definition given in Bunder (1996). The proof of Lemma 2.6 (subject reduction for HRM-terms) is similar to the proof of Property 2.4 on page 375 of Broda *et al.* (2004). The right-to-left implication of Lemma 2.10 can be deduced from Property 2.20 on page 390 of Broda *et al.* (2004), though our proof method seems to be simpler.

2.1. *Lambda calculus*

Let $\mathscr{X}$ be a countably infinite set of variables $x, y, z \ldots$ together with a one-to-one function $\mathscr{O}$ from $\mathscr{X}$ to $\mathbb{N}$. For all x, y in $\mathscr{X}$, we write $x < y$ when $\mathscr{O}(x) < \mathscr{O}(y)$. Throughout the rest of the paper, when we write *term*, we will always mean a term of lambda calculus built over these variables. For each term M, we write $\mathsf{Free}(M)$ for the strictly increasing sequence of all free variables of M.

Terms are *not* identified modulo α-conversion – apart from in Section 2, all terms considered will be in normal form, and we will even use the Greek letters α and β with a new meaning at the beginning of Section 3. However, we do adopt the usual convention according to which two distinct λ's may not bind the same variable in a term, and no variable can be simultaneously free and bound in the same term.

2.2. *Hereditarily right-maximal terms*

Definition 2.1. The set of *hereditarily right-maximal* (HRM) terms is inductively defined as follows:

(1) Each variable x is HRM.
(2) If M is HRM and x is the greatest free variable of M, then $\lambda x.M$ is HRM.
(3) If M, N are HRM, and for each free variable x of M there exists a free variable y of N such that $x \leqslant y$, then (MN) is HRM.

The second rule ensures that all HRM-terms are λ_I-terms, that is, terms in which every subterm $\lambda x.M$ is such that x is free in M. As a consequence, the set of free variables of an HRM-term is preserved under β-reduction. As we shall see in Lemma 2.6, right-maximality can also be preserved at the cost of appropriate bound variable renamings.

In the third rule, if N is closed, then so is M. When M and N are non-closed terms, the greatest free variable of M is less than or equal to the greatest free variable of N. For instance, if $f < g < x$ and $h < x$, then $\lambda fgx.f(gx)$, $\lambda fgx.g(fx)$, $\lambda x.x$, $\lambda hx.hxx$ are HRM, whereas $\lambda yz.zy$ is not, no matter whether $y < z$ or $y > z$.

Definition 2.2. Let Ω be a function mapping each variable to a formula in such a way that $\Omega^{-1}(\phi)$ is an infinite set for each ϕ. We extend this function to the set of all strictly increasing finite sequences of variables by letting $\Omega(x_1, \ldots, x_n) = (\Omega(x_1), \ldots, \Omega(x_n))$.

Definition 2.3. The judgment $M : \phi$ (in words 'M has type ϕ with respect to Ω') is defined by:

— If $\Omega(x) = \phi$, then $x : \phi$.
— If $x : \chi$, $M : \psi$ and $\lambda x.M$ is HRM, then $\lambda x.M : \chi \to \psi$.
— If $M : \chi \to \psi$, $N : \chi$ and (MN) is HRM, then $(MN) : \psi$.

The function Ω will remain fixed throughout the exposition. Accordingly, the type of a term M with respect to Ω will be called *the* type of M without any further reference to the choice of Ω. Note that every typed term is HRM.

Definition 2.4. We write Λ_{NF} for the set of all typed terms in β-normal form. We say any closed term $M \in \Lambda_{\mathrm{NF}}$ of type ϕ is an Λ_{NF}-*inhabitant of* ϕ.

The next lemma is the well-known subformula property of simply typed lambda calculus:

Lemma 2.5 (subformula property). Let M be a Λ_{NF}-inhabitant of ϕ. The types of the subterms of M are subformulas of ϕ.

2.3. *Subject reduction of hereditarily right-maximal terms*

Lemma 2.6. If there exists a closed $M : \phi$, then ϕ is Λ_{NF}-inhabited.

Proof.

(1) The proof of the fact that for every variable y and for every $N : \phi$, there exists $N' \equiv_\alpha N$ such that $N' : \phi$ and every bound variable of N' is strictly greater than y is left as an exercise.

(2) We prove the following proposition by induction on P. Let P, Q be typed HRM-terms. Suppose:

— x and Q have the same type,

— If Q is closed and $x \in \mathsf{Free}(P)$, then $x = \min(\mathsf{Free}(P))$

— If Q is not closed, then for all $z \in \mathsf{Free}(P)$:

 – If $z < x$, then $z \leqslant \max(\mathsf{Free}(Q))$.

 – If $x < z$, then $\max(\mathsf{Free}(Q)) < z$.

 – if Q is not closed, then $\max(\mathsf{Free}(Q)) < z$ for all bound variables z of P.

Then $R = P\langle x \leftarrow Q\rangle$ is defined, HRM and have the same type as P. We consider cases:

— P is a variable:
The proposition is clear in this case

— $P = \lambda z.P'$:
Then $\mathsf{Free}(P') = \mathsf{Free}(P) \cdot (z)$. By the induction hypothesis, $R' = P'\langle x \leftarrow Q\rangle$ is defined, HRM and have the same type as P'. The variable z is still the greatest free variable of R' and z is not free in Q, hence $R = \lambda z.R'$.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

— $P = (P_1P_2)$:
By the induction hypothesis, $R_i = P_i\langle x \leftarrow Q\rangle$ is defined, HRM and have the same type as P_i for each $i \in \{1,2\}$. We still need to check that $R = (R_1R_2)$ is HRM. Assume x is free in P and P_1 is not closed. There are three sub-cases:

– $\max(\mathsf{Free}(P_1)) > x$:
Then

$$\max(\mathsf{Free}(P_1)) = \max(\mathsf{Free}(R_1)) \leqslant \max(\mathsf{Free}(P_2)) = \max(\mathsf{Free}(R_2)).$$

– $\max(\mathsf{Free}(P_1)) < x$:
The term Q cannot be closed, and

$$\max(\mathsf{Free}(P_1)) = \max(\mathsf{Free}(R_1)) \leqslant \max(\mathsf{Free}(Q)).$$

We have either

$$\max(\mathsf{Free}(P_2)) = x$$
$$\max(\mathsf{Free}(R_2)) = \max(\mathsf{Free}(Q))$$

or

$$\max(\mathsf{Free}(P_2)) > x$$
$$\max(\mathsf{Free}(P_2)) = \max(\mathsf{Free}(R_2)).$$

– $\max(\mathsf{Free}(P_1)) = x$:
There are two subcases:

- $\max(\mathsf{Free}(P_2)) > x$:
Then $\max(\mathsf{Free}(P_2)) = \max(\mathsf{Free}(R_2))$. If Q is closed, then $\mathsf{Free}(P_1) = (x)$ and R_1 is closed. Otherwise,

$$\max(\mathsf{Free}(R_1)) = \max(\mathsf{Free}(Q)) \leqslant \max(\mathsf{Free}(P_2)).$$

- $\max(\mathsf{Free}(P_2)) = x$:
If Q is closed, then $\mathsf{Free}(P_1) = \mathsf{Free}(P_2) = (x)$ and R_1, R_2 are closed. Otherwise,

$$\max(\mathsf{Free}(R_1)) = \max(\mathsf{Free}(R_2) = \max(\mathsf{Free}(Q)).$$

(3) Assume $N : \phi$ and N is not in normal form. We will prove the existence of $N' : \phi$ such that $N\beta N'$ by induction on N.
If $N = \lambda x.P$ or $N = (N_1N_2)$ with N_1 or N_2 not in normal form, then the existence of N' follows from the induction hypothesis and the fact that β-reduction preserves the set of free variables of an HRM-term. Otherwise, $N = (\lambda x.P)Q$, where for each free variable z of $\lambda x.P$, we have $z < x$ and there exists a free variable y of Q such that $z < y$. By (1), there exists $P' \equiv_\alpha P$ such that $P' : \phi$ and no bound variable of P' is less than or equal to a free variable of Q. The variable x is the greatest free variable of P'. By (2), the term $N' = P'\langle x \leftarrow Q\rangle$ is well defined, HRM and of the type ϕ. Moreover, $N\beta N'$.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

(4) We can now prove the lemma. The term M is a simply typable HRM-term. The strong normalisation property implies the existence of a normal form N of M. The term N is still a closed term. Finally, by (1), there exists $N' \equiv_\alpha N$ such that $N' : \phi$, that is, ϕ is Λ_{NF}-inhabited. □

2.4. *Equivalence between inhabitation in* BB′IW *and* Λ_{NF}*-inhabitation*

In the next three lemmas we write $\phi_1 \ldots \phi_n \to \psi$ to mean the formula

$$(\phi_1 \to (\ldots(\phi_n \to \psi)\ldots))$$

if $n > 0$, and the formula ψ otherwise. We write $\vdash_{\mathsf{BB'IW}} \phi$ for the judgment 'there exists within the basis BB′IW a combinator of type ϕ'.

Lemma 2.7. If $\vdash_{\mathsf{BB'IW}} \phi$, then ϕ is Λ_{NF}-inhabited.

Proof. If $f < g < x$ and $h < x$, then $\lambda x.x$, $\lambda fgx.f(gx)$, $\lambda fgx.g(fx)$ and $\lambda hx.hxx$ are HRM. For each type ϕ of an atomic combinator, the variables f, g, h, x can be chosen so that one of these terms has type ϕ. The set of all formulas ϕ for which there exists a closed M of type ϕ is closed under *modus ponens*. By Lemma 2.6, every such formula is Λ_{NF}-inhabited. □

Lemma 2.8. If $\vdash_{\mathsf{BB'IW}} \chi \to \psi$, then

$$\vdash_{\mathsf{BB'IW}} (\phi_1 \ldots \phi_n \to \chi) \to (\phi_1 \ldots \phi_n \to \psi)$$

for all $\phi_1, \ldots, \phi_n$.

Proof. The proof is by induction on n using left-applications of B. □

Lemma 2.9. Suppose $(i_1, \ldots, i_n)$, $(j_1, \ldots, j_m)$, $(k_1, \ldots, k_p)$ are strictly increasing sequences of integers, $\{k_1, \ldots, k_p\} = \{i_1, \ldots, i_n, j_1, \ldots, j_m\}$, $n = 0$ or ($n > 0$, $m > 0$, $i_n \leqslant j_m$). If

(1) $\vdash_{\mathsf{BB'IW}} \omega_{i_1} \ldots \omega_{i_n} \to (\chi \to \psi)$,

(2) $\vdash_{\mathsf{BB'IW}} \omega_{j_1} \ldots \omega_{j_m} \to \chi$,

then $\vdash_{\mathsf{BB'IW}} \omega_{k_1} \ldots \omega_{k_p} \to \psi$.

Proof. The proof is by induction on $n{+}m$. The proposition is obviously true if $n = m = 0$, so we assume $n + m > 0$. So $m > 0$. We consider cases:

— $n = 0$:

Then $(j_i, \ldots, j_m) = (k_1, \ldots, k_p)$. We have:

$$\vdash_{\mathsf{BB'IW}} (\chi \to \psi) \to ((\omega_{j_m} \to \chi) \to (\omega_{j_m} \to \psi)) \quad \text{(i)}$$

$$\vdash_{\mathsf{BB'IW}} (\omega_{j_m} \to \chi) \to (\omega_{j_m} \to \psi), \quad \text{(ii)}$$

where (i) is a type for B and (ii) follows from (i), (1) and *modus ponens*.

If $m = 1$, then $\vdash_{\mathsf{BB'IW}} \omega_{j_1} \to \psi$ follows from (ii), (2) and *modus ponens*. Otherwise,

$$\vdash_{\mathsf{BB'IW}} \omega_{j_1} \ldots \omega_{j_m} \to \psi$$

follows from (ii), (2) and the induction hypothesis.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

— $n > 0$:
We consider two sub-cases:

– $m > 1$ and $i_n \leqslant j_{m-1}$:
Hence:

$$\vdash_{\mathsf{BB'IW}} (\chi \to \psi) \to ((\omega_{j_m} \to \chi) \to (\omega_{j_m} \to \psi)) \qquad \text{(iii)}$$
$$\vdash_{\mathsf{BB'IW}} (\omega_{i_1} \dots \omega_{i_n} \to (\chi \to \psi)) \to (\omega_{i_1} \dots \omega_{i_n} \to ((\omega_{j_m} \to \chi) \to (\omega_{j_m} \to \psi))) \qquad \text{(iv)}$$
$$\vdash_{\mathsf{BB'IW}} \omega_{i_1} \dots \omega_{i_n} \to ((\omega_{j_m} \to \chi) \to (\omega_{j_m} \to \psi)), \qquad \text{(v)}$$

where (iii) is a type for B and (iv) follows from (iii) and Lemma 2.8. (v) then follows from (iv), (1) and *modus ponens*.
We now have $k_p = j_m$ and

$$\{k_1, \dots, k_{p-1}\} = \{i_1, \dots, i_n, j_1, \dots, j_{m-1}\}.$$

Since $i_n \leqslant j_{m-1}$, we have

$$\vdash_{\mathsf{BB'IW}} \omega_{k_1} \dots \omega_{k_{p-1}} \to (\omega_{j_m} \to \psi)$$

by (v), (2) and the induction hypothesis.

– $m = 1$ or ($m > 1$ and $i_n > j_{m-1}$):
Hence:

$$\vdash_{\mathsf{BB'IW}} (\omega_{j_m} \to \chi) \to ((\chi \to \psi) \to (\omega_{j_m} \to \psi)) \qquad \text{(vi)}$$
$$\vdash_{\mathsf{BB'IW}} (\omega_{j_1} \dots \omega_{j_m} \to \chi) \to (\omega_{j_1} \dots \omega_{j_{m-1}} \to ((\chi \to \psi) \to (\omega_{j_m} \to \psi))) \qquad \text{(vii)}$$
$$\vdash_{\mathsf{BB'IW}} \omega_{j_1} \dots \omega_{j_{m-1}} \to ((\chi \to \psi) \to (\omega_{j_m} \to \psi)) \qquad \text{(viii)}$$
$$\vdash_{\mathsf{BB'IW}} \omega_{n_1} \dots \omega_{n_q} \to (\omega_{j_m} \to \psi), \qquad \text{(ix)}$$

where (vi) is a type for B' and (vii) follows from (vi) and Lemma 2.8. (viii) then follows from (vii), (2) and *modus ponens*. Then, writing

$$\{n_1, \dots, n_q\} = \{j_1, \dots, j_{m-1}, i_1, \dots, i_n\},$$

(ix) follows from (viii), (1) and the induction hypothesis.
If $j_m > i_n$,

$$(n_1, \dots, n_q, j_m) = (k_1, \dots, k_p).$$

Otherwise,

$$j_m = i_n$$
$$n_q = i_n$$
$$(n_1, \dots n_q) = (k_1, \dots, k_p)$$

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

and

$$\vdash_{\mathsf{BB'IW}} \omega_{k_1} \dots \omega_{k_{p-1}} \to (\omega_{i_n} \to (\omega_{i_n} \to \psi)) \quad \text{(x)}$$

$$\vdash_{\mathsf{BB'IW}} (\omega_{i_n} \to (\omega_{i_n} \to \psi)) \to (\omega_{i_n} \to \psi) \quad \text{(xi)}$$

$$\vdash_{\mathsf{BB'IW}} (\omega_{k_1} \dots \omega_{k_{p-1}} \to (\omega_{i_n} \to (\omega_{i_n} \to \psi))) \to (\omega_{k_1} \dots \omega_{k_{p-1}} \to (\omega_{i_n} \to \psi)) \quad \text{(xii)}$$

$$\vdash_{\mathsf{BB'IW}} \omega_{k_1} \dots \omega_{k_{p-1}} \to (\omega_{i_n} \to \psi), \quad \text{(xiii)}$$

where (x) is just (ix), and (xi) is a type for W. Then (xii) follows from (xi) and Lemma 2.8, and (xiii) follows from (x), (xii) and *modus ponens*.
Finally, (xiii) is just $\vdash_{\mathsf{BB'IW}} \omega_{k_1} \dots \omega_{k_p} \to \psi$. □

Lemma 2.10. For every formula ϕ, we have $\vdash_{\mathsf{BB'IW}} \phi$ if and only if ϕ is Λ_{NF}-inhabited.

Proof. The left to right implication is Lemma 2.7. Using Lemma 2.9 when M is an application, an immediate induction on M shows that if $M : \psi$, $\mathsf{Free}(M) = (x_1, \dots, x_n)$ and $x_1 : \chi_1, \dots, x_n : \chi_n$, then $\vdash_{\mathsf{BB'IW}} \chi_1 \dots \chi_n \to \psi$. □

3. Stable parts and blueprints

3.1. *Introduction*

Lemma 2.10 showed that the decidability of type inhabitation for BB′IW is equivalent to the decidability of Λ_{NF}-inhabitation. The rest of the paper is devoted to the elaboration of a decision algorithm for the latter problem.

The problem we shall examine throughout Sections 3 and 4 is that if an inhabitant is not of minimal size, is there any way to transform it (with the help of grafts and/or another compression of some sort) into a smaller inhabitant with the same type? This question is not easy because we are dealing with a lambda calculus restricted with strong structural constraints (right-maximality). There are, however, simple situations in which an inhabitant is obviously not of minimal size.

Consider a Λ_{NF}-inhabitant M and two subterms N, P of M such that P is a strict subterm of N. Suppose:

— N, P are applications with the same type or abstractions with the same type.
— $\mathsf{Free}(N) = X = (x_1, \dots, x_n)$.
— $\mathsf{Free}(P) = Y = (y^1_0, \dots, y^1_{p_1}, \dots, y^n_0, \dots, y^n_{p_n})$.
— $\Omega(X) = (\chi_1, \dots, \chi_n)$.
— $\Omega(Y) = (\chi^1_0, \dots, \chi^1_{p_1}, \dots, \chi^n_0, \dots, \chi^n_{p_n})$.
— $\chi^i_j = \chi_i$ for each i, j.

