

# DISCRIMINATING LAMBDA-TERMS USING CLOCKED BÖHM TREES

JÖRG ENDRULLIS, DIMITRI HENDRIKS, JAN WILLEM KLOP, AND ANDREW POLONSKY

VU University Amsterdam, Dept. of Computer Science, De Boelelaan 1081A, 1081 HV Amsterdam  
{j.endrullis,r.d.a.hendriks,j.w.klop,a.polonsky}@vu.nl

**ABSTRACT.** As observed by Intrigila [16], there are hardly techniques available in the  $\lambda$ -calculus to prove that two  $\lambda$ -terms are not  $\beta$ -convertible. Techniques employing the usual Böhm Trees are inadequate when we deal with terms having the same Böhm Tree (BT). This is the case in particular for fixed point combinators, as they all have the same BT. Another interesting equation, whose consideration was suggested by Scott [24], is  $\text{BY} = \text{BYS}$ , an equation valid in the classical model  $\mathcal{P}\omega$  of  $\lambda$ -calculus, and hence valid with respect to BT-equality  $=_{\text{BT}}$ , but nevertheless the terms are  $\beta$ -inconvertible.

To prove such  $\beta$ -inconvertibilities, we employ ‘clocked’ BT’s, with annotations that convey information of the tempo in which the data in the BT are produced. Böhm Trees are thus enriched with an intrinsic clock behaviour, leading to a refined discrimination method for  $\lambda$ -terms. The corresponding equality is strictly intermediate between  $=_{\beta}$  and  $=_{\text{BT}}$ , the equality in the model  $\mathcal{P}\omega$ . An analogous approach pertains to Lévy–Longo and Berarducci Trees.

Our refined Böhm Trees find in particular an application in  $\beta$ -discriminating fixed point combinators (fpc’s). It turns out that Scott’s equation  $\text{BY} = \text{BYS}$  is the key to unlocking a plethora of fpc’s, generated by a variety of production schemes of which the simplest was found by Böhm, stating that new fpc’s are obtained by postfixing the term  $\text{Sl}$ , also known as Smullyan’s Owl. We prove that all these newly generated fpc’s are indeed new, by considering their clocked BT’s. Even so, not all pairs of new fpc’s can be discriminated this way. For that purpose we increase the discrimination power by a precision of the clock notion that we call ‘atomic clock’.

## 1. INTRODUCTION

Böhm Trees constitute a well-known method to discriminate  $\lambda$ -terms  $M, N$ : if  $\text{BT}(M)$  and  $\text{BT}(N)$  are not identical, then  $M$  and  $N$  are  $\beta$ -inconvertible,  $M \neq_{\beta} N$ . But how do we prove  $\beta$ -inconvertibility of  $\lambda$ -terms with the same Böhm Tree? This question was raised in Scott [24] for the interesting equation  $\text{BY} = \text{BYS}$  between terms that as Scott noted are presumably  $\beta$ -inconvertible, yet BT-equal ( $=_{\text{BT}}$ ). Scott used his Induction Rule to prove

*1998 ACM Subject Classification:* F.4.1 [Mathematical logic and formal languages]: Lambda calculus and related systems.

*Key words and phrases:*  $\lambda$ -calculus, Böhm Trees,  $\beta$ -inconvertibility, fixed point combinators.

This is a modified and extended version of [13] which appeared in the proceedings of LICS 2010. The research has been partially funded by the Netherlands Organisation for Scientific Research (NWO) under grant numbers 612.000.934, 639.021.020, and 612.001.002.

that  $BY = BYS$ ; instead we will employ below the infinitary  $\lambda$ -calculus with the same effect, but with more convenience for calculations as a direct generalization of finitary  $\lambda$ -calculus. Often one can solve such a  $\beta$ -discrimination problem by finding a suitable invariant for all the  $\beta$ -reducts of  $M, N$ . Below we will do this by way of preparatory example for the fixed point combinators (fpc's) in the Böhm sequence. But a systematic method for this discrimination problem has been lacking, and such a method is one of the two contributions of this paper.

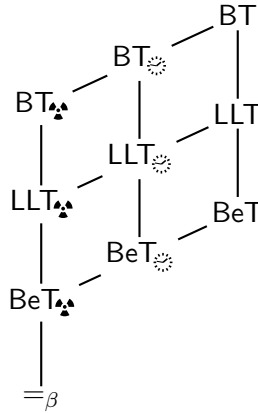


Figure 1: Comparison of (atomic) clock semantics and unclocked semantics. Higher means more identifications.

Actually, the need for such a strategic method was forced upon us, by the other contribution, because Scott's equation  $BY = BYS$  turned out to be the key unlocking a plethora of new fpc's. The new generation schemes are of the form: if  $Y$  is an fpc, then  $YP_1 \dots P_n$  is an fpc, abbreviated as  $Y \Rightarrow YP_1 \dots P_n$ . So  $\square P_1 \dots P_n$  is an 'fpc-generating' vector, and can be considered as a building block to make new fpc's. But are they indeed new? A well-known example of a (singleton)-fpc-generating vector is  $\square \delta$ , where  $\delta = SI$ , giving rise when starting from Curry's fpc to the Böhm sequence of fpc's. Here another interesting equation is turning up, namely  $Y = Y\delta$ , for an arbitrary fpc  $Y$ , considered by Statman and Intrigila [16]. In fact, it is implied by Scott's equation, for an arbitrary fpc  $Y$ :

$$BY = BYS \implies BYI = BYSI \iff Y = Y\delta$$

The first equation  $BY = BYS$  will yield many new fpc's, built in a modular way; the last equation  $Y = Y\delta$  addresses the question whether they are indeed new. Finding ad hoc invariant proofs for their novelty is too cumbersome. But fortunately, it turns out that although the new fpc's all have the same BT, namely  $\lambda f.f^\omega$ , they differ in the way this BT is formed, in the 'tempo of formation', where the ticks of the clock are head reduction steps. We can thus discern a clock-like behaviour of BT's, and we refine BT's to 'clocked' BT's by annotating them with this information. These then enable us to discriminate the terms in question.

This refined discrimination method works best for what we call 'simple' terms (or, for terms that reduce to simple terms). A term is 'simple' if its reduction to the Böhm Tree does not duplicate redexes. The class of simple terms is still fairly extensive; it includes all fpc's that are constructed in the modular way that we present, thereby solving our novelty problem. In fact, we gain some more ground: though our discrimination method works

best for pairs of simple terms, it can also fruitfully be applied to compare a simple term with a non-simple term, and with some more effort, we can even compare and discriminate non-simple terms; see Section 5 for an example.

Even so, many pairs of fpc's cannot yet be discriminated, because they not only have the same BT, they also have the same clocked BT. Therefore, in a final grading up of the precision, we introduce 'atomic clocks', where the actual position of a head reduction step is administrated. All this pertains not only to the BT-semantics, but also to Lévy–Longo Trees (LLT) (or lazy trees), and Berarducci Trees (BeT) (or syntactic trees). Many problems stay open, in particular problems generalizing the equation of Statman and Intrigila, when arbitrary fpc's are considered.

**Overview.** After defining preliminary notions in Section 2, in Section 3 we are concerned with constructing new fpc's from old, by some generating schemes. In Section 4 we define clocked Böhm Trees. The main results here are Theorems 26 and 31. The first states that if no reduct of  $M$  has a clock that is at least as fast as the clock of  $N$ , then  $M$  and  $N$  are inconvertible, whereas Theorem 31 states that if  $M$  is a simple term then it suffices that the clock  $M$  is not eventually faster than the clock of  $N$ . In the paper we are mainly concerned with  $\lambda$ -terms which have simple reducts. An exception is Section 5, where we answer a question of Plotkin (see Related Work below) which involves arbitrary (not necessarily simple) fpc's. Another elaborate example is given in Section 6, where we compute the clocks of three enumerators for Combinatory Logic. In Section 7, we give a refinement of the clock method by not only recording the number of head reduction steps but also their positions. As an application we show that every combination of the fixed point generating vectors  $\square(SS)S^{\sim n}$  introduced in Section 3 give rise to *new* fpc's. We briefly mention how the theory can be extended to the other well-known semantics of  $\lambda$ -calculus, namely Lévy–Longo and Berarducci Trees, in Section 8. We conclude in Section 9 with directions for future research.