In this case M is not of minimal size. Indeed, we can rename the free variables of P (letting $\rho(y^i_j) = x_i$) so as to obtain a term P' with the same size as P and the same type, and with the same free variables as N. The subterm N of M can be replaced with P' in M. The resulting term is a Λ_{NF}-inhabitant with the same type but of strictly smaller size.

However, this simple property is far from sufficient to characterise the minimal inhabitants of a formula: there are indeed formulas with inhabitants of abitrary size in which this situation never occurs. What we need is a more flexible way to reduce the size of non-minimal inhabitants. In particular, we need a better understanding of the available freedom of action if we are to rename the free variables of a term – possibly occurrence by occurrence – and if we want to ensure that right-maximality is preserved. This section is devoted to the proof of two key lemmas that delimit this freedom:

— In Sections 3.2, 3.3 and 3.3 we show how to build from any term $M \in \Lambda_{NF}$ a partial tree labelled with formulas. This partial tree is called the *blueprint* of M. This blueprint can be seen as a synthesised version of M that contains all and only the information required to determine whether a (non-uniform) renaming of the free variables of M will preserve hereditarily right-maximality.
— In Sections 3.5 and 3.6 we introduce a rewriting relation on blueprints that allows us to 'extract' sequences of formulas from a blueprint.
— In section 3.7 we prove our two key lemmas:
  – Lemma 3.15 clarifies the link between the blueprints of M and $\lambda x.M$ (provided both are in Λ_{NF}). This lemma proves, in particular, that the sequence of the types of the free variables of M (that is, $\Omega(\mathsf{Free}(M))$) can always be extracted from its blueprint.
  – Lemma 3.16 shows that for *every* sequence of formulas $\overline{\phi}$ that can be extracted from the blueprint of M, there exists a (non-uniform) renaming of the free variables of M that will produce a term N with the same type and the same blueprint as M, and such that $\Omega(\mathsf{Free}(N)) = \overline{\phi}$.

As a continuation of our first example, we will examine the consequences of this last result. Consider again a Λ_{NF}-inhabitant M and two subterms N, P of M such that P is a strict subterm of N and N, P are applications with the same type or abstractions with the same type. Suppose:

— The sequence $\Omega(\mathsf{Free}(N))$ can be extracted from the blueprint of P.

This situation is a generalisation of the preceding one (in our first example $\Omega(X)$ could also be extracted from the blueprint of P – see Definition 3.10). The term M is still not of minimal size. Indeed, we may use the second key lemma to prove the existence of a (non-uniform) renaming of the free variables of P that will produce a term P' with the same type as P such that $\mathsf{Free}(P') = \mathsf{Free}(N)$. The term N can be replaced with P' in M.

3.2. *Partial trees and trees*

Definition 3.1. Let $(\mathbb{A}, \leqslant)$ be the set of all finite sequences over the set $\mathbb{N}_+$ of natural numbers, ordered by prefix ordering. Elements of $\mathbb{A}$ are called *addresses*. We define a *partial tree* to be any function π whose domain is a set of addresses. For each partial tree π and for each address a, we let $\pi_{|a}$ denote the partial tree $c \mapsto \pi(a \cdot c)$ of domain $\{c \mid a \cdot c \in \mathrm{dom}(\pi)\}$.

Definition 3.2. For all partial trees π, π' and for every address a, we use $\pi[a \leftarrow \pi']$ to denote the partial tree π'' such that $\pi''_{|a} = \pi'$ and $\pi''(b) = \pi(b)$ for all $b \in \mathrm{dom}(\pi)$ such that $a \not\leqslant b$.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

Definition 3.3. A *tree domain* is a set $A \subseteq \mathbb{A}$ such that for all $a \in A$:

— Every prefix of a is in A.
— For every integer $i > 0$, if $a \cdot (i) \in A$, then $a \cdot (j) \in A$ for each $j \in \{1, \ldots, i-1\}$.

A tree domain A is *finitely branching* if and only if for each $a \in A$, there exists an $i > 0$ such that $a \cdot (i)$ is undefined. We define a *tree* to be any function whose domain is a tree domain.

In the rest of this paper, terms will be freely identified with trees. We identify:

— x with the tree mapping ε to x;
— $\lambda x.M$ with the tree τ mapping ε to λx and such that $\tau_{|(1)}$ is the tree of M;
— $(M_1 M_2)$ with the tree τ mapping ε to @ and such that $\tau_{|(i)}$ is the tree of M_i for each $i \in \{1, 2\}$.

3.3. *Blueprints*

Definition 3.4. Let $\mathfrak{S}$ be the signature consisting of all formulas and all symbols of the form $@_\phi$ where ϕ is a formula. Each formula is considered as a symbol of null arity. Each $@_\phi$ has arity 2.

We define a *blueprint* to be any finite partial tree $\alpha : A \to \mathfrak{S}$ satisfying the condition that for each $a \in A$, if $\alpha(a) = @_\phi$, then $\alpha_{|a\cdot(1)}$ and $\alpha_{|a\cdot(2)}$ are of non-empty domains. We define a *rooted* blueprint to be a blueprint α such that $\varepsilon \in \mathrm{dom}(\alpha)$.

For each $\mathscr{S} \subseteq \mathfrak{S}$, we define a *$\mathscr{S}$-blueprint* to be any blueprint whose image is a subset of $\mathscr{S}$. We write $\mathbb{B}(\mathscr{S})$ for the set of all $\mathscr{S}$-blueprints, and $\mathbb{B}_\varepsilon(\mathscr{S})$ for the set of all rooted $\mathscr{S}$-blueprints.

Definition 3.5. For every blueprint α and every address a, the *relative depth of a in α* is the number of $b \in \mathrm{dom}(\alpha)$ such that $b < a$. The *relative depth of* α is defined as 0 if α has an empty domain, and the maximal relative depth of an address in α otherwise.

In the rest of this paper, we will use the following notation for blueprints (see Figure 2):

— $\varnothing_{\mathbb{B}}$ denotes the blueprint of the empty domain.
— ϕ abbreviates $\varepsilon \mapsto \phi$.
— $@_\phi(\alpha_1, \alpha_2)$ denotes the (rooted) blueprint α such that

$$\alpha(\varepsilon) = \phi$$
$$\alpha_{|(1)} = \alpha_1$$
$$\alpha_{|(2)} = \alpha_1.$$

— For any sequence $\overline{a} = (a_1, \ldots, a_k)$ of pairwise incomparable addresses, $*_{\overline{a}}(\alpha_1, \ldots, \alpha_k)$ denotes the blueprint α of minimal domain such that $\alpha_{|a_i} = \alpha_i$ for each $i \in [1, \ldots, k]$.
— $*(\alpha_1, \ldots, \alpha_k)$ denotes the blueprint $*_{\overline{a}}(\alpha_1, \ldots, \alpha_k)$ such that $\overline{a} = ((1), \ldots, (k))$.

For each blueprint α, the choice of $\overline{a}, \alpha_1, \ldots, \alpha_k$ such that $\alpha = *_{\overline{a}}(\alpha_1, \ldots, \alpha_k)$ is obviously not unique. The sequence $(\alpha_1, \ldots, \alpha_k)$ may contain an arbitrary number of empty blueprints, hence the sequence $\overline{a}$ may be of arbitrary length. Also, α can be roooted (if $k = 1$, $a_1 = \varepsilon$ and α_1 is rooted) or empty (if $k = 0$ or $\alpha_1 = \cdots = \alpha_k = \varnothing_{\mathbb{B}}$). These ambiguities will not

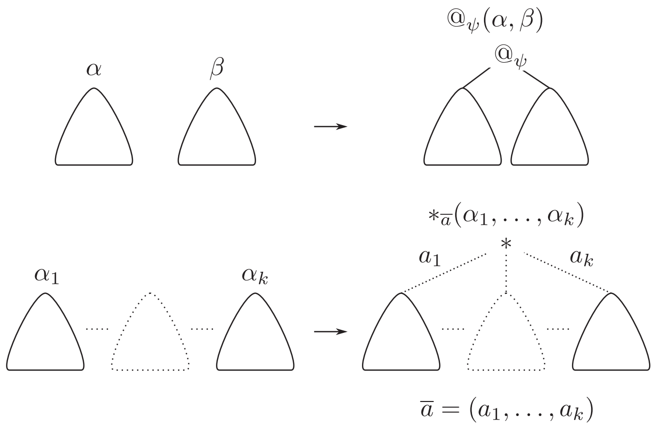

Fig. 2. The construction of blueprints using the notation of Section 3.3. In the upper diagram, the blueprints α and β must be non-empty. Although $\alpha_1, \dots, \alpha_k$ are displayed from left to right, the sequence $(a_1, \dots, a_k)$ need not be lexicographically ordered.

be difficult to deal with, but they will require us to take a few precautions in our proofs and definitions by induction on blueprints.

3.4. *Blueprint of a term*

Definition 3.6. For all $M \in \Lambda_{\mathrm{NF}}$, the *stable part* of M is the set of all $a \in \mathrm{dom}(M)$ such that $\mathsf{Free}(M_{|a}) \subseteq \mathsf{Free}(M)$ and $M_{|a}$ is a variable or an application.

It is easy to check that our conventions (no variable is simultaneously free and bound in a term) ensure that the stable part of a term does not depend on the choice of variable names. Since M is in normal form, M has an empty stable part if and only if it is closed.

Definition 3.7. For all $M \in \Lambda_{\mathrm{NF}}$, we define the *blueprint of M* to be the function α mapping each a in the stable part of M to:

— ψ if $M_{|a}$ is a variable of type ψ,
— $@_\psi$ if $M_{|a}$ is an application of type ψ.

We will use $M \Vdash \alpha$ to denote the judgment 'M has blueprint α' (Figure 3).

If $M = (M_1 M_2) \in \Lambda_{\mathrm{NF}}$, $M : \phi$, $M_1 \Vdash \alpha_1$, $M_2 \Vdash \alpha_2$, then each α_i has a non-empty domain and $(M_1 M_2) \Vdash @_\phi(\alpha_1, \alpha_2)$ – in other words, the so-called blueprint of M is indeed a blueprint, provided the blueprints of M_1 and M_2 are blueprints. When $M = \lambda x.M_1$, the blueprint of M has the form $*(\alpha)$ – the relation between α and the blueprint of M_1 in this case will be clarified by Lemma 3.15.

Lemma 3.8. For all $M \in \Lambda_{\mathrm{NF}}$ and forall $a \cdot b \in \mathrm{dom}(M)$:

(1) If $\mathsf{Free}(M_{|a\cdot b}) \subseteq \mathsf{Free}(M)$, then $\mathsf{Free}(M_{|a\cdot b}) \subseteq \mathsf{Free}(M_{|a})$.
(2) If $M_{|a} \Vdash \alpha$ and $M_{|a\cdot b} \Vdash \beta$, then $\alpha_{|b} = \beta$.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

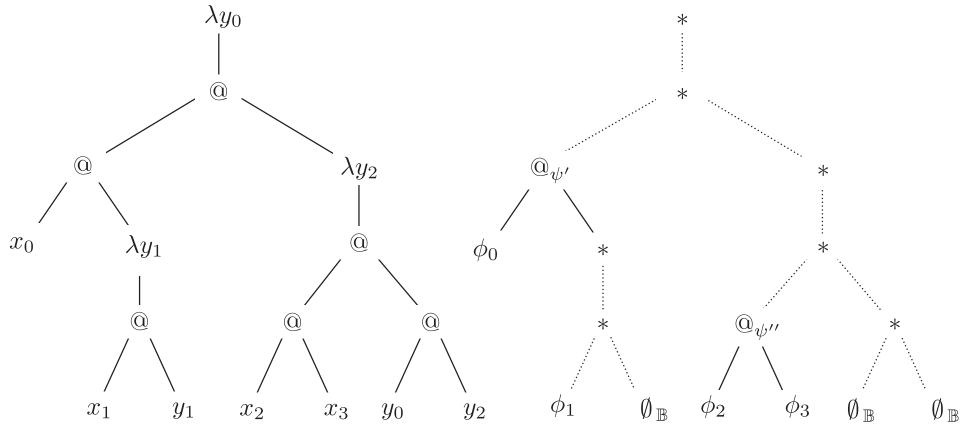

Fig. 3. An element of Λ_{NF} with its blueprint ($x_0 < x_1 < y_1$, $x_2 < x_3 < y_0 < y_2$, $x_1 < y_0 < y_2$).

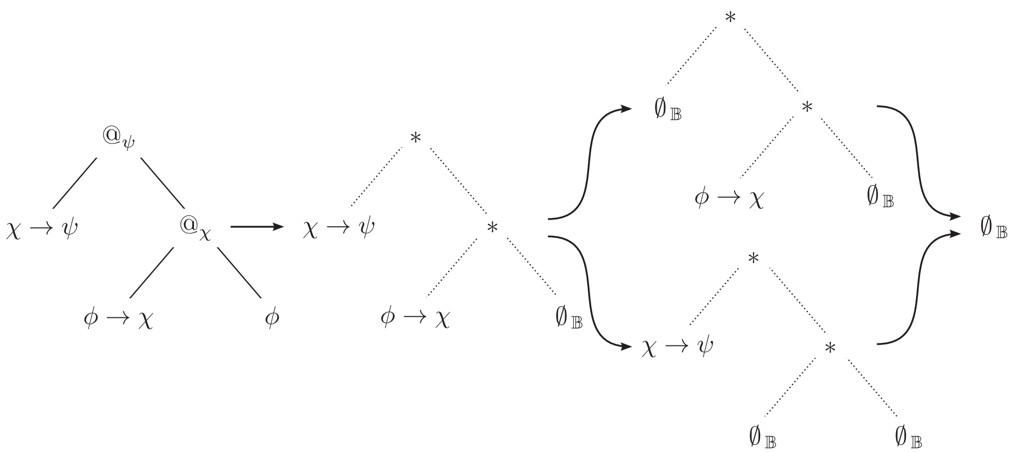

Fig. 4. Full reductions of $@_\psi(\chi \to \psi, @_\chi(\phi \to \chi, \phi))$ to $\emptyset_{\mathbb{B}}$.

Proof. The first part is a consequence of our bound variable convention (see Section 2.1), since if

$$\mathsf{Free}(M) = X$$
$$\mathsf{Free}(M_{|a}) = X' \cup Y$$

where $X' \subseteq X$ and X, Y are disjoint, then every element of $\mathsf{Free}(M_{|a\cdot b})$ in X is also an element of X'. Thus if $a \cdot b$ is in the stable part of M, then b is also in the stable part of $M_{|a}$.

The second part is equivalent to the first. □

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

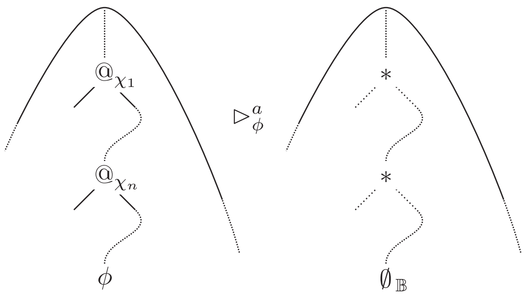

Fig. 5. Principle of blueprint reduction.

3.5. *Extraction of the formulas of a blueprint*

Definition 3.9. The judgment 'β is the blueprint obtained by extracting the formula ϕ at the address a in the blueprint α' (written $\alpha \rhd^a_\phi \beta$) is defined inductively by:

(1) $\phi \rhd^\varepsilon_\phi \varnothing_{\mathbb{B}}$.
(2) If $\alpha \rhd^a_\phi \beta$, then $@_\psi(\gamma, \alpha) \rhd^{(2)\cdot a}_\phi *(\gamma, \beta)$.
(3) If $\alpha \rhd^a_\phi \beta$, then $*_{(b,c_1,\ldots,c_k)}(\alpha, \gamma_1, \ldots, \gamma_k) \rhd^{b\cdot a}_\phi *_{(b,c_1,\ldots,c_k)}(\beta, \gamma_1, \ldots, \gamma_n)$.

In (2) we assume, of course, that α and γ are non-empty. In (3) we assume $b \neq \varepsilon$ in order to avoid circularity.

For instance (see Figure 4):

— $$@_\psi(\chi \to \psi, @_\chi(\phi \to \chi, \phi)) \rhd^{(2,2)}_\phi *(\chi \to \psi, *(\phi \to \chi, \varnothing_{\mathbb{B}})) \rhd^{(2,1)}_{\phi\to\chi} *(\chi \to \psi, *(\varnothing_{\mathbb{B}}, \varnothing_{\mathbb{B}})) \rhd^{(1)}_{\chi\to\psi} *(\varnothing_{\mathbb{B}}, *(\varnothing_{\mathbb{B}}, \varnothing_{\mathbb{B}})) = \varnothing_{\mathbb{B}}.$$

— $$@_\psi(\chi \to \psi, @_\chi(\phi \to \chi, \phi)) \rhd^{(2,2)}_\phi *(\chi \to \psi, *(\phi \to \chi, \varnothing_{\mathbb{B}})) \rhd^{(1)}_{\chi\to\psi} *(\varnothing_{\mathbb{B}}, *(\phi \to \chi, \varnothing_{\mathbb{B}})) \rhd^{(2,1)}_{\phi\to\chi} *(\varnothing_{\mathbb{B}}, *(\varnothing_{\mathbb{B}}, \varnothing_{\mathbb{B}})) = \varnothing_{\mathbb{B}}.$$

When $\alpha \rhd^a_\phi \beta$, the blueprint β can be seen as α in which the formula ϕ at a is erased together with all @'s in the path to a. At each @, this path must follow the right-hand branch of @ (see Figure 5). The constraints on the construction of blueprints imply the existence of at least one such path in every non-empty blueprint, even if it is not the blueprint of a term.

3.6. *Sets of extractible sequences*

Definition 3.10. For each formula ϕ, let $\rhd_\phi$ be the relation defined by $\alpha \rhd_\phi \beta$ if and only if there exists a such that $\alpha \rhd^a_\phi \beta$. We write $\rhd^+_\phi$ for the transitive closure of $\rhd_\phi$. For each α, we write $\mathbb{F}(\alpha)$ for the set of all sequences $(\phi_1, \ldots, \phi_n)$ such that $\alpha \rhd^+_{\phi_n} \cdots \rhd^+_{\phi_1} \varnothing_{\mathbb{B}}$.

The set $\mathbb{F}(\alpha)$ is what we called 'set of extractible sequences of α' in Section 3.1. Note that $\mathbb{F}(\varnothing_{\mathbb{B}}) = \{\varepsilon\}$. If $\alpha \neq \varnothing_{\mathbb{B}}$, then all elements of $\mathbb{F}(\alpha)$ are non-empty sequences. Note also that each $\rhd$-reduction strictly decreases the cardinality of the domain of a blueprint,

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

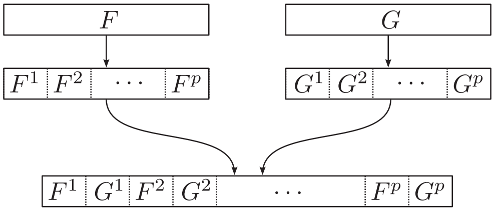

Fig. 6. Shuffling of two sequences. The chunks of F and G need not have the same size – some of them can be empty. Every contraction of the resulting sequence belongs to $\circledast(F, G)$. Each contraction also belongs to $\circledcirc(F, G)$ when F, G are non-empty and the last chunk G^p of G is non-empty.

therefore $\mathbb{F}(\alpha)$ is a finite set for all α. We now introduce the notion of a *shuffle*, which will allow us to characterise $\mathbb{F}(\alpha)$ according to the structure of α.

Definition 3.11. A *contraction* of a sequence F is either the sequence F or a sequence $G \cdot (f) \cdot H$ where $G \cdot (f) \cdot (f) \cdot H$ is a contraction of F.

Definition 3.12. For all finite sequences $F_1, \ldots, F_n$, we define a *shuffle* of $(F_1, \ldots, F_n)$ to be any sequence $F_1^1 \cdot \cdots \cdot F_n^1 \cdot \cdots \cdot F_1^p \cdot \cdots \cdot F_n^p$ such that $F_i^1 \cdot \cdots \cdot F_i^p = F_i$ for each i. For each tuple of sets of finite sequences $(\mathscr{F}_1, \ldots, \mathscr{F}_n)$, we write $\circledast(\mathscr{F}_1, \ldots, \mathscr{F}_n)$ for the closure under contraction of the set of shuffles of elements of $\mathscr{F}_1 \times \cdots \times \mathscr{F}_n$.

Definition 3.13. Given two non-empty finite sequences F_1, F_2, we derfine a *right-shuffle* of (F_1, F_2) to be any sequence $F_1^1 \cdot F_2^1 \cdot \cdots \cdot F_1^p \cdot F_2^p$ such that $F_i^1 \cdot \ldots F_i^p = F_i$ for each i and $F_2^p \neq \varepsilon$. For each pair of sets of non-empty finite sequences $(\mathscr{F}_1, \mathscr{F}_2)$, we write $\circledcirc(\mathscr{F}_1, \mathscr{F}_2)$ for the closure under contraction of the set of right-shuffles of elements $\mathscr{F}_1 \times \mathscr{F}_2$.

The principle of (right-)shuffling is shown in Figure 6. The following properties follow from our definitions, and will be used without further reference:

(1) If $\alpha = \varnothing_{\mathbb{B}}$, then $\mathbb{F}(\alpha) = \{\varepsilon\}$.
(2) If $\alpha = \phi$, then $\mathbb{F}(\alpha) = \{(\phi)\}$.
(3) If $\alpha = *_{\overline{a}}(\alpha_1, \ldots, \alpha_k)$, then $\mathbb{F}(\alpha) = \circledast(\mathbb{F}(\alpha_1), \ldots, \mathbb{F}(\alpha_k))$.
(4) If $\alpha = @_{\phi}(\alpha_1, \alpha_2)$, then $\mathbb{F}(\alpha) = \circledcirc(\mathbb{F}(\alpha_1), \mathbb{F}(\alpha_2))$.

3.7. *Abstraction versus extraction*

Lemma 3.14. Suppose $\{a_1, \ldots, a_p\} = \{b_1, \ldots, b_p\}$, and:
— $\alpha \rhd_{\chi}^{a_1} \cdots \rhd_{\chi}^{a_p} \beta$.
— $\alpha \rhd_{\chi}^{b_1} \cdots \rhd_{\chi}^{b_p} \beta'$.
Then $\beta = \beta'$.

Proof. The proof is by an easy induction on α. □

Recall that for every strictly increasing sequence of variables $X = (x_1, \ldots, x_n)$, we write $\Omega(X)$ for the sequence of the types of $x_1, \ldots, x_n$. We will now clarify the link between the blueprint α of a term M and the blueprint of $\lambda x.M$.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

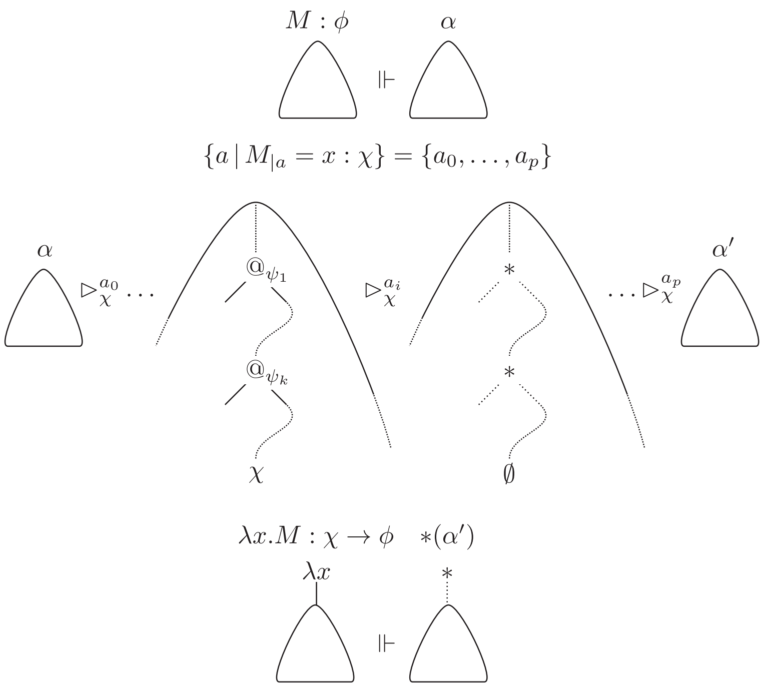

Fig. 7. How the blueprint of M evolves into the blueprint of $\lambda x.M$

In particular, the next lemma shows that if $M, \lambda x.M \in \Lambda_{\mathrm{NF}}$, then M and $\lambda x.M$ have blueprints α and β if and only if there exist $a_0, \dots, a_p$ such that

$$\{a_0, \dots, a_p\} = \{a \mid M_{|a} \ = \ x\}$$
$$\alpha \rhd^{a_0}_{\chi} \cdots \rhd^{a_p}_{\chi} \alpha'$$
$$\beta = *(\alpha')$$

(see Figure 7).

Lemma 3.15. Suppose $M \in \Lambda_{\mathrm{NF}}$ has blueprint α, with $\mathsf{Free}(M) = (x_1, \dots, x_n)$ and $\Omega(x_1, \dots, x_n) = (\chi_1, \dots, \chi_n)$. For each $i \in [0, \dots, n]$:

— Let α_i be the restriction of α to $\mathrm{dom}(\alpha) \cap \{a \mid \mathsf{Free}(M_{|a}) \subseteq \{x_1, \dots, x_i\}\}$.
— Let β_i be the blueprint of $\lambda x_{i+1} \dots x_n.M$.

Then:

(1) For each $i \in [0, \dots, n]$, we have $\mathrm{dom}(\beta_i) = \{1^{n-1} \cdot a \mid a \in \mathrm{dom}(\alpha_i)\}$ and $\beta_{i|1^{n-i}} = \alpha_i$.
(2) For each $i \in]0, \dots, n]$:
 (a) There exist $a^i_0, \dots, a^i_{p_i}$ such that

$$\{a^i_0, \dots, a^i_{p_i}\} = \{a \mid M_{|a} = x_i\}$$

and

$$\alpha_i \rhd^{a^i_0}_{\chi_i} \cdots \rhd^{a^i_{p_i}}_{\chi_i} \alpha_{i-1}.$$

(b) If

$$\{b_0, \dots, b_{p_i}\} = \{a \mid M_{|a} = x_i\}$$

and

$$\alpha_i \rhd^{b_0}_{\chi_i} \cdots \rhd^{b_{p_i}}_{\chi_i} \alpha',$$

then $\alpha' = \alpha_{i-1}$.

(3) $(\chi_1, \dots, \chi_n) \in \mathbb{F}(\alpha)$.

Proof. Property (1) follows immediately from the definition of a blueprint. Since $\alpha_n = \alpha$ and $\alpha_0 = \varnothing_{\mathbb{B}}$, Property (3) follows from Property (2a). Property (2b) follows from Property (2a) and Lemma 3.14. To prove (2a), we first introduce the following notation.

For each $N \in \Lambda_{\mathrm{NF}}$, we let ρ_N be the least partial function satisfying the following conditions:

— For every blueprint γ, we have $\rho_N(\varepsilon, \gamma) = \gamma$.
— For every finite sequence of variables Y and for every blueprint γ, if

$$\rho_N(Y, \gamma) = \delta$$
$$\{b \mid N_{|b} = y\} = \{b_0, \dots, b_m\}$$

and

$$\delta \rhd^{b_0}_{\chi} \cdots \rhd^{b_m}_{\chi} \delta',$$

then $\rho_M((y) \cdot Y, \gamma) = \delta'$.

By Lemma 3.14, if

$$\{b \mid N_{|b} = y\} = \{b_0, \dots, b_m\} = \{c_0, \dots, c_m\}$$

and

$$\delta \rhd^{b_0}_{\chi} \cdots \rhd^{b_m}_{\chi} \delta'$$
$$\delta \rhd^{c_0}_{\chi} \cdots \rhd^{c_m}_{\chi} \delta'',$$

then $\delta' = \delta''$, so ρ_N is indeed a function. For each finite sequence of variables Y' and for each blueprint γ, we let $\mu_N(Y', \gamma)$ be the restriction of γ to $\mathrm{dom}(\gamma) \cap \{b \mid \mathsf{Free}(N_{|b}) \subseteq Y'\}$.

We shall prove by induction on M that for all pairs (X, X') such that $\mathsf{Free}(M) = X \cdot X'$, we have $\mu_M(X, \alpha) = \rho_N(X', \alpha)$. In particular, for all $i > 0$, we have

$$\begin{aligned}
\alpha_{i-1} &= \mu_M((x_1, \dots, x_{i-1}), \alpha) \\
&= \rho_M((x_i, \dots, x_n), \alpha) \\
&= \rho_M((x_i), \rho_M((x_{i+1} \dots, x_n), \alpha)) \\
&= \rho_M((x_i), \mu_M((x_1 \dots, x_i), \alpha)) \\
&= \rho_M((x_i), \alpha_i),
\end{aligned}$$

so (2a) holds.

The case $X' = \varepsilon$ is immediate, so we will assume that X' is a non-empty suffix of $\mathsf{Free}(M)$ and consider cases:

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

— M is equal to a variable:
This case follows immediately from the definitions.

— $M = (M_1 M_2)$, $M_1 \Vdash \gamma_1$ and $M_2 \Vdash \gamma_2$:
There exist X_1, X_2, X'_1, X'_2 such that:

$$\begin{aligned} X_1 \cup X_2 &= X \\ X'_1 \cup X'_2 &= X' \\ \mathsf{Free}(M_j) &= X_j \cdot X'_j \quad \text{for each } j \in \{1,2\}. \end{aligned}$$

We have $\alpha = @_\psi(\gamma_1, \gamma_2)$ where ψ is the type of M, and

$$\mu_M(X, \alpha) = *(\mu_{M_1}(X_1, \gamma_1), \mu_{M_2}(X_2, \gamma_2)).$$

By the induction hypothesis,

$$\mu_{M_i}(X_i, \gamma_i) = \rho_{M_i}(X'_i, \gamma_i)$$

for each i. The sequence X' is non-empty, so the final elements of X' and X'_2 are equal. Assume $X' = X'' \cdot (x)$ and $X'_2 = X''_2 \cdot (x)$. If x is not the final element of X'_1,

$$\begin{aligned} \rho_M(X', \alpha) &= \rho_M(X'' \cdot (x), @_\psi(\gamma_1, \gamma_2)) \\ &= \rho_M(X'_1 \cup X''_2, *(\gamma_1, \rho_{M_2}((x), \gamma_2))) \\ &= *(\rho_{M_1}(X'_1, \gamma_1), \rho_{M_2}(X''_2, \rho_{M_2}((x), \gamma_2))) \\ &= *(\rho_{M_1}(X'_1, \gamma_1), \rho_{M_2}(X''_2 \cdot (x), \gamma_2)) \\ &= *(\rho_{M_1}(X'_1, \gamma_1), \rho_{M_2}(X'_2, \gamma_2)). \end{aligned}$$

Otherwise, $X'_1 = X''_1 \cdot (x)$ and we have

$$\begin{aligned} \rho_M(X', \alpha) &= \rho_M(X', @_\psi(\gamma_1, \gamma_2)) \\ &= \rho_M(X''_1 \cup X''_2, *(\rho_{M_1}((x), \gamma_1), \rho_{M_2}((x), \gamma_2))) \\ &= *(\rho_{M_1}(X''_1, \rho_{M_1}((x), \gamma_1)), \rho_{M_2}(X''_2, \rho_{M_2}((x), \gamma_2))) \\ &= *(\rho_{M_1}(X''_1 \cdot (x), \gamma_1), \rho_{M_2}(X''_2 \cdot (x), \gamma_2)) \\ &= *(\rho_{M_1}(X'_1, \gamma_1), \rho_{M_2}(X'_2, \gamma_2)). \end{aligned}$$

In either case

$$\begin{aligned} \rho_M(X', \alpha) &= *(\rho_{M_1}(X'_1, \gamma_1), \rho_{M_2}(X'_2, \gamma_2)) \\ &= *(\mu_{M_1}(X_1, \gamma_1), \mu_{M_2}(X_2, \gamma_2)) \\ &= \mu_M(X, \alpha). \end{aligned}$$

— $M = \lambda x.M_1$, $M_1 \Vdash \gamma_1$:
By the induction hypothesis,

$$\begin{aligned} \mu_{M_1}(X, \gamma_1) &= \rho_{M_1}(X' \cdot (x), \gamma_1) \\ &= \rho_{M_1}(X', \rho_{M_1}((x), \gamma_1)) \\ &= \rho_{M_1}(X', \mu(X \cdot X', \gamma_1)) \\ &= \rho_{M_1}(X', \alpha_{|(1)}). \end{aligned}$$

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

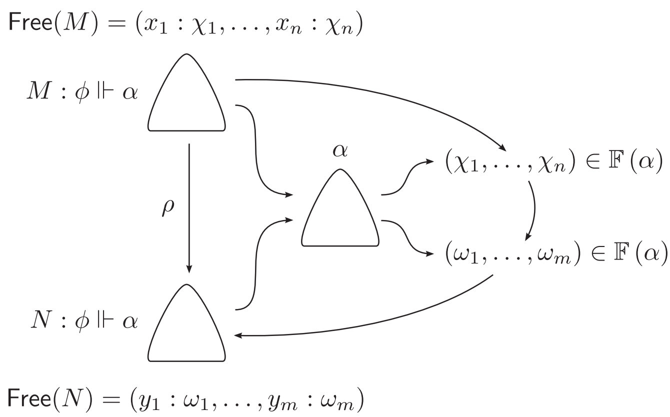

Fig. 8. A non-uniform renaming of the variables of M, based on an alternate extraction of the formulas of its blueprint.

Moreover,

$$\mu_{M_1}(X, \gamma_1) = \mu_{M_1}(X, \mu_{M_1}(X \cdot X', \gamma_1)) = \mu_{M_1}(X, \alpha_{|(1)}).$$

Therefore,

$$\mu_{M_1}(X, \alpha_{|(1)}) = \rho_{M_1}(X', \alpha_{|(1)}),$$

so

$$\mu_{M_1}(X, \alpha) = \rho_{M_1}(X', \alpha). \qquad \square$$

Thus the full sequence of the types of the free variables of M can be extracted from its blueprint. The next lemma shows that, conversely, for each sequence $\overline{\chi}$ in $\mathbb{F}(\alpha)$, there exists a term N with the same domain and blueprint, that has the same type as M, and such that the sequence of types of the free variables of N is equal to $\overline{\chi}$ – see Figure 8.