**Related Work.** The present paper is an extension and elaboration of [13]. In particular, as an example application of the main theorems (Theorem 26), we now answer the following question of Plotkin [22]:

*Is there a fixed point combinator  $Y$  such that  $Y(\lambda z.fzz) =_{\beta} Y(\lambda x.Y(\lambda y.fxy))$ ?*

An idea similar to clocked Böhm Trees is employed in the excellent paper [2], pointed out to us by Tarmo Uustalu at the 2010 LICS conference in Edinburgh. In [2], Aehlig and Joachimski study continuous normalization of the coinductive  $\lambda$ -calculus (see [19, 17]) extended with a void 'wait' constructor  $\mathcal{R}$ . This extra constructor is returned whenever the head constructor of a term cannot immediately be read off from the argument. If the head constructor of an application  $rs$  is to be found, it depends on a recursive call to investigate whether  $r$  is a variable, an abstraction or an application again; in this case an  $\mathcal{R}$  is returned. Thus it is guaranteed that the procedure of building a non-wellfounded term over the extended grammar is productive by making the recursion guarded. Every  $\mathcal{R}$  is matched either by a  $\beta$ -step necessary to reach the normal form (as represented by the Böhm Tree) or by an application node in the Böhm Tree. When  $r = \lambda x.r'$  then  $\mathcal{R}$  matches a  $\beta$ -step and so  $s$  is used in the substitution  $r'[x := s]$ . Or we find that  $r$  is a variable and so  $\mathcal{R}$  matches an application node in the Böhm Tree. So the number of  $\mathcal{R}$ 's in the normal form of a term  $t$  is precisely related to the number of reduction steps and the

‘size’ of the resulting Böhm Tree. Summarizing, similar to the clocked Böhm Trees that we introduce in Section 4, in [2] Böhm Trees are enriched with information about how the tree is constructed. However, in [2] the refinement is used for proving their normalization function to be continuous, and not for discriminating  $\lambda$ -terms, the goal we pursue here.

In [15] a heuristic procedure in finitary  $\lambda$ -calculus is given to construct fpc’s in a uniform way.

## 2. PRELIMINARIES

To make this paper moderately self-contained, and to fix notations, we lay out some ingredients. For  $\lambda$ -calculus we refer to [3] and [7]. For an introduction to Böhm, Berarducci and Lévy–Longo Trees, we refer to [3, 1, 8, 5].

**Definition 1.** Let  $\mathcal{X}$  be a countably infinite set of variables. The set  $Ter(\lambda)$  of *finite  $\lambda$ -terms* is defined inductively by the following grammar:

$$M ::= x \mid \lambda x.M \mid M \cdot M \quad (x \in \mathcal{X})$$

We use  $x, y, z, \dots$  for variables, and  $M, N, \dots$  to range over the elements of  $Ter(\lambda)$ .

Thus  $\lambda$ -terms are variables, abstractions or applications. Usually we suppress the application symbol in a term  $M \cdot N$  and just write  $(MN)$ . We also adopt the usual conventions for omitting brackets, i.e., we let application associate to the left, so that  $N_1 N_2 \dots N_k$  denotes  $(\dots (N_1 N_2) \dots N_k)$ , and we let abstraction associate to the right:  $\lambda x_1 \dots x_n.M$  stands for  $(\lambda x_1.(\dots (\lambda x_n.(M))))$ .

**Definition 2.** The set  $Ter^\infty(\lambda)$  of (finite and) *infinite  $\lambda$ -terms* is defined by interpreting the grammar from Definition 1 *coinductively*, that is,  $Ter^\infty(\lambda)$  is the largest set  $X$  such that every element  $M \in X$  is either a variable  $x$ , an abstraction  $\lambda x.M'$  or an application  $M_1 M_2$  with  $M', M_1, M_2 \in X$ . (See further [23] for a precise treatment of coinductive definition and proof principles.)

Böhm, Lévy–Longo and Berarducci Trees are infinite  $\lambda$ -terms with an additional clause for  $\perp$  in the grammar, where  $\perp$  stands for the different notions of ‘undefined’ in these semantics, see Definitions 21, 54 and 55.

**Definition 3.** The relation  $\rightarrow_\beta$  on  $Ter(\lambda)$  or  $Ter^\infty(\lambda)$ , called  *$\beta$ -reduction*, is the compatible closure (i.e., closure under term formation) of the  $\beta$ -rule:

$$(\lambda x.M)N \rightarrow M[x := N] \quad (\beta)$$

where  $M[x := N]$  denotes the result of substituting  $N$  for all free occurrences of  $x$  in  $M$ . Furthermore, we write  $\rightarrow_\beta^n$  for the  $n$ -fold composition of  $\rightarrow_\beta$ , defined by  $M \rightarrow_\beta^0 M$ , and  $M \rightarrow_\beta^{n+1} N$  if  $M \rightarrow_\beta P$  and  $P \rightarrow_\beta^n N$  for some  $P \in Ter(\lambda)$ . We use  $\rightarrow_\beta$  to denote the reflexive-transitive closure of  $\rightarrow_\beta$ ,  $M \twoheadrightarrow_\beta N$  iff  $\exists n. M \rightarrow_\beta^n N$ ; we let  $\rightarrow_{\bar{\beta}}$  denote the reflexive closure of  $\rightarrow_\beta$ ,  $\rightarrow_{\bar{\beta}} = \rightarrow_\beta^0 \cup \rightarrow_\beta^1$ . We write  $M =_\beta N$  to denote that  $M$  is  $\beta$ -convertible with  $N$ , i.e.,  $=_\beta$  is the equivalence closure of  $\rightarrow_\beta$ . For syntactic equality (modulo renaming of bound variables), we use  $\equiv$ .

Apart from the ‘many-steps’ relation  $\rightarrow_\beta$ , we introduce ‘multi-steps’  $\twoheadrightarrow_\beta$  [30, 3], which arise from complete developments.

**Definition 4.** We write  $\multimap_{\beta}$  for *multi-steps*, that is, complete developments of a set of redex occurrences in a term. A *development* of a set of redex occurrences  $U$  in a term contracts exclusively descendants (residuals) of  $U$ , and it is called *complete* if there are no residuals of  $U$  left.

A complete development of  $U$  can alternatively be viewed as contracting all redex occurrences in  $U$  in an inside-out fashion.

We will often omit the subscript  $\beta$  in  $\rightarrow_{\beta}$ ,  $\twoheadrightarrow_{\beta}$  and  $\multimap_{\beta}$ .

**Definition 5.** A  $\lambda$ -term  $M$  is called a *normal form* if there exists no  $N$  with  $M \rightarrow N$ . We say that a term  $M$  *has a normal form* if it reduces to one. For  $\lambda$ -terms  $M$  having a normal form we write  $\underline{M}$  for the unique normal form  $N$  with  $M \twoheadrightarrow N$  (uniqueness follows from confluence of the  $\lambda$ -calculus).

Some commonly used combinators are:

$$\mathbb{I} = \lambda x.x \quad \mathbb{K} = \lambda xy.x \quad \mathbb{S} = \lambda xyz.xz(yz) \quad \mathbb{B} = \lambda xyz.x(yz)$$

**Definition 6.** A *position* is a sequence over  $\{0, 1, 2\}$ . For a finite or infinite  $\lambda$ -term  $M$ , the *subterm*  $M|_p$  of  $M$  at *position*  $p$  is defined by:

$$\begin{aligned} M|_{\epsilon} &= M & (MN)|_{1p} &= M|_p \\ (\lambda x.M)|_{0p} &= M|_p & (MN)|_{2p} &= N|_p \end{aligned}$$

$\text{Pos}(M)$  is the set of positions  $p$  such that  $M|_p$  is defined.

**Definition 7.**

- (i) A term  $Y$  is an *fpc* if  $Yx =_{\beta} x(Yx)$ .
- (ii) An fpc  $Y$  is *k-reducing* if  $Yx \rightarrow^k x(Yx)$ .
- (iii) An fpc  $Y$  is *reducing* if  $Y$  is  $k$ -reducing for some  $k \in \mathbb{N}$ .
- (iv) A term  $Z$  is a *weak fpc (wfpc)* if  $Zx =_{\beta} x(Z'x)$  where  $Z'$  is again a wfpc.

The definition of weak fpc's in (iv) is essentially coinductive [23], that is, implicitly employing a 'largest set' semantics. In long form, the definition means the following: the set of weak fpc's is the largest set  $W \subseteq \text{Ter}(\lambda)$  such that for every  $Z \in W$  we have  $Zx =_{\beta} x(Z'x)$  for some  $Z' \in W$ .

A wfpc is alternatively defined as a term having the same Böhm Tree as an fpc, namely  $\lambda x.x^{\omega} \equiv \lambda x.x(x(x(\dots)))$ . Weak fpc's are known in foundational studies of type systems as *looping combinators*; see, e.g., [10] and [14].

**Example 8.** Define by double recursion,  $Z$  and  $Z'$  such that  $Zx = x(Z'x)$  and  $Z'x = x(Zx)$ . Then  $Z, Z'$  are both wfpc's, and  $Zx = x(x(Zx))$ . So  $Z$  delivers its output twice as fast as an ordinary fpc, but the generator flipflops.