Lemma 3.16. Let $M \in \Lambda_{\mathrm{NF}}$ be a term with blueprint α. Suppose

$$\alpha \rhd^{b_0^m}_{\omega_m} \cdots \rhd^{b_{p_m}^m}_{\omega_m} \quad \ldots \quad \rhd^{b_0^1}_{\omega_1} \cdots \rhd^{b_{p_1}^1}_{\omega_1} \varnothing_{\mathbb{B}}.$$

Then for every strictly increasing sequence of variables $Y = (y_1, \ldots, y_m)$ such that $\Omega(Y) = (\omega_1, \ldots, \omega_m)$, there exists N with the same domain and blueprint and with the same type as M, and such that $\mathsf{Free}(N) = Y$ and $\{b \mid N_{|b} = y_i\} = \{b_1^i, \ldots, b_{p_i}^i\}$ for each i.

Proof. The proof is by induction on M:

— M is a variable:
This case is obvious.

— $M = (M_1 M_2)$:
This case follows easily from the induction hypothesis.

— $M = \lambda x.M_1 : \phi \to \psi$ with $M_1 \Vdash \gamma$:
Let $Y' = (y_1, \ldots, y_m, x)$. By Lemma 3.15 (2a), there exist $a_1, \ldots, a_p$ such that

$$\{a_1, \ldots, a_p\} = \{a \mid M_{|a} = x\}$$

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

and

$$\gamma \rhd_{\phi}^{a_0} \cdots \rhd_{\phi}^{a_p} \gamma' = \alpha_{|1}.$$

Now

$$\alpha \rhd_{\omega_m}^{b_0^m} \cdots \rhd_{\omega_m}^{b_{pm}^m} \quad \ldots \quad \rhd_{\omega_1}^{b_0^1} \cdots \rhd_{\omega_1}^{b_{p1}^1} \varnothing_{\mathbb{B}},$$

so each b_j^i has the form $(1) \cdot c_j^i$. Furthermore,

$$\gamma \rhd_{\phi}^{a_0} \cdots \rhd_{\phi}^{a_p} \rhd_{\omega_m}^{c_0^m} \cdots \rhd_{\omega_m}^{c_{pm}^m} \quad \ldots \quad \rhd_{\omega_1}^{c_0^1} \cdots \rhd_{\omega_1}^{c_{p1}^1} \varnothing_{\mathbb{B}}.$$

By the induction hypothesis, there exists N_1 with the same domain and blueprint and with the same type as M_1 such that

$$\begin{aligned} \mathsf{Free}(N_1) &= Y' \\ \{a \mid N_{1|a} = x\} &= \{a_0, \ldots, a_p\} \\ \{c \mid N_{1|c} = y_i\} &= \{c_0^i, \ldots, c_{p_i}^i\} \text{ for each } i. \end{aligned}$$

By Lemma 3.15 (2b), we have $\lambda x.N_1 \Vdash \alpha$, so we may take $N = \lambda x.N_1$. □

4. Vertical compressions and compact terms

This section provides a partial characterisation of minimal inhabitants. In Section 4.1, we make a simple observation on the relative depths of their blueprints, together with the following easy consequence of the subformula property (Lemma 2.5). If M is a minimal Λ_{NF}-inhabitant of ϕ, then for all addresses a in M the blueprint of $M_{|a}$ has relative depth at most $k \times p$, where:

— k is the number of λ in the path from the root to M to a.
— p is the number of subformulas of ϕ.

We say any Λ_{NF}-inhabitant satisfying this condition is *locally compact*. In Section 4.2 we introduce the notion of a *vertical compression* of a blueprint. A (strict) vertical compression of β is obtained by taking any address b in β and then grafting $\beta_{|b}$ at any address $a < b$ such that $\beta(a) = \beta(b)$. The vertical compressions of β are all blueprints obtained by applying this transformation to β zero or more times. The key property of these compressions is (see Figure 9):

— If M has blueprint β and α is a vertical compression of β, the compression of β into α can be mimicked by a compression of M into an HRM-term in the following sense. Assuming $\alpha = \beta[a \leftarrow \beta_{|b}]$ (the base case), the term $Q = M[a \leftarrow M_{|b}]$ is *not* in general an HRM-term. However, *there exists* an HRM-term M' with the same domain as Q and with the same type as M. Moreover, M' and M are applications with the same type or abstractions with the same type.

We again consider a Λ_{NF}-inhabitant M and two addresses a, b such that $a < b$, and $M_{|a}$ and $M_{|b}$ are applications with the same type or abstractions with the same type. Suppose:

— there exists a vertical compression α' of the blueprint of $M_{|b}$ such that the sequence $\Omega(\mathsf{Free}(M_{|a}))$ can be extracted from α'.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

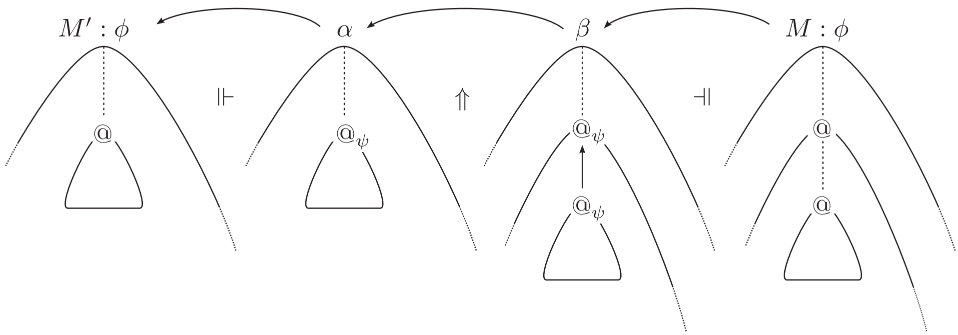

Fig. 9. How the compression of terms is able to follow the compression of blueprints.

This situation is a generalisation of the last example in Section 3.1 (in which α' was equal to the blueprint of $M_{|b}$, which is thereby a trivial compression of this blueprint). The term M is not minimal. Indeed, the key property above implies the existence of a term N of blueprint α' whose size is not greater than the size of $M_{|b}$, and such that $N, M_{|b}, M_{|a}$ are applications with the same type, or abstractions with the same type. By Lemma 3.16, there exists a term P with the same type and domain as N such that $\mathsf{Free}(P) = \mathsf{Free}(M_{|a})$. The graft of P at a yields an inhabitant of strictly smaller size.

We will say all inhabitants in which the preceding situation does not occur are *compact*. All inhabitants of minimal size are of course compact. As we shall see in Section 6, we will not need a sharper characterisation of minimal inhabitants. For every formula ϕ, the set of compact inhabitants of ϕ is actually a *finite* set, and our decision method will consist of the exhaustive computation of their domains.

4.1. *Depths of the blueprints of minimal inhabitants*

Definition 4.1. Two terms $M, M' \in \Lambda_{\mathrm{NF}}$ are *of the same kind* if and only if they are both variables or both applications, or both abstractions, and if they have the same type.

Definition 4.2. For all formulas ϕ, we write $\mathsf{Sub}(\phi)$ for the set of all subformulas of ϕ.

Definition 4.3. Let $M \in \Lambda_{\mathrm{NF}}$. Let a be any address in M. Let $(a_1, \ldots, a_m)$ be the strictly increasing sequence of all prefixes of a. Let $(\lambda x_1, \ldots, \lambda x_k)$ be the subsequence of $(M(a_1), \ldots, M(a_m))$ consisting of all labels of the form λx. We write $\Lambda(M, a)$ for $(x_1, \ldots, x_k)$.

Definition 4.4. Let M be a Λ_{NF}-inhabitant of ϕ. We say that M is *locally compact* if for all addresses a in M, the blueprint of $M_{|a}$ has relative depth at most $|\Lambda(M, a)| \times |\mathsf{Sub}(\phi)|$.

Lemma 4.5. Let M be a Λ_{NF}-inhabitant of ϕ. If M is not locally compact, then there exist two addresses b, b' such that $b < b'$, $M_{|b}$ and $M_{|b'}$ are of the same kind and $\mathsf{Free}(M_{|b}) = \mathsf{Free}(M_{|b'})$. Moreover, M is not a Λ_{NF}-inhabitant of ϕ of minimal size.

Proof. For each address a in $\mathrm{dom}(M)$, let α_a be the blueprint of $M_{|a}$ and let $X_a = \mathsf{Free}(M_{|a})$. We assume the existence of an α_a with relative depth

$$n > |\Lambda(M, a)| \times |\mathsf{Sub}(\phi)|.$$

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

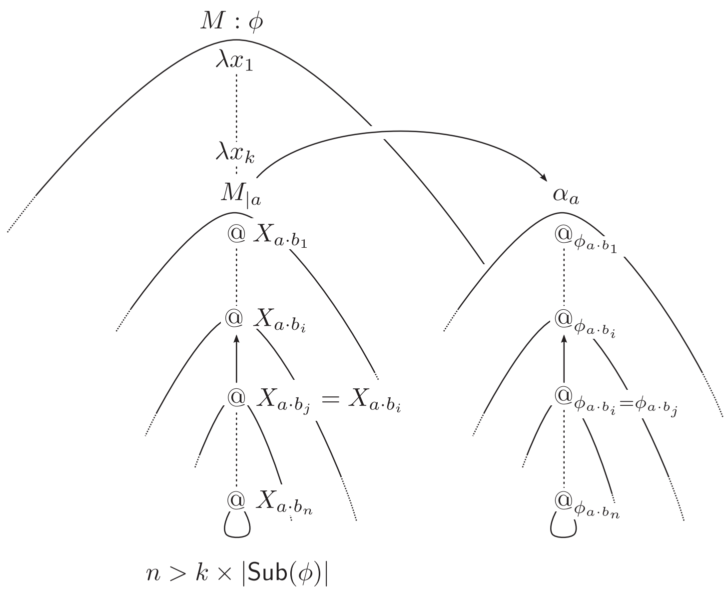

Fig. 10. Proof of Lemma 4.5.

So there exist $b_1, \dots, b_{n+1} \in \mathrm{dom}(\alpha_a)$ such that $b_1 < \cdots < b_n < b_{n+1}$. By Lemma 3.8 (1), we have

$$X_{a\cdot b_n} \subseteq \cdots \subseteq X_{a\cdot b_1} \subseteq \Lambda(M, a).$$

By Lemma 2.5, each $\phi_{a\cdot b_i}$ is a subformula of ϕ. Hence there exist i, j such that $i < j$ and

$$(X_{a\cdot b_i}, \phi_{a\cdot b_i}) = (X_{a\cdot b_j}, \phi_{a\cdot b_j}),$$

that is, $M_{|a\cdot b_i}$ and $M_{|a\cdot b_j}$ are applications with the same type and with the same free variables (see Figure 10). Now, let $M' = M[a \cdot b_i \leftarrow M_{|a\cdot b_j}]$. The term M' is a Λ_{NF}-inhabitant of ϕ of strictly smaller size. □

4.2. *Vertical compression of a blueprint*

Definition 4.6. Let $\Uparrow$ be the least reflexive and transitive binary relation on blueprints satisfying the condition that if $a, b \in \mathrm{dom}(\beta)$, $a < b$ and $\beta(a) = \beta(b)$, then $\beta[a \leftarrow \beta_{|b}] \Uparrow \beta$.

Lemma 4.7. Suppose $M \in \Lambda_{\mathrm{NF}}$, $M : \phi$, $M \Vdash \beta$ and $\alpha \Uparrow \beta$. There exists a term $M' \in \Lambda_{\mathrm{NF}}$ of the same kind as M, with blueprint α and such that $|\mathrm{dom}(M')| \leqslant |\mathrm{dom}(M)|$.

Proof. It is enough to consider the case of $\alpha = \beta[a \leftarrow \beta_{|b}]$ with $a, b \in \mathrm{dom}(\beta)$, $a < b$ and $\beta(a) = \beta(b)$.

We prove the existence of M' by induction on the length of a.

— $a = \varepsilon$:
In this case M is necessarily an application and $\beta(\varepsilon) = \beta(b) = @_\phi$, so $M_{|b}$ is an application with type ϕ, and we can take $M' = M_{|b}$.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

— $a \neq \varepsilon$.
We consider cases:

– $M = (M_1 M_2)$, $M_1 \Vdash \beta_1$, $M_2 \Vdash \beta_2$, $a = (i) \cdot a_i$ and $b = (i) \cdot b_i$:
By the induction hypothesis, there exists M_i' with blueprint

$$\alpha_i = \beta_i[a_i \leftarrow \beta_{i|b_i}] = \beta_i[a_i \leftarrow \beta_{|b}],$$

of the same kind as M_i and such that $\mathrm{dom}(M_i') \leqslant \mathrm{dom}(M_i)$. Let $j = 1$ if $i = 2$, otherwise, let $j = 2$. Let $(M_j', \alpha_j) = (M_j, \beta_j)$. Let $X = (x_1, \ldots, x_n)$ be the strictly increasing sequence of all variables free or bound in M_2'. Let $Y = (y_1, \ldots, y_n)$ be a strictly increasing sequence of variables such that $\Omega(X) = \Omega(Y)$ and y_1 is greater than or equal to the greatest variable of M_1'. Let M_2'' be the term obtained by replacing each x_i by y_i in M_2'. We can take $M' = (M_1' M_2'')$.

– $M = \lambda x.M_1$, $M_1 \Vdash \beta_1$, $x : \chi$, $a = (1) \cdot a_1$ and $b = (1) \cdot b_1$:
Since $a, b \in \mathrm{dom}(\beta)$, we also have $a_1, b_1 \in \mathrm{dom}(\beta_1)$. By the induction hypothesis, there exists M_1' of the same kind as M_1, with blueprint

$$\alpha_1 = \beta_1[a_1 \leftarrow \beta_{1|b_1}]$$

and such that $\mathrm{dom}(M_1') \leqslant \mathrm{dom}(M_1)$. By Lemma 3.15 (2a), there exist $\gamma_1, c_0, \ldots, c_p$ such that

$$\{c_0, \ldots, c_p\} = \{c \mid M_{|c} = x\}$$
$$\beta_1 \rhd_\chi^{c_0} \cdots \rhd_\chi^{c_p} \gamma_1$$
$$\beta = *(\gamma_1).$$

Since $a, b \in \mathrm{dom}(\alpha)$, we know a_1 and c_i are incomparable addresses for all i. Hence

$$\alpha_1 = \beta_1[a_1 \leftarrow \beta_{1|b_1}] \rhd_\chi^{c_0} \cdots \rhd_\chi^{c_p} \gamma_1[a_1 \leftarrow \beta_{1|b_1}] = \beta[a \leftarrow \beta_{|b}]_{|(1)} = \alpha_{|1}.$$

By Lemma 3.16, there exists a term M_1'' with the same type and domain as M_1' such that the greatest variable y free in M_1'' has type χ and

$$\{c \mid M''_{1\,|c} = y\} = \{c_0, \ldots, c_p\}.$$

By Lemma 3.15 (2b), we have $\lambda y.M_1'' \Vdash \alpha$, so we may take $M' = \lambda y.M_1''$. □

Definition 4.8. We say a term $M \in \Lambda_{\mathrm{NF}}$ is *compact* if there are no a, b, α' such that $a < b$, $M_{|a}$ and $M_{|b}$ are of the same kind, $M_{|b} \Vdash \alpha_b$, $\alpha' \Uparrow \alpha_b$ and $\Omega(\mathsf{Free}(M_{|a})) \in \mathbb{F}(\alpha')$.

Lemma 4.9.

(1) Every Λ_{NF}-inhabitant of minimal size is compact.
(2) Every compact Λ_{NF}-inhabitant of ϕ is locally compact.

Proof. Let M by an arbitrary Λ_{NF}-inhabitant of ϕ.

(1) Assume M is not compact. Let a, b be such that $a < b$, $M_{|a}$ and $M_{|b}$ are of the same kind, $M_{|b} \Vdash \alpha_b$, $\alpha' \Uparrow \alpha_b$, $\mathsf{Free}(M_{|a}) = X_a$ and $\Omega(X_a) \in \mathbb{F}(\alpha')$ – see Figure 11. By Lemma 4.7, there exists a term $N \in \Lambda_{\mathrm{NF}}$ with blueprint α', of the same kind as $M_{|b}$ and such that $|\mathrm{dom}(N)| \leqslant |\mathrm{dom}(M_{|b})|$. By Lemma 3.16, there exists $P \in \Lambda_{\mathrm{NF}}$ with

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

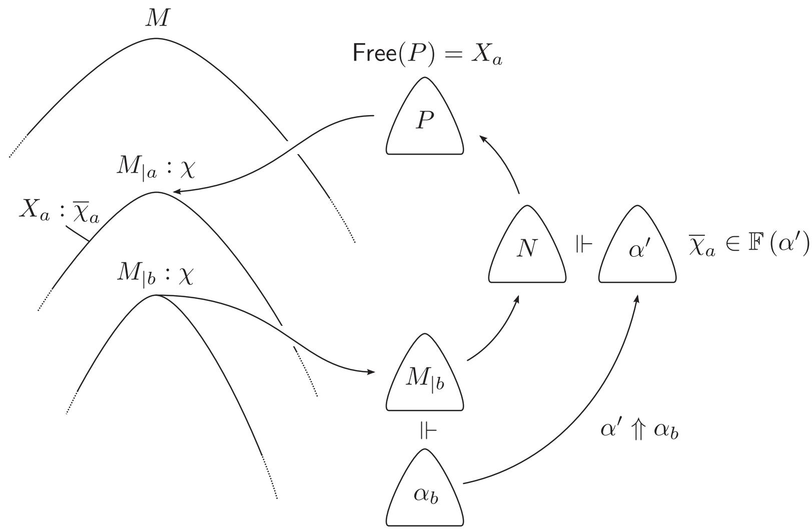

Fig. 11. Proof of Lemma 4.9, part (1).

blueprint α', of the same kind as N, such that $\mathrm{dom}(P) = \mathrm{dom}(N)$ and $\mathsf{Free}(P) = X_a$. The term $M[a \leftarrow P]$ is then a Λ_{NF}-inhabitant of ϕ of smaller size.

(2) Suppose M meets the conditions of Lemma 4.5. Let $\alpha_{b'}$ be the blueprint of $M_{|b'}$. By Lemma 3.15 (3), we have $\Omega(\mathsf{Free}(M_{|b})) = \Omega(\mathsf{Free}(M_{|b'}) \in \mathbb{F}(\alpha_{b'})$. Since the relation $\Uparrow$ is reflexive, M is not compact. $\square$

5. Shadows

So far we have isolated two properties shared by all minimal inhabitants (Lemma 4.9). We shall now exploit these properties to design a decision method for the inhabitation problem.

In Sections 5.1 and 5.2 we show how to associate with each locally compact inhabitant M of a formula ϕ, a tree with the same domain as M, which we call the *shadow* of M. At each address a, this tree is labelled with a triple of the form $(\overline{\chi}_a, \gamma_a, \phi_a)$ where ϕ_a is the type of $M_{|a}$, the sequence $\overline{\chi}_a$ is $\Omega(\mathsf{Free}(M_{|a}))$, and γ_a is a 'transversal compression' of the blueprint α_a of $M_{|a}$ (Definitions 5.1 and 5.2). Recall that $\overline{\chi}_a \in \mathbb{F}(\alpha_a)$ – see Lemma 3.15 (3). The blueprint γ_a can be viewed as a synthesised version of α_a with the same relative depth but smaller 'width', and such that $\overline{\chi}_a \in \mathbb{F}(\gamma_a) \subseteq \mathbb{F}(\alpha_a)$.

Each tree prefix of the shadow of M belongs to a finite set that is effectively computable from ϕ and the domain of this prefix. In particular, we can compute all possible values for its labels, even without full knowledge of M – or even without knowledge of the existence of M. The key property satisfied by this shadow at every address a is:

— For each $\gamma' \Uparrow \gamma_a$, there exists $\alpha' \Uparrow \alpha_a$ such that $\mathbb{F}(\gamma') \subseteq \mathbb{F}(\alpha')$.

This property is sufficient to detect the non-compactness of M for a pair of addresses (a, b) simply from the knowledge of $\overline{\chi}_a, \phi_a, \gamma_b, \phi_b$ and the arity of the nodes at a and b.

Indeed, we suppose $a < b$, $\phi_a = \phi_b$ and the nodes at a and b have the same arity (1 or 2), and assume:

— There exists $\gamma' \Uparrow \gamma_b$ such that $\overline{\chi}_a \in \mathbb{F}(\gamma')$.

Then $M_{|a}$ and $M_{|b}$ are of the same kind and there exists $\alpha' \Uparrow \alpha_b$ such that

$$\overline{\chi}_a = \Omega(\mathsf{Free}(M_{|a})) \in \mathbb{F}(\gamma') \subseteq \mathbb{F}(\alpha'),$$

therefore M is not compact.

In Section 5.2, what we call a *shadow* is merely a tree $a \mapsto (\overline{\chi}_a, \gamma_a, \phi_a)$ of a certain shape, regardless of whether this tree is the shadow of a term or not. This shadow is *compact* if there is no pair (a, b) as above. Of course, the shadow of a compact term is always compact in this sense.

In Section 6 we will prove that for every formula ϕ, the set of shadows of the compact inhabitants of ϕ is a finite set that is effectively computable from ϕ (hence the same property holds for the set of compact inhabitants of ϕ), and we will deduce from this key property the decidability of type inhabitation for HRM-terms.

5.1. *Blueprint equivalence and transversal compression*

Definition 5.1. We let $\equiv$ be the least binary relation on blueprints such that:

(1) $\varnothing_{\mathbb{B}} \equiv \varnothing_{\mathbb{B}}$.
(2) $\phi \equiv \phi$.
(3) If $\alpha_1 \equiv \beta_1$, $\alpha_2 \equiv \beta_2$, then $@_\phi(\alpha_1, \alpha_1) \equiv @_\phi(\beta_1, \beta_2)$.
(4) If $|\overline{a}| = |\overline{b}| = n$ and $\alpha_i \equiv \beta_i$ for each $i \in [1, \ldots, n]$, then

$$*_{\overline{a}}(\alpha_1, \ldots, \alpha_n) \equiv *_{\overline{b}}(\beta_1, \ldots, \beta_n).$$

In (3), we assume $\alpha_1, \alpha_2, \beta_1, \beta_2$ are non-empty, and in (4), we assume that the elements of each sequence $\overline{a}$, $\overline{b}$ are pairwise incomparable addresses. To avoid circularity, we also assume $a \neq \varepsilon$ or $b \neq \varepsilon$, and $\alpha_i, \beta_i \neq \varnothing_{\mathbb{B}}$ for at least one i.

To some extent, this equivalence allows us to consider blueprints regardless of the exact values of addresses. For instance,

$$*_{\overline{a}}(\alpha_1, \ldots, \alpha_n) \equiv *(\alpha_1, \ldots, \alpha_n) \equiv *(\alpha_n, \ldots, \alpha_1),$$

also

$$*(*(\alpha, \beta), \gamma) \equiv *(\alpha, \beta, \gamma) \equiv *(\alpha, *(\beta, \gamma)),$$

and so on. It is easy to check that $\alpha \equiv \beta$ implies $\mathbb{F}(\alpha) = \mathbb{F}(\beta)$ – this property will be used without further reference.

Definition 5.2. For each $m \in \mathbb{N}$, we let $\backsim_m$ be the least binary relation such that:

(1) if $\gamma_1 \equiv \cdots \equiv \gamma_m \equiv \gamma_{m+1} \not\equiv \varnothing_{\mathbb{B}}$, then $*_{\overline{a}}(\gamma_1, \ldots, \gamma_m) \backsim_m *_{\overline{a} \cdot (b)}(\gamma_1, \ldots, \gamma_m, \gamma_{m+1})$,
(2) if $\alpha = *_{\overline{a}}(\alpha_1, \ldots, \alpha_n)$, $\beta = *_{\overline{b}}(\beta_1, \ldots, \beta_p)$ and $\alpha \backsim_m \beta$, then:

  (a) $@_\phi(\alpha, \gamma) \backsim_m @_\phi(\beta, \gamma)$,

  (b) $@_\phi(\gamma, \alpha) \backsim_m @_\phi(\gamma, \beta)$,

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

(c) $*_{\overline{a}\cdot(c)}(\alpha_1,\ldots,\alpha_n,\gamma) \backsim_m *_{\overline{b}\cdot(c)}(\beta_1,\ldots,\beta_p,\gamma)$.

We define an *m-compression of* β to be any α such that $\alpha \backsim_m \beta$. The *width* of β is defined as the least $m \in \mathbb{N}$ for which there is no α such that $\alpha \backsim_m \beta$.

Again the elements of $\overline{a}\cdot(b)$, $\overline{a}\cdot(c)$ and $\overline{b}\cdot(c)$ must be pairwise incomparable addresses, and α,β,γ must be non-empty. Note that for all non-empty β, we have $\varnothing_{\mathbb{B}} \backsim_0 \beta$, so the empty blueprint is the only blueprint of null width. If β has width $m>0$, then for all addresses a, for $\beta_{|a} = *_{\overline{a}}(\gamma_1,\ldots,\gamma_k)$ and for each $\gamma_i \neq \varnothing_{\mathbb{B}}$, the sequence $(\gamma_1,\ldots,\gamma_k)$ contains no more than m blueprints $\equiv$-equivalent to γ_i. For instance, if ϕ,ψ,χ are distinct formulas, $*(\phi,\phi,\phi,\psi,\psi,\chi)$ has width 3, $*(\omega, @_\omega(*(\phi,\psi),\phi), @_\omega(*(\psi,\phi),\phi))$ has width 2, and so on.

Definition 5.3. For each $m \in \mathbb{N}$, we write $\sqsubseteq_m$ for the reflexive and transitive closure of the union of $\equiv$ and $\backsim_m$. We use $\sqsubseteq_m^{\max}$ to denote the subset of the relation $\sqsubseteq_m$ of all pairs with a left-hand side of width at most m.

For instance, if ϕ,ψ,χ are distinct formulas:

$$\varnothing_{\mathbb{B}} \sqsubseteq_0^{\max} *(\psi,\chi,\phi) \sqsubseteq_1^{\max} *(\chi,\phi,\phi,\psi,\psi) \sqsubseteq_2^{\max} *(\phi,\phi,\phi,\psi,\psi,\chi).$$

Of course, $\alpha \sqsubseteq_m \beta$ implies $\alpha \sqsubseteq_j \beta$ for all $j \in [1,\ldots,m]$ and, clearly, $\alpha \backsim_m \beta$ implies $|\mathrm{dom}(\alpha)| < |\mathrm{dom}(\beta)|$, so $\backsim_m$ is well founded.

Definition 5.4. For all $\mathscr{S} \subseteq \mathfrak{S}$, all $d \in \mathbb{N}$ and all $m \in \mathbb{N}$:

— We let $\mathbb{B}(\mathscr{S},d,\infty)$ be the set of $\mathscr{S}$-blueprints with relative depth at most d.
— We let $\mathbb{B}(\mathscr{S},d,m)$ be the set of all blueprints in $\mathbb{B}(\mathscr{S},d,\infty)$ of width at most m.

Lemma 5.5. For all finite $\mathscr{S} \subseteq \mathfrak{S}$, all $d \in \mathbb{N}$ and all $m \in \mathbb{N}$:

(1) The set $\mathbb{B}(\mathscr{S},d,m)/_{\equiv}$ is a finite set.
(2) A selector $\mathbb{R}(\mathscr{S},d,m)$ for $\mathbb{B}(\mathscr{S},d,m)/_{\equiv}$ is effectively computable from $(\mathscr{S},d,m)$.

Proof.

(1) Let $\mathbb{B}_\varepsilon(\mathscr{S},d,m)$ be the set of all rooted blueprints in $\mathbb{B}(\mathscr{S},d,m)$. Assuming that $\mathbb{B}_\varepsilon(\mathscr{S},d,m)/_{\equiv}$ is a finite set and that a selector $\mathbb{R}_\varepsilon(\mathscr{S},d,m)$ for $\mathbb{B}_\varepsilon(\mathscr{S},d,m)/_{\equiv}$ is effectively computable from $(\mathscr{S},d,m)$, we prove that $\mathbb{B}(\mathscr{S},d,m)/_{\equiv}$ and $\mathbb{B}_\varepsilon(\mathscr{S},d+1,m)/_{\equiv}$ are finite sets and show how to compute a selector for each set.
Let $(\alpha_1,\ldots,\alpha_k)$ be an enumeration of $\mathbb{R}_\varepsilon(\mathscr{S},d,m)$, and Σ_d be the set of all functions from $\{1,\ldots,k\}$ to $\{0,\ldots,m\}$. For each $\beta \in \mathbb{B}(\mathscr{S},d,m)$, there exist $\beta_1,\ldots,\beta_n \in \mathbb{B}_\varepsilon(\mathscr{S},d,m)$ and $\overline{b}$ such that $\beta = *_{\overline{b}}(\beta_1,\ldots,\beta_n)$. We let σ_β be the function mapping each $i \in \{1,\ldots,k\}$ to the number of occurrences of an element $\equiv$-equivalent to α_i in the sequence $(\beta_1,\ldots,\beta_n)$. Clearly, $\sigma_\beta \in \Sigma_d$ and, furthermore, for all $\beta' \in \mathbb{B}(\mathscr{S},d,m)$, we have $\beta \equiv \beta'$ if and only if $\sigma_\beta = \sigma_{\beta'}$, so $\mathbb{B}(\mathscr{S},d,m)$ is a finite set.
For each $\tau \in \Sigma_d$, let

$$\rho_\tau = *(\alpha_1^1,\ldots,\alpha_1^{\tau(1)},\ldots,\alpha_k^1,\ldots,\alpha_k^{\tau(k)})$$

where each α_i^j is equal to α_i. We have $\rho_\tau \in \mathbb{B}(\mathscr{S},d,m)$ and $\sigma(\rho_\tau) = \tau$, that is, if $\tau,\tau' \in \Sigma_d$ and $\tau \neq \tau'$, then $\rho_\tau \not\equiv \rho_{\tau'}$. Hence we may define $\mathbb{R}(\mathscr{S},d,m)$ as $\{\rho_\tau \mid \tau \in \Sigma_d\}$.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

Now the finiteness of $\mathbb{B}_\varepsilon(\mathscr{S}, d+1, m)/_\equiv$ follows immediately from the finiteness of $\mathbb{B}(\mathscr{S}, d, m)$ and the fact that if $\beta = @_\phi(\beta_1, \beta_2)$ and $\beta' = @_\psi(\beta'_1, \beta'_2)$ are elements of $\mathbb{B}_\varepsilon(\mathscr{S}, d+1, m)$, then $\beta_1, \beta_2, \beta'_1, \beta'_2$ are non-empty elements of $\mathbb{B}(\mathscr{S}, d, m)$ and, furthermore, $\beta \equiv \beta'$ if and only if $\beta_1 \equiv \beta'_1$ and $\beta_2 \equiv \beta'_2$. The same property allows us to define $\mathbb{R}_\varepsilon(\mathscr{S}, d+1, m)$ as the set of all blueprints of the form $@_\phi(\gamma_1, \gamma_2)$ where $@_\phi \in S$ and each γ_i is a non-empty element of $\mathbb{R}(\mathscr{S}, d, m)$.

(2) The second part of the lemma follows by induction on d, using part (1) and the fact that $\mathbb{B}_\varepsilon(\mathscr{S}, 0, 0)$ is empty (hence $\mathbb{B}(\mathscr{S}, d, 0) = \{\varnothing_{\mathbb{B}}\}$ for all d) and if $m \in \mathbb{N}_+$, then $\mathbb{B}_\varepsilon(\mathscr{S}, 0, m)$ is the finite set of all formulas of $\mathscr{S}$. □

5.2. *Shadow of a term*

Definition 5.6. Let ϕ be a formula and $\mathscr{S}_\phi$ be the union of $\mathsf{Sub}(\phi)$ (Definition 4.2) and the set of all $@_\psi$ such that $\psi \in \mathsf{Sub}(\phi)$. For each integer k and each formula ϕ, we let $\mathfrak{R}(\phi, k) = \mathbb{R}(\mathscr{S}_\phi, k \times |\mathsf{Sub}(\phi)|, k)$, where $\mathbb{R}$ is the function introduced in Lemma 5.5 (2).

Definition 5.7. A *shadow* is a finite tree in which each node has arity at most 2 and is labelled with a triple of the form $(\overline{\chi}, \gamma, \psi)$, where $\overline{\chi}$ is a sequence of formulas, γ is a blueprint and ψ is a formula.

We define a *ϕ-shadow* to be any shadow Ξ satisfying the following conditions, where we assume $\Xi(\varepsilon) = (\varepsilon, \varnothing_{\mathbb{B}}, \phi)$, and for each $a \in \mathrm{dom}(\Xi)$, we let k_a be the number of $b < a$ such that the node of Ξ at b is unary, and we let $(\overline{\chi}_a, \gamma_a, \psi_a) = \Xi(a)$:

— $\overline{\chi}_a$ is a sequence of subformulas of ϕ of length at most k_a.
— $\gamma_a \in \mathfrak{R}(\phi, k_a)$.
— $\overline{\chi}_a \in \mathbb{F}(\gamma_a)$.
— ψ_a is a subformula of ϕ.

Definition 5.8. Let M be a locally compact Λ_{NF}-inhabitant of ϕ. For each $a \in \mathrm{dom}(M)$, let:

— $\overline{\chi}_a = \Omega(\mathsf{Free}(M_{|a}))$;
— α_a be the blueprint of $M_{|a}$;
— $\gamma_a \in \mathfrak{R}(\phi, |\Lambda(M, a)|)$ be such that $\gamma_a \sqsubseteq^{\max}_{|\Lambda(M,a)|} \alpha_a$;
— ϕ_a be the type of $M_{|a}$.

We will say the tree Ξ mapping each $a \in \mathrm{dom}(M)$ to $(\overline{\chi}_a.\gamma_a, \phi_a)$ is *the shadow of M*.

Recall that if M is a locally compact Λ_{NF}-inhabitant of ϕ, then for each address a in M, the blueprint α_a of $M_{|a}$ has relative depth at most $|\Lambda(M, a)| \times |\mathsf{Sub}(\phi)|$. Every maximal $|\Lambda(M, a)|$-compression of α_a produces a shadow α'_a with the same relative depth and of width at most $|\Lambda(M, a)|$, to which some element of $\mathfrak{R}(\phi, |\Lambda(M, a)|)$ is equivalent, thus the shadow of M is well defined. Note that the choice of γ_a is not required to be unique (though it is, since $\mathbb{R}$ is a selector and we can actually prove that $\gamma \sqsubseteq^{\max}_m \alpha$ and $\gamma' \sqsubseteq^{\max}_m \alpha$ implies $\gamma \equiv \gamma'$, but this property is irrelevant to our discussion); we just assume that *some* γ_a is chosen for each address a in M (see Figure 12).

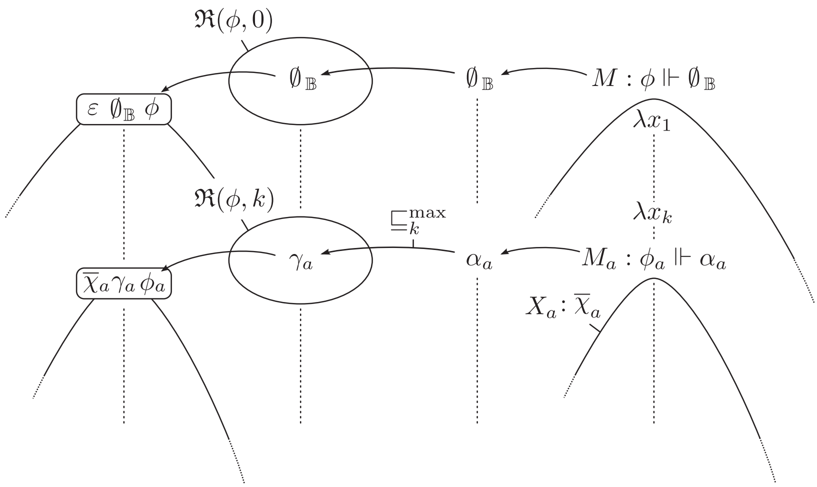

Fig. 12. A locally compact inhabitant and its shadow.

It is obvious that the shadow of M satisfies the first, second and fourth conditions in the definition of ϕ-shadows given above – in the next section, we prove that it also satisfies the third.

5.3. *Compact shadows and compact inhabitants*

Definition 5.9. A shadow Ξ is *compact* if and only if there are no a, b such that:

— $a < b$;
— the nodes of Ξ at a and b have the same arity;
— $\Xi(a) = (\overline{\chi}_a, \gamma_a, \psi)$;
— $\Xi(b) = (\overline{\chi}_b, \gamma_b, \psi)$;
— there exists $\gamma' \Uparrow \gamma_b$ such that $\overline{\chi}_a \in \mathbb{F}(\gamma')$.

This definition shoud be compared with the definition of compactness for a term (Definition 4.8). With the help of three auxiliary lemmas, we now prove the key lemma of Section 5: if M is a compact inhabitant – *a fortiori* locally compact by Lemma 4.9 – then the shadow of M is a compact ϕ-shadow.

Lemma 5.10. If $\alpha \Uparrow \beta \sqsubseteq_1 \beta'$, then there exists α' such that $\alpha \sqsubseteq_1 \alpha' \Uparrow \beta'$.

Proof.

(1) An immediate induction on $|\mathrm{dom}(\beta')|$ shows that if $\alpha = \beta[a \leftarrow \beta_{|b}]$ and $\beta \equiv \beta'$, then there exist a', b' such that $a' < b'$ and $\alpha \equiv \alpha' = \beta'[a' \leftarrow \beta'_{|b'}]$. As a consequence, an immediate induction on the length of the derivation of $\alpha \Uparrow \beta$ shows that the lemma holds if $\beta \equiv \beta'$.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

(2) Another induction on $|\text{dom}(\beta')|$ shows that if $\alpha \Uparrow \beta \backsim_1 \beta'$, then there exists α' such that $\alpha \backsim_1 \alpha' \Uparrow \beta'$. The only non-trivial case is

$$\begin{aligned} \alpha &= *_{(a_1)}(\alpha_1) \\ \beta &= *_{(a_1)}(\beta_1) && \text{with } \alpha_1 \Uparrow \beta_1 \\ \beta' &= *_{(a_1,a_2)}(\beta_1, \beta_2) && \text{with } \beta_1 \equiv \beta_2. \end{aligned}$$

Since $\alpha_1 \Uparrow \beta_1 \equiv \beta_2$, by (1) there exists α_2 such that $\alpha_1 \equiv \alpha_2 \Uparrow \beta_2$. Hence

$$\alpha = *_{(a_1)}(\alpha_1) \backsim_1 *_{(a_1,a_2)}(\alpha_1, \alpha_2) \Uparrow *_{(a_1,a_2)}(\beta_1, \beta_2) = \beta'.$$

(3) Using (1) and (2), the lemma follows by induction on the length of an arbitrary sequence $(\beta_0, \ldots, \beta_n)$ such that $\beta_0 = \beta$, $\beta_n = \beta'$ and $\beta_{i-1} \equiv \beta_i$ or $\beta_{i-1} \backsim_1 \beta_i$ for each $i \in [1, \ldots, n]$. □

Lemma 5.11. If $\alpha \sqsubseteq_1 \beta$, then $\mathbb{F}(\alpha) \subseteq \mathbb{F}(\beta)$.

Proof. The proof is by induction on $|\text{dom}(\beta)|$. Since $\gamma \equiv \gamma'$ implies $\mathbb{F}(\gamma) = \mathbb{F}(\gamma')$ and $|\text{dom}(\gamma)| = |\text{dom}(\gamma')|$, it is enough to consider the case where α is a 1-compression of β. The case $\alpha = *_{(a_1)}(\alpha_1)$ and $\beta = *_{(a_1,a_2)}(\alpha_1, \alpha_2)$ is clear, and the remaining cases follow easily from the induction hypothesis. □

Lemma 5.12. If $\alpha \sqsubseteq_m \beta$, the set of all elements of $\mathbb{F}(\beta)$ of length at most m is a subset of $\mathbb{F}(\alpha)$.

Proof. The proof is by induction on $|\text{dom}(\beta)|$. Again, we will only examine the case $\alpha \backsim_m \beta$. The proposition is trivially true if $m = 0$. So we assume $m > 0$. The only non-trivial case is

$$\begin{aligned} \alpha &\equiv *_{\overline{a}}(\gamma_1, \ldots, \gamma_m) \\ \beta &\equiv *_{\overline{a}}(\gamma_1, \ldots, \gamma_m, \gamma_{m+1}) \end{aligned}$$

with $\gamma_i \equiv \gamma$ for all i. Let $\Phi = \mathbb{F}(\gamma)$. For each integer k, let $\Phi^{(k)} = \circledast(\Phi_1, \ldots, \Phi_k)$ where $\Phi_i = \mathbb{F}(\gamma)$ for each i. Let $\overline{\phi} = (\phi_1, \ldots, \phi_p) \in \mathbb{F}(\beta)$ be such that $p \leqslant m$. We have to prove that $\overline{\phi} \in \mathbb{F}(\alpha)$. For each $J \subseteq \{1, \ldots, p\}$, let $(j_1, \ldots, j_q)$ be the strictly increasing enumeration of all elements of J and let $f(J) = (\phi_{j_1}, \ldots, \phi_{j_q})$. We have $\overline{\phi} \in \mathbb{F}(\beta) = \Phi^{(m+1)}$, so there exist $J_1, \ldots, J_{m+1}$ such that

$$J_1 \cup \cdots \cup J_{m+1} = \{1, \ldots, p\},$$

and $f(J_i) \in \mathbb{F}(\gamma)$ for each $i \in \{1, \ldots, m+1\}$. For each $j \in \{1, \ldots, p\}$, let k_j be any element of $\{1, \ldots, m+1\}$ such that $j \in J_{k_j}$. Then $J_{k_1} \cup \cdots \cup J_{k_p} = \{1, \ldots, p\}$, so

$$\overline{\phi} \in \circledast(\{f(J_{k_1})\}, \ldots, \{f(J_{k_p})\}) \subseteq \Phi^{(p)} \subseteq \Phi^{(m)} = \mathbb{F}(\alpha).$$

□

Lemma 5.13. Let M be a locally compact Λ_{NF}-inhabitant of ϕ.

(1) The shadow of M is a ϕ-shadow.
(2) If M is compact, this shadow is also compact.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

Proof.

(1) For each address a in M, the sequence $\overline{\chi}_a = \Omega(\mathsf{Free}(M_{|a}))$ is a subsequence of $\Omega(\Lambda(M, a))$, so the first part of the lemma follows from the definition of the shadow of M, Lemma 2.5, Lemma 3.15 (3) and Lemma 5.12.

(2) Let Ξ be shadow of M and assume Ξ is not compact. There exist $a, b \in \mathrm{dom}(\Xi) = \mathrm{dom}(M)$ such that:

— $\Xi(a) = (\overline{\chi}_a, \gamma_a, \psi)$.

— $\Xi(b) = (\overline{\chi}_b, \gamma_b, \psi)$.

— The nodes at a and b in Ξ have the same arity.

— There exists $\gamma' \Uparrow \gamma_b$ such that $\overline{\chi}_a \in \mathbb{F}(\gamma')$.

We have $M_{|a}$, $M_{|b}$ of the same kind. Let α_a, α_b be the blueprints of $M_{|a}$, $M_{|b}$. Since $\gamma_b \sqsubseteq^{\max}_{|\Lambda(M, a \cdot b)|} \alpha_b$, we have $\gamma' \Uparrow \gamma_b \sqsubseteq_1 \alpha_b$. By Lemma 5.10, there exists α' such that $\gamma' \sqsubseteq_1 \alpha' \Uparrow \alpha_b$. By Lemma 5.11, we have $\overline{\chi}_a \in \mathbb{F}(\gamma') \subseteq \mathbb{F}(\alpha')$, so M is not compact. □

6. Finiteness of the set of compact ϕ-shadows

Our final step is to prove that for each formula ϕ, the set of all compact ϕ-shadows is a finite set effectively computable from ϕ.

In Definition 6.1, we will introduce a final binary relation $\Subset$ on blueprints. The key lemma of this section (Lemma 6.14) shows that whenever $\mathscr{S} \subset \mathfrak{S}$ is a finite set (in particular, when $\mathscr{S}$ is the set of all subformulas of ϕ and all @'s tagged with a subformula of ϕ), the relation $\Subset$ is an almost full relation (Bezem *et al.* 2003) on the set of all $\mathscr{S}$-blueprints: for every infinite sequence $\gamma_1, \gamma_2, \ldots$ over $\mathbb{B}(\mathscr{S})$, there exists i, j such that $i < j$ and $\gamma_i \Subset \gamma_j$. We will prove this result with the help of Melliès' Axiomatic Kruskal Theorem (Melliès 1998). The finiteness of the set of compact ϕ-shadows follows from this key lemma with the help of König's Lemma (Lemma 6.15). The ability to compute these shadows follows directly from their definition.

By Lemma 5.13, another consequence of this result is the finiteness for each ϕ of the set of all compact Λ_{NF}-inhabitants of ϕ, although our decision method is based on the computation of *shadows* of compact terms rather than a direct computation of those terms. It is worth mentioning that the proof of Theorem 6.13 is non-constructive and that it gives no information about the complexity of our proof-search method – this question might itself become another open problem.

6.1. *Almost full relations and Higman's Theorem*

Definition 6.1. We let $\Subset$ be the relation on blueprints defined by $\alpha \Subset \beta$ if and only if for all $\overline{\chi} \in \mathbb{F}(\alpha)$, there exists $\gamma \Uparrow \beta$ such that $\overline{\chi} \in \mathbb{F}(\gamma)$.

Definition 6.2. Let $\mathscr{U}$ be an arbitrary set. An *almost full relation (AFR) on* $\mathscr{U}$ is a binary relation $\ll$ such that for every infinite sequence $(u_i)_{i \in \mathbb{N}}$ over $\mathscr{U}$, there exist i, j such that $i < j$ and $u_i \ll u_j$.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

The main aim of Section 6 is to prove the final key lemma of the paper, from which we will easily infer the decidability of Λ_{NF}-inhabitation: for each finite $\mathscr{S} \subseteq \mathfrak{S}$, the relation $\Subset$ is an AFR on $\mathbb{B}(\mathscr{S})$.

Proposition 6.3.

(1) If $\ll$ and $\ll'$ are AFRs on $\mathscr{U}$, then $\ll \cap \ll'$ is an AFR on $\mathscr{U}$.

(2) Suppose $\ll_{\mathscr{U}}$ is an AFR on $\mathscr{U}$ and $\ll_{\mathscr{V}}$ is an AFR on $\mathscr{V}$. Let $\ll_{\mathscr{U}\times\mathscr{V}}$ be the relation defined by $(U, V) \ll_{\mathscr{U}\times\mathscr{V}} (U', V')$ if and only if $U \ll_{\mathscr{U}} U'$ and $V \ll_{\mathscr{V}} V'$. Then $\ll_{\mathscr{U}\times\mathscr{V}}$ is an AFR on $\mathscr{U} \times \mathscr{V}$.

Proof. Both results appear in Melliès (1998) in Step 4 of the proof of Theorem 1 (page 523), as a corollary of Lemma 4 (page 520). □

Definition 6.4. Let $\mathscr{U}$ be a set and $\ll$ be a binary relation. We write $\mathbb{S}(\mathscr{U})$ to denote the set of all finite sequences over $\mathscr{U}$. The relation $\ll_{\mathbb{S}}$ *induced by* $\ll$ *on* $\mathbb{S}(\mathscr{U})$ is defined by $(U_1, \ldots, U_n) \ll_{\mathbb{S}} (V_1, \ldots, V_m)$ if and only if there exists a strictly monotone function $\eta : \{1, \ldots, n\} \to \{1, \ldots, m\}$ such that $U_i \ll V_{\eta(i)}$ for each $i \in \{1, \ldots, n\}$.

Theorem 6.5 (Higman). If $\ll$ is an AFR on $\mathscr{U}$, then $\ll_{\mathbb{S}}$ is an AFR on $\mathbb{S}(\mathscr{U})$.

Proof. See Higman (1952), Kruskal (1972) and Melliès (1998). □

6.2. *From rooted to unrooted blueprints*

Melliès' Axiomatic Kruskal Theorem allows us to conclude that a relation is an AFR (a 'well binary relation' in Melliès (1998)) if it satisfies a set of five properties or 'axioms' (there were six axioms in the original version of the theorem – see Melliès' remarks at the end of his proof explaining why five axioms suffice). The details of those axioms will be given in Section 6.3.

Four of those five axioms are relatively easy to check, but the other one is more problematic. This rather technical section is entirely devoted to the proof of Lemma 6.11, which will ensure that this final axiom is satisfied. To do this, we want to prove the following proposition:

Let $\mathscr{S}$ be a finite subset of $\mathfrak{S}$ and $\mathscr{B}_\varepsilon$ be a subset of $\mathbb{B}_\varepsilon(\mathscr{S})$.
Let $\mathscr{B} = \{*_{\overline{a}}(\beta_1, \ldots, \beta_n) | \, \forall i \in [1, \ldots, n], \beta_i \in \mathscr{B}_\varepsilon\}$.
If $\Subset$ is an AFR on $\mathscr{B}_\varepsilon$, then $\Subset$ is an AFR on $\mathscr{B}$.

Recall that $\mathbb{B}_\varepsilon(\mathscr{S})$ stands for the set of all rooted $\mathscr{S}$-blueprints. We want to be able to extend the property that $\Subset$ is an AFR on a given set of rooted blueprints to the set of all blueprints that have those rooted blueprints at their minimal addresses.

Higman's Theorem suffices to show that $\Subset_{\mathbb{S}}$ (Definition 6.4) is an AFR on the set of finite sequences over $\mathscr{B}_\varepsilon$. However, if we consider an infinite sequence $(\beta_i)_{i\in\mathbb{N}}$ over $\mathscr{B}$ and transform each $\beta_i = *_{\overline{a}_i}(\beta^i_1, \ldots, \beta^i_{n_i})$ where $\beta^i_1, \ldots, \beta^i_{n_i} \in \mathscr{B}_\varepsilon$ into $\sigma(\beta_i) = (\beta^i_1, \ldots, \beta^i_{n_i})$, the theorem will only provide two integers i, j and a strictly monotone function η such that $i < j$ and $\beta^i_k \Subset \beta^j_{\eta(k)}$ for each $k \in \{1, \ldots, n_i\}$. This is sufficient to ensure that

$$\beta_i = *_{\overline{a}_i}\big(\beta^i_1, \ldots, \beta^i_{n_i}\big) \Subset *_{\overline{b}}\big(\beta^j_{\eta(1)}, \ldots, \beta^j_{\eta(n_i)}\big),$$

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

but not $\beta_i \Subset \beta_j$ in general.

To bypass this difficulty, we show how for each blueprint $\beta \in \mathbb{B}(\mathscr{S})$, we can extract from the set of all vertical compressions of β a complete set of 'followers' of β of minimal size (Lemma 6.7). This set $\{\alpha_1, \ldots, \alpha_p\}$ has the property that for each $\overline{\phi} \in \mathbb{F}(\beta)$, there exists at least one α_i such that $\mathbb{F}(\alpha_i)$ contains a *subsequence* of $\overline{\phi}$ – but not necessarily $\overline{\phi}$ itself. The relative depth of each α_i does not depend on the relative depth of β, but only on $\mathscr{S}$: it is at most $\Sigma_{i=1}^{1+|\mathscr{S}_@|} i$, where $\mathscr{S}_@$ is the set of all binary symbols in $\mathscr{S}$. The lemma is proved in four steps:

(1) We first observe that the set of all $\alpha \Uparrow \beta$ with relative depth at most $\Sigma_{i=1}^{1+|\mathscr{S}_@|} i$ is a complete set of followers. If we consider the set of all γ such that $\gamma \sqsubseteq_1^{\max} \alpha$ for at least one such α, we obtain a (possibly infinite) set closed under $\equiv$ and finite up to $\equiv$. We call it the set of $\mathscr{S}$*-residuals* of β.

(2) We prove that the set of $\mathscr{S}$-residuals of β is a complete set of followers of β in the same sense, that is, for each $\overline{\phi} \in \mathbb{F}(\beta)$, there exists an $\mathscr{S}$-residual γ of β such that $\mathbb{F}(\gamma)$ contains a subsequence of $\overline{\phi}$ (Lemma 6.9).

(3) We prove that if

$$\beta = *_{\overline{a}}(\beta_1, \ldots, \beta_n)$$
$$\beta' = *_{\overline{b}}(\beta'_1, \ldots, \beta'_n, \beta'_{n+1}, \ldots, \beta'_{n+k})$$

are such that $\beta_i \Subset \beta'_i$ for each $i \in [1, \ldots, n]$, and if, furthermore, β, β' have the same set of $\mathscr{S}$-residuals, then $\beta \Subset \beta'$ (Lemma 6.10).

(4) The final step is the proof of the lemma itself. The set of $\mathscr{S}$-residuals is finite up to $\equiv$ (Lemma 5.5), so there are only a finite number of possible values for the set of residuals of each $\mathscr{S}$-blueprint. As a consequence, it is always possible to extract from an infinite sequence over $\mathscr{B}$ an infinite sequence of blueprints with the same set of residuals. The conclusion then follows from the third step and Higman's Theorem.

Definition 6.6. For every $\mathscr{S} \subseteq \mathfrak{S}$, we write $\mathscr{S}_@$ to denote the set of all binary symbols in $\mathscr{S}$.

Lemma 6.7. Let $\mathscr{S}$ be a finite subset of $\mathfrak{S}$. For all $\beta \in \mathbb{B}(\mathscr{S})$ and all $\overline{\psi} \in \mathbb{F}(\beta)$, there exists α with relative depth at most $\Sigma_{i=1}^{1+|\mathscr{S}_@|} i$ such that $\alpha \Uparrow \beta$ and such that $\mathbb{F}(\alpha)$ contains a subsequence of $\overline{\psi}$.

Proof. We define an $\mathscr{S}$*-linearisation* to be any pair $(\gamma, \overline{\chi})$ such that $\gamma \in \mathbb{B}(\mathscr{S})$ and $\overline{\chi} \in \mathbb{F}(\gamma)$. We say a *starting address for* $(\gamma, \overline{\chi})$ is any address b for which there exist ϕ, γ' such that $\gamma \rhd_\phi^b \gamma'$ and $\overline{\chi} \in \circledcirc(\mathbb{F}(\gamma'), (\phi))$. We define the *path to b in γ* to be the maximal sequence $(b_1, \ldots, b_n, b_{n+1})$ over $\mathrm{dom}(\gamma)$ such that $b_1 < \cdots < b_n < b_{n+1} = b$.

Given an arbitrary $\mathscr{S}$-linearisation $(\beta, \overline{\psi})$, we use induction on $|\mathrm{dom}(\beta)|$ to prove the following properties simultaneously:

(1) There exists an $\mathscr{S}$-linearisation $(\gamma, \overline{\chi})$ such that:

(a) $\gamma \Uparrow \beta$ and $\overline{\chi}$ is a subsequence of $\overline{\psi}$.

(b) γ has relative depth at most $1 + \Sigma_{i=1}^{|\mathscr{S}_@|} i$.

(2) There exists an $\mathscr{S}$-linearisation $(\alpha, \overline{\phi})$ such that:

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

(a) $\alpha \Uparrow \beta$, $\overline{\phi}$ is a subsequence of $\overline{\psi}$, and if $\psi \neq \varepsilon$, then the last elements of $\overline{\phi}, \overline{\psi}$ are equal.

(b) For each starting address b for $(\alpha, \overline{\phi})$ and for $(b_1, \ldots, b_n, b_{n+1})$ equal to the path to b in α, the values $\alpha(b_1), \ldots, \alpha(b_n)$ are pairwise distinct.

(c) For all c incomparable with each starting address for $(\alpha, \overline{\phi})$, the relative depth of $(\alpha_{|c})$ is at most $1 + \Sigma_{i=1}^{|\mathscr{S}_@|} i$.

Note that the conjunction of (2b) and (2c) implies that every address d in α has relative depth at most

$$|S_@| + 1 + \Sigma_{i=1}^{|\mathscr{S}_@|} i = \Sigma_{i=1}^{1+|\mathscr{S}_@|} i.$$

Indeed, suppose d has maximal relative depth and not a starting address for $(\alpha, \overline{\phi})$. Then d must be incomparable with each starting address for $(\alpha, \overline{\phi})$. Let e be the shortest prefix of d in $\mathrm{dom}(\alpha)$ that is incomparable with each starting address for $(\alpha, \overline{\phi})$. The address e has relative depth at most $|\mathscr{S}_@|$ in α, since otherwise there would exist in $\mathrm{dom}(\alpha)$ an address $f < e$ with relative depth $|\mathscr{S}_@|$ and a starting adress for $(\alpha, \overline{\phi})$ of the form $f \cdot f'$, with relative depth strictly greater than $|\mathscr{S}_\alpha|$, which would give a contradiction. Moreover, the relative depth of d is the sum of the relative depth of e in α and the relative depth of $\alpha_{|e}$.

The case $\beta = \varnothing_{\mathbb{B}}$ is immediate, and if $\beta = *_{\overline{a}}(\beta_1, \ldots, \beta_n)$, $i \neq j$ and $\beta_i, \beta_j \neq \varnothing_{\mathbb{B}}$, the conclusion follows easily from the induction hypothesis. So we suppose $\beta = @_\psi(\beta_1, \beta_2)$.

(1) Let d be an address of maximal length in $\beta^{-1}(@_\psi)$, and let $\delta = @_\psi(\delta_1, \delta_2) = \beta_{|d}$. By assumption, ε is the only element of $\delta^{-1}(@_\psi)$. As $\overline{\psi} \in \mathbb{F}(\beta)$, there exist

$$\overline{\psi}_0 \in \mathbb{F}(\delta)$$
$$\overline{\psi}_1 \in \mathbb{F}(\delta_1)$$
$$\overline{\psi}_2 \in \mathbb{F}(\delta_2)$$

such that $\overline{\psi}_0$ is a subsequence $\overline{\psi}$ and $\overline{\psi}_0 \in \circledcirc(\{\overline{\psi}_1\}, \{\overline{\psi}_2\})$. By the induction hypothesis, there exists an $(\mathscr{S} - \{@_\psi\})$-linearisation $(\gamma_1, \overline{\chi}_1)$ satisfying conditions (1a) and (1b) with respect to $(\delta_1, \overline{\psi}_1)$, and an $(\mathscr{S} - \{@_\psi\})$-linearisation $(\gamma_2, \overline{\chi}_2)$ satisfying conditions (2a), (2b) and (2c) with respect to $(\delta_2, \overline{\psi}_2)$. Let $\gamma = @_\psi(\gamma_1, \gamma_2)$. We have $\gamma \Uparrow \delta$ and $\beta(\varepsilon) = \delta(\varepsilon) = \gamma(\varepsilon)$, so $\gamma \Uparrow \beta$. The blueprint γ_1 has relative depth at most

$$1 + \Sigma_{i=1}^{|\mathscr{S}_@|-1} i \leqslant \Sigma_{i=1}^{|\mathscr{S}_@|} i.$$

The blueprint γ_2 has relative depth at most

$$|\mathscr{S}_@| + \Sigma_{i=1}^{|\mathscr{S}_@|-1} = \Sigma_{i=1}^{|\mathscr{S}_@|} i.$$

Therefore, γ has relative depth at most

$$1 + \Sigma_{i=1}^{|\mathscr{S}_@|} i.$$

Now $\overline{\chi}_2$ is a subsequence of $\overline{\psi}_2$ with the same final element, so there exists in

$$\circledcirc(\{\overline{\chi}_1\}, \{\overline{\chi}_2\}) \subseteq \mathbb{F}(@_\psi(\gamma_1, \gamma_2))$$

a subsequence $\overline{\chi}$ of $\overline{\psi}_0$. Thus $(\gamma, \overline{\chi})$ satisfies (1a) and (1b) with respect to $(\beta, \overline{\psi})$.

(2) As $\overline{\psi} \in \mathbb{F}(\beta)$, there exist

$$\overline{\psi}_1 \in \mathbb{F}(\beta_1)$$
$$\overline{\psi}_2 \in \mathbb{F}(\beta_2)$$

such that

$$\overline{\psi} \in \circledcirc(\{\overline{\psi}_1\}, \{\overline{\psi}_2\}).$$

By the induction hypothesis, there exists an $\mathscr{S}$-linearisation $(\alpha_1, \overline{\phi}_1)$ satisfying conditions (1a) and (1b) with respect to $(\beta_1, \overline{\psi}_1)$, and an $\mathscr{S}$-linearisation $(\alpha_2, \overline{\phi}_2)$ satisfying conditions (2a), (2b) and (2c) with respect to $(\beta_2, \overline{\psi}_2)$.

Let $\alpha_0 = @_\psi(\alpha_1, \alpha_2)$. We have $\alpha_0 \Uparrow \beta$. The final elements of $\overline{\phi}_2$ and $\overline{\psi}_2$ are equal and

$$\circledcirc(\{\overline{\phi}_1\}, \{\overline{\phi}_2\}) \subseteq \mathbb{F}(\alpha_0).$$

Hence, there exists in $\mathbb{F}(\alpha_0)$ a subsequence $\overline{\phi}_0$ of $\overline{\psi}$ with the same final element as $\overline{\psi}$. Thus $(\alpha_0, \overline{\phi}_0)$ satisfies (2a).

For all c incomparable with each starting address for $(\alpha_0, \overline{\phi}_0)$, either $c = (1) \cdot c'$ and $c' \in \mathrm{dom}(\alpha_1)$, or $c = (2) \cdot c''$ and $c'' \in \mathrm{dom}(\alpha_2)$ is incomparable with each starting address in α_2. As a consequence, the choice of α_1, α_2 ensures that $(\alpha_0, \overline{\phi}_0)$ satisfies (2c). If $(\alpha_0, \overline{\phi}_0)$ satisfies (2b), we may take $(\alpha, \overline{\phi}) = (\alpha_0, \overline{\phi}_0)$. Otherwise, some starting address b for $(\alpha_0, \overline{\phi}_0)$ does not satisfy condition (2b). Let $(b_1, \ldots, b_n, b_{n+1})$ be the path to b in α. We have $b_1 = \varepsilon$, and for each $i > 0$, there exists d_i such that $b_i = (2) \cdot d_i$. The sequence $(d_2, \ldots, d_{n+1})$ is then a path to $d = d_{n+1}$ in α_2, and d is a starting address for $(\alpha_2, \overline{\phi}_2)$. The values $\alpha_2(d_2), \ldots, \alpha_2(d_n)$ are pairwise distinct, so there must exist $i > 1$ such that $\alpha(b_i) = @_\psi$. Since b_i is in the path to b, there exists in $\mathbb{F}(\alpha_{2|d_i})$ a subsequence $\overline{\phi}'_0$ of $\overline{\phi}_0$ with the same last element as $\overline{\phi}_0$. For $\alpha'_0 = \alpha_0[\varepsilon \leftarrow \alpha_{2|d_i}]$, we have $\alpha'_0 \Uparrow \beta$ and $\overline{\phi}'_0 \in \mathbb{F}(\alpha'_0)$ and the final elements of $\overline{\phi}'_0$, $\overline{\phi}_0$ and $\overline{\psi}$ are equal. By the induction hypothesis, there exists an $\mathscr{S}$-linearisation $(\alpha, \overline{\phi})$ satisfying (2a), (2b) and (2c) with respect to $(\alpha'_0, \overline{\phi}'_0)$. The pair $(\alpha, \overline{\phi})$ also satisfies these conditions with respect to $(\beta, \overline{\psi})$. □

Definition 6.8. Let $\mathscr{S}$ be a finite subset of $\mathfrak{S}$. For all $\beta \in \mathbb{B}(\mathscr{S})$ and all $\alpha \Uparrow \beta$ with relative depth at most $\Sigma_{i=1}^{1+|\mathscr{S}_@|} i$, we define an *$\mathscr{S}$-residual* of β to be any α_0 such that $\alpha_0 \sqsubseteq_1^{\max} \alpha$.

Note that the set of $\mathscr{S}$-residuals of β is $\{\varnothing_{\mathbb{B}}\}$ if $\beta = \varnothing_{\mathbb{B}}$. Otherwise, it is an infinite set – even if $\beta = \phi$, the set of residuals of β is the $\equiv$-equivalence class of β and contains all blueprints of the form $*_a(\phi)$ (recall that $\equiv$ is a subset of $\sqsubseteq_1$ – see Definition 5.3).

Lemma 6.9. Let $\mathscr{S}$ be a finite subset of $\mathfrak{S}$. For all $\beta \in \mathbb{B}(\mathscr{S})$ and all $\overline{\psi} \in \mathbb{F}(\beta)$, there exists an $\mathscr{S}$-residual α_0 of β such that $\mathbb{F}(\alpha_0)$ contains a subsequence of $\overline{\psi}$.

Proof.

(1) Let γ, δ be arbitrary blueprints. Suppose $\gamma \curvearrowleft_1 \delta$. We prove by induction on δ that for all $\overline{\phi} \in \mathbb{F}(\delta)$, there exists in $\mathbb{F}(\gamma)$ a subsequence of $\overline{\phi}$. In order to deal with the case $\delta = @_\phi(\delta_1, \delta_2)$, we need to prove a slightly more precise property: for all $\overline{\phi} \in \mathbb{F}(\delta)$, there exists in $\mathbb{F}(\gamma)$ a subsequence $\overline{\psi}$ of $\overline{\phi}$ such that the final elements of $\overline{\phi}$ and $\overline{\psi}$

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

are equal. The base case is $\delta = *_{(a_1,a_2)}(\gamma_1, \gamma_2)$, $\gamma_1 \equiv \gamma_2$ and $\gamma = *_{a_1}(\gamma_1)$, and this case is clear. Other cases follow easily from the induction hypothesis.

(2) We can now prove the lemma. By Lemma 6.7 and the definition of an $\mathscr{S}$-residual, there exist α_0 and α such that $\alpha_0 \sqsubseteq_1 \alpha \Uparrow \beta$, $\mathbb{F}(\alpha)$ contains a subsequence of $\overline{\psi}$ and α_0 is an $\mathscr{S}$-residual. It then follows from (1) that $\mathbb{F}(\alpha_0)$ contains a subsequence of $\overline{\psi}$. □

Lemma 6.10. Let $\mathscr{S}$ be a finite subset of $\mathfrak{S}$. Suppose:

— $\beta = *_{\overline{a}}(\beta_1, \ldots, \beta_n) \in \mathbb{B}(\mathscr{S})$.
— $\beta' = *_{\overline{b}}(\beta'_1, \ldots, \beta'_n, \beta'_{n+1}, \ldots, \beta'_{n+k}) \in \mathbb{B}(\mathscr{S})$.
— $\beta_i \Subset \beta'_i$ for each $i \in \{1, \ldots, n\}$.
— The sets of $\mathscr{S}$-residuals of β and β' are equal.

Then $\beta \Subset \beta'$.

Proof. Let $\overline{\psi} \in \mathbb{F}(\beta)$. So there exists for each $i \in [1, \ldots, n]$ a sequence $\overline{\psi}_i \in \mathbb{F}(\beta_i)$ such that $\overline{\psi} \in \circledast(\{\overline{\psi}_1\}, \ldots, \{\overline{\psi}_n\})$. By assumption, there exists for each $i \in [1, \ldots, n]$ an $\alpha_i \Uparrow \beta'_i$ such that $\overline{\psi}_i \in \mathbb{F}(\alpha_i)$. As a consequence, $\overline{\psi} \in \mathbb{F}(*(\alpha_1, \ldots, \alpha_n))$.

By Lemma 6.9, there exists an $\mathscr{S}$-residual α_0 of β such that $\mathbb{F}(\alpha_0)$ contains a subsequence $\overline{\phi}$ of $\overline{\psi}$. By assumption, α_0 is also an $\mathscr{S}$-residual of β', so there exist $\alpha'_1, \ldots, \alpha'_{n+k}$ and $\overline{b}$ such that

$$\alpha_0 \sqsubseteq_1 *_{\overline{b}}(\alpha'_1, \ldots, \alpha'_{n+k}) \Uparrow \beta'.$$

By Lemma 5.11, we have

$$\overline{\phi} \in \mathbb{F}(*_{\overline{b}}(\alpha'_1, \ldots, \alpha'_{n+k})).$$

Hence, for each $i \in [1, \ldots, n+k]$, there exists in $\mathbb{F}(\alpha'_i)$ a subsequence of $\overline{\phi}$, which is also a subsequence of $\overline{\psi}$. Let

$$\alpha = *_{\overline{b}}(\alpha_1, \ldots, \alpha_n, \alpha'_{n+1}, \ldots, \alpha'_{n+k}).$$

Then $\alpha \Uparrow \beta'$, $\overline{\psi} \in \mathbb{F}(*(\alpha_1, \ldots \alpha_n))$, and for each $j \in [1, \ldots, k]$, there exists in $\mathbb{F}(\alpha'_{n+j})$ a subsequence of $\overline{\psi}$. Hence, $\overline{\psi} \in \mathbb{F}(\alpha)$. □

Lemma 6.11. Let $\mathscr{S}$ be a finite subset of $\mathfrak{S}$ and $\mathscr{B}_\varepsilon$ be a subset of $\mathbb{B}_\varepsilon(\mathscr{S})$. Let

$$\mathscr{B} = \{*_{\overline{a}}(\beta_1, \ldots, \beta_n) \mid \forall i \in [1, \ldots, n], \beta_i \in \mathscr{B}_\varepsilon\}.$$

If $\Subset$ is an AFR on $\mathscr{B}_\varepsilon$, then $\Subset$ is an AFR on $\mathscr{B}$.

Proof. Let $\mathscr{R} = \mathbb{B}(\mathscr{S}, \Sigma_{i=1}^{1+|\mathscr{S}_@|} i, 1)$ (see Definition 5.4). For each $\beta \in \mathscr{B}$, let $\rho(\beta)$ be the set of $\mathscr{S}$-residuals of β. We have $\rho(\beta) \subseteq \mathscr{R}$. Moreover, $\rho(\beta)$ is closed under $\equiv$ (since $\equiv$ is a subset of $\sqsubseteq_1$ – see Definition 5.3), that is, $\rho(\beta)$ is a union of the elements of a subset of $\mathscr{R}/_\equiv$. By Lemma 5.5 (1), the latter is a finite set, so $\{\rho(\beta) \mid \beta \in \mathscr{B}\}$ is a finite set.

For each $\beta = *_{\overline{a}}(\beta_1, \ldots, \beta_n) \in \mathscr{B}$, where $\overline{a}$ is increasing with respect to the lexicographic ordering of addresses and $\beta_1, \ldots, \beta_n \in \mathscr{B}_\varepsilon$, let $\sigma(\beta) = (\beta_1, \ldots, \beta_n)$ – recall that we can take $\overline{a} = \varepsilon$, $n = 0$ if $\beta = \varnothing_{\mathbb{B}}$, and $\overline{a} = (\varepsilon)$, $n = 1$ if β is a rooted blueprint. Since $\{\rho(\beta) \mid \beta \in \mathscr{B}\}$ is a finite set, every infinite sequence over $\mathscr{B}$ contains an infinite subsequence of blueprints with the same set of $\mathscr{S}$-residuals. By assumption, $\Subset$ is an AFR on $\mathscr{B}_\varepsilon$, so, by Theorem 6.5, $\Subset_{\mathbf{S}}$ is an AFR on $\{\sigma(\beta) \mid \beta \in \mathscr{B}\}$.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

Thus, for every infinite sequence $(\beta_i)_{i\in\mathbb{N}}$ over $\mathscr{B}$, there exist i, j such that $i < j$, $\sigma(\beta_i) \Subset_{\mathbb{S}} \sigma(\beta_j)$, and β_i and β_j have the same set of residuals. For

$$\sigma(\beta_i) = \left(\beta_1^i, \ldots, \beta_n^i\right)$$
$$\sigma(\beta_j) = \left(\beta_1^j, \ldots, \beta_{n+k}^j\right),$$

there exists a subsequence $(\beta_{l_1}^i, \ldots, \beta_{l_n}^i)$ of $\sigma(\beta_j)$ such that

$$\beta_1^i \Subset \beta_{l_1}^i, \ldots, \beta_n^i \Subset \beta_{l_n}^i.$$

There also exist $l_{n+1}, \ldots, l_{n+k}$ and two sequences $\overline{a}$ and $\overline{b}$ such that

$$\beta_i = *_{\overline{a}}\left(\beta_1^i, \ldots, \beta_n^i\right)$$
$$\beta_j = *_{\overline{b}}\left(\beta_{l_1}^j, \ldots, \beta_{l_n}^j, \beta_{l_{n+1}}^j, \ldots, \beta_{l_{n+k}}^j\right).$$

By Lemma 6.10, we have $\beta_i \Subset \beta_j$. □

6.3. *Axiomatic Kruskal Theorem and main key lemma*

The following definition is borrowed from Melliès (1998).

Definition 6.12. An *abstract decomposition system* is an 8-tuple

$$(\mathscr{T}, \mathscr{L}, \mathscr{V}, \preceq_{\mathscr{T}}, \preceq_{\mathscr{L}}, \preceq_{\mathscr{V}}, \overset{\cdot}{\longrightarrow}, \vdash)$$

where:

— $\mathscr{T}$ is a set of *terms* noted $t, u, \ldots$ equipped with a binary relation $\preceq_{\mathscr{T}}$.
— $\mathscr{L}$ is a set of *labels* noted $f, g, \ldots$ equipped with a binary relation $\preceq_{\mathscr{L}}$.
— $\mathscr{V}$ is a set of *vectors* noted $T, U, \ldots$ equipped with a binary relation $\preceq_{\mathscr{V}}$.
— $\overset{\cdot}{\longrightarrow}$ is a relation on $\mathscr{T} \times \mathscr{L} \times \mathscr{V}$, for example, $t \overset{f}{\longrightarrow} T$.
— $\vdash$ is a relation on $\mathscr{V} \times \mathscr{T}$, for example, $T \vdash t$.

For each such system, we let $\rhd_{\mathscr{T}}$ be the binary relation on $\mathscr{T}$ defined by

$$t \rhd_{\mathscr{T}} u \iff \exists (f, T) \in \mathscr{L} \times \mathscr{V}, \quad t \overset{f}{\longrightarrow} T \vdash u.$$

An *elementary term* t is a term minimal with respect to $\rhd_{\mathscr{T}}$, that is, a term for which there exists no u such that $t \rhd_{\mathscr{T}} u$.

Theorem 6.13 (Melliès). Suppose $(\mathscr{T}, \mathscr{L}, \mathscr{V}, \preceq_{\mathscr{T}}, \preceq_{\mathscr{L}}, \preceq_{\mathscr{V}}, \overset{\cdot}{\longrightarrow}, \vdash)$ satisfies the following properties:

Axiom I: There is no infinite chain $t_1 \rhd_{\mathscr{T}} t_2 \rhd_{\mathscr{T}} \ldots$
Axiom II: The relation $\preceq_{\mathscr{T}}$ is an AFR on the set of elementary terms.
Axiom III: For all t, u, u', if $t \preceq_{\mathscr{T}} u'$ and $u \rhd_{\mathscr{T}} u'$, then $t \preceq_{\mathscr{T}} u$.
Axiom IV-bis: For all t, u, f, g, T, U, if $t \overset{f}{\longrightarrow} T$ and $u \overset{g}{\longrightarrow} U$ and $f \preceq_{\mathscr{L}} g$ and $T \preceq_{\mathscr{V}} U$, then $t \preceq_{\mathscr{T}} u$.
Axiom V: For all $\mathscr{W} \subseteq \mathscr{V}$, for $\mathscr{W}_{\vdash} = \{t \in \mathscr{T} \mid \exists T \in \mathscr{W}, T \vdash t\}$, if $\preceq_{\mathscr{T}}$ is an AFR on $\mathscr{W}_{\vdash}$, then $\preceq_{\mathscr{V}}$ is an AFR on $\mathscr{W}$.

If, furthermore, $\preceq_{\mathscr{L}}$ is an AFR on $\mathscr{L}$, then $\preceq_{\mathscr{T}}$ is an AFR on $\mathscr{T}$.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

Proof. See Melliès (1998)[†]. □

Lemma 6.14. For each finite $\mathscr{S} \subseteq \mathfrak{S}$, the relation $\Subset$ is an AFR on $\mathbb{B}(\mathscr{S})$.

Proof. According to Lemma 6.11, it is sufficient to prove that $\Subset$ is an AFR on $\mathbb{B}_\varepsilon(\mathscr{S})$. Let $(\mathscr{T}, \mathscr{L}, \mathscr{V}, \preceq_\mathscr{T}, \preceq_\mathscr{L}, \preceq_\mathscr{V}, \stackrel{\cdot}{\longrightarrow}, \vdash)$ be the abstract decomposition system defined as follows:

— The set $\mathscr{T}$ is $\mathbb{B}_\varepsilon(\mathscr{S})$ – we let $\alpha \preceq_\mathscr{T} \beta$ if and only if there exists an address c such that $\alpha \Subset (\beta_{|c})$ and $\alpha(\varepsilon) = (\beta_{|c})(\varepsilon)$.

— The set $\mathscr{L}$ is the set of all @'s in $\mathscr{S}$ – the relation $\preceq_\mathscr{L}$ is the identity relation on this set.

— The set $\mathscr{V}$ is $\mathbb{B}(\mathscr{S}) \times \mathbb{B}(\mathscr{S})$ – the relation $\preceq_\mathscr{V}$ is defined by $(\alpha_1, \alpha_2) \preceq_\mathscr{V} (\beta_1, \beta_2)$ if and only if $\alpha_1 \Subset \beta_1$ and $\alpha_2 \Subset \beta_2$.

— The relation $\stackrel{\cdot}{\longrightarrow}$ is defined by $\alpha \stackrel{@_\phi}{\longrightarrow} (\beta_1, \beta_2)$ if and only if $\alpha = @_\phi(\beta_1, \beta_2)$.

— The relation $\vdash$ is the least relation satisfying the following condition: if $V = (\alpha_1, \alpha_2)$, $i \in \{1, 2\}$, $\beta_1, \ldots, \beta_n \in \mathbb{B}_\varepsilon(\mathscr{S})$ and $\alpha_i = *_{\overline{a}}(\beta_1, \ldots, \beta_n)$, then $V \vdash \beta_j$ for each $j \in [1, \ldots, n]$.

Note that the elements of $\mathscr{V}$ are pairs of blueprints that may be rootless. However, if $V \vdash \beta$, the blueprint β is always a rooted blueprint, so the relation $\vdash$ is indeed a subset of $\mathscr{V} \times \mathscr{T}$.

(A) For all $\mathscr{T}' \subseteq \mathscr{T}$, the relation $\Subset$ is an AFR on $\mathscr{T}'$ if and only if $\preceq_\mathscr{T}$ is an AFR on $\mathscr{T}'$. Indeed, consider an arbitrary infinite sequence $\overline{\alpha}$ over $\mathscr{T}'$. This sequence contains an infinite subsequence $(\alpha)_{i \in \mathbb{N}}$ such that all $\alpha_i(\varepsilon)$ are equal. Clearly, $\alpha_i \Subset \alpha_j$ implies $\alpha_i \preceq_\mathscr{T} \alpha_j$. Conversely, if $\alpha_i \preceq_\mathscr{T} \alpha_j$, there exists c such that $\alpha_i \Subset \alpha_{j_{|c}}$ and $\alpha_i(\varepsilon) = \alpha_j(\varepsilon) = \alpha_j(c)$. So $\alpha_i \Subset \alpha_{j_{|c}} \Uparrow \alpha_j$, hence $\alpha_i \Subset \alpha_j$.

(B) We now check that all axioms of Theorem 6.13 are satisfied:

— It is clear that Axiom I is satisfied.

— The set of elementary terms is the set of all blueprints consisting of single formulas of $\mathscr{S}$. The relation $\preceq_\mathscr{T}$ is, of course, an AFR on the set of elementary terms, that is, Axiom II is satisfied.

— Axiom III is immediate.

— If $(\alpha_1, \alpha_2) \preceq_\mathscr{V} (\beta_1, \beta_2)$, then $\alpha_1 \Subset \beta_1$ and $\alpha_2 \Subset \beta_2$, hence

$$@_\psi(\alpha_1, \alpha_2) \Subset @_\psi(\beta_1, \beta_2),$$

a fortiori

$$@_\psi(\alpha_1, \alpha_2) \preceq_\mathscr{T} @_\psi(\beta_1, \beta_2),$$

so Axiom IV-bis is satisfied.

— To prove that Axiom V is satisfied, we let $\mathscr{W} \subseteq \mathscr{V}$. Then, by definition,

$$\mathscr{W}_\vdash = \{\beta \in \mathscr{T} \mid \exists(\alpha_1, \alpha_2) \in \mathscr{W}, (\alpha_1, \alpha_2) \vdash \beta\}.$$

[†] Mellies' result was actually established for a different list of axioms (numbered from I to VI), but Mellies mentions the possibility of dropping Axiom VI and replacing Axiom IV with Axiom IV-bis in a remark following his proof of the main theorem.

https://doi.org/10.1017/S0960129512000412 Published online by Cambridge University Press

Assuming $\preceq_{\mathscr{T}}$ is an AFR on $\mathscr{W}_{\vdash}$, we will prove that $\preceq_{\mathscr{V}}$ is an AFR on $\mathscr{W}$. By part (A), the relation $\Subset$ is an AFR on $\mathscr{W}_{\vdash} \subseteq \mathbb{B}_{\varepsilon}(\mathscr{S})$. Let

$$\mathscr{B} = \{*_{\overline{a}}(\beta_1, \dots, \beta_n) | \, \forall i \in [1, \dots, n], \beta_i \in \mathscr{W}_{\vdash}\}.$$

By Lemma 6.11, the relation $\Subset$ is an AFR on $\mathscr{B}$. Moreover, $\mathscr{W} \subseteq \mathscr{B} \times \mathscr{B}$. By Proposition 6.3 (2), the relation $\preceq_{\mathscr{V}}$ is an AFR on $\mathscr{B} \times \mathscr{B}$, and is thus an AFR on $\mathscr{W}$. □

Lemma 6.15. For each formula ϕ, the set of all compact ϕ-shadows is a finite set effectively computable from ϕ.

Proof. For each compact ϕ-shadow Ξ and each address a such that a is a leaf in Ξ, we define a *step-continuation at a* of Ξ to be any compact ϕ-shadow Ξ' such that

$$\mathrm{dom}(\Xi') \subsetneq \mathrm{dom}(\Xi) \cup \{a \cdot (1), a \cdot (2)\}$$

and Ξ and Ξ' take the same values on dom(Ξ). Let $\rightsquigarrow$ be the relation defined by $\Xi \rightsquigarrow \Xi'$ if and only if Ξ' is a step continuation of Ξ. By Lemma 5.5 and the fact that the set of subformulas of ϕ is a finite set, for all Ξ, the set of all Ξ' such that $\Xi \rightsquigarrow \Xi'$ is a finite set effectively computable from Ξ. Let $\mathscr{C}$ be the closure under $\rightsquigarrow$ of $\{(\varepsilon \mapsto (\varepsilon, \varnothing_{\mathbb{B}}, \phi))\}$ The set of all compact ϕ-shadows is clearly equal to this set, so it is enough to prove that $\mathscr{C}$ is a finite set. In order to show a contradiction, we assume that $\mathscr{C}$ is infinite. By König's Lemma, there exists an infinite sequence $\Xi_0 \rightsquigarrow \Xi_1 \rightsquigarrow \dots$ over $\mathscr{C}$. The union $\Xi_\infty = \cup_{i \geqslant 0} \Xi_i$ is a tree of infinite domain. By König's Lemma again, there exists an infinite chain of addresses $a_1 < a_2 < \dots$ such that all a_i are nodes of Ξ_∞ with the same arity and labelled with the same subformula of ϕ. If $i < j$ and a_i and a_j are labelled with $(\overline{\chi}_i, \gamma_i, \psi)$ and $(\overline{\chi}_i, \gamma_j, \psi)$, we cannot have $\gamma_i \Subset \gamma_j$, since otherwise there would exist a k such that Ξ_k is not compact. A contradiction then follows from Lemma 6.14. □

7. From the shadows to the light

Theorem 7.1. Ticket Entailment is decidable.

Proof. The following propositions are equivalent:

— The formula ϕ is provable in the logic $T_{\rightarrow}$.
— The formula ϕ is inhabited by a combinator within the basis BB′IW.
— The formula ϕ is Λ_{NF}-inhabited (Lemma 2.10).
— There exists a compact Λ_{NF}-inhabitant of ϕ (Lemma 4.9).
— There exists a compact ϕ-shadow with the same tree domain as a Λ_{NF}-inhabitant of ϕ (Lemmas 4.9 and 5.13).

By Lemma 6.15, the set of compact ϕ-shadows is effectively computable from ϕ. By the subformula property (Lemma 2.5), for each shadow Ξ in this set, up to the choice of bound variables, there are only a finite number of Λ_{NF}-inhabitant of ϕ with the same domain as Ξ. Moreover, this set of inhabitants is clearly computable from Ξ and ϕ. Hence the existence of a Λ_{NF}-inhabitant of ϕ is decidable. □

Acknowledgments

This work could not have been achieved without countless helpful comments and invaluable support from Paweł Urzyczyn, Paul-André Melliès and Pierre-Louis Curien. I am also deeply indebted to the anonymous referees for their remarkably careful reading of the paper.

References

Anderson, A. R. (1960) Entailment shorn of modality. *Journal of Symbolic Logic* **25** (4) 388.

Anderson, A. R. and Belnap Jr, N. D. (1975) *Entailment: The Logic of Relevance and Necessity*, Volume 1, Princeton University Press.

Anderson, A. R., Belnap Jr, N. D. and Dunn, J. M. (1990) *Entailment: The Logic of Relevance and Necessity*, Volume 2, Princeton University Press.

Barendregt, H. (1984) *The Lambda Calculus: Its Syntax and Semantics* (Revised edition), Studies in Logic and the Foundations of Mathematics **103**, North Holland.

Barwise, J. (ed.) (1977) *Handbook of Mathematical Logic*, Studies in Logic and Foundations of Mathematics, North-Holland.

Bezem, M., Klop, J. W. and de Vrijer, R. (eds.)[†] (2003) Term Rewriting Systems. *Cambridge Tracts in Theoretical Computer Science* **55**, Cambridge University Press.

Bimbó, K. (2005) Types of I-free hereditary right maximal terms. *Journal of Philosophical Logic* **34** (5-6) 607–620.

Broda, S., Damas, L., Finger, M. and Silva e Silva, P. S. (2004) The decidability of a fragment of $BB'IW$-logic. *Theoretical Computer Science* **318** (3) 373–408.

Bunder, M. W. (1996) Lambda Terms Definable as Combinators. *Theoretical Computer Science* **169** (1) 3–21.

Higman, G. (1952) Ordering by divisibility in abstract algebra. *Proceedings of the London Mathematical Society* **3** (2) 326–336.

Kripke, S. (1959) The problem of entailment. *Journal of Symbolic Logic* **24** (4) 324.

Krivine, J.-L. (1993) *Lambda-calculus, types and models*, Masson.

Kruskal, J. B. (1972) The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory, Series A* **13** (3) 297–305.

Melliès, P.-A. (1998) On a duality between Kruskal and Dershowitz theorems. In: Larsen, K. G, Skyum, S. and Winskel, G. (eds.) ICALP. *Springer-Verlag Lecture Notes in Computer Science* **1443** 518–529.

Trigg, P., Hindley, J. R. and Bunder, M. W. (1994) Combinatory abstraction using B, B' and friends. *Theoretical Computer Science* **135** (2) 405–422.

Urquhart, A (1984) The undecidability of entailment and relevant implication. *Journal of Symbolic Logic* **49** (4) 1059–1073.

[†] Also known as 'Terese'.