As to 'double recursion', [20] collects several proofs of the double fixed point theorem, including some in [3, 28].

**Definition 9.**

- (i) A *head reduction step*  $\rightarrow_h$  is a  $\beta$ -reduction step of the form:  
 $\lambda x_1 \dots x_n. (\lambda y.M) N N_1 \dots N_m \rightarrow \lambda x_1 \dots x_n. (M[y := N]) N_1 \dots N_m$  with  $n, m \geq 0$ .
- (ii) Accordingly, a *head normal form (hnf)* is a  $\lambda$ -term of the form:  
 $\lambda x_1 \dots \lambda x_n. y N_1 \dots N_m$  with  $n, m \geq 0$ .

- (iii) A *weak head normal form (whnf)* is an hnf or an abstraction, that is, a whnf is a term of the form  $xM_1 \dots M_m$  or  $\lambda x.M$ .
- (iv) A term *has a (weak) hnf* if it reduces to one.
- (v) We call a term *root-stable* if it does not reduce to a redex:  $(\lambda x.M)N$ . A term is called *root-active* if it does not reduce to a root-stable term.

**Infinitary  $\lambda$ -calculus  $\lambda^\infty\beta$ .** We will only use the infinitary  $\lambda$ -calculus  $\lambda^\infty\beta$  for some simple calculations such as  $(\lambda ab.(ab)^\omega)l =_{\lambda^\infty\beta} \lambda b.(lb)^\omega =_{\lambda^\infty\beta} \lambda b.b^\omega$ . Here  $M^\omega$  denotes the infinite  $\lambda$ -term  $M(M(M(\dots)))$  obtained as the solution of  $M^\omega = MM^\omega$ . For a proper setup of  $\lambda^\infty\beta$  we refer to [6, 19, 18, 5]. In a nutshell,  $\lambda^\infty\beta$  extends finitary  $\lambda$ -calculus by admitting infinite  $\lambda$ -terms, the set of which is called  $Ter^\infty(\lambda)$ , and infinite reductions (in [18, 5] possibly transfinitely long, in [6] of length  $\leq \omega$ ). Limits of infinite reduction sequences are obtained by a strengthening of Cauchy-convergence, stipulating that the depth of contracted redexes must tend to infinity. The  $\lambda^\infty\beta$ -calculus is not infinitary confluent ( $CR^\infty$ ), but still has unique infinite normal forms ( $UN^\infty$ ). Böhm Trees (BT's) without  $\perp$  are infinite normal forms in  $\lambda^\infty\beta$ . But beware, the reverse does not hold, e.g.  $\lambda x.(\lambda x.(\lambda x.\dots))$  is an infinite normal form, but not a BT; it is in fact an LLT (Lévy–Longo Tree, and also a BeT (Berarducci Tree). The notions BT, LLT, BeT are defined e.g. in [5], and in [8]. These notions are also defined in Sections 4 and 8, via their clocked versions.

**Definition 10.** For terms  $A, B$  we define  $AB^{\sim n}$  and  $A^n B$ :

$$\begin{aligned} AB^{\sim 0} &= A & A^0 B &= B \\ AB^{\sim n+1} &= ABB^{\sim n} & A^{n+1} B &= A(A^n B) \end{aligned}$$

A context of the form  $\square B^{\sim n}$  is called a *vector*. For the vector notation, it is to be understood that term formation gets highest priority, i.e.,  $ABC^{\sim n} = (AB)C^{\sim n}$ .

### 3. FIXED POINT COMBINATORS

*The theory of sage birds (technically called fixed point combinators) is a fascinating and basic part of combinatory logic; we have only scratched the surface.*

R. Smullyan [28].

**3.1. The Böhm Sequence.** There are several ways to make fpc's. For heuristics behind the construction of Curry's fpc  $Y_0 = \lambda f.\omega_f\omega_f$ , with  $\omega_f = \lambda x.f(xx)$ , and Turing's fpc  $Y_1 = \eta\eta$  with  $\eta = \lambda x.f.f(xxf)$ , see [3, 20].

It is well-known, as observed by C. Böhm [9, 3], that the class of fpc's coincides exactly with the class of fixed points of the peculiar term  $\delta = \lambda ab.b(ab)$ , convertible with SI. The notation  $\delta$  is convenient for calculations and stems from [16]. This term also attracted the attention of R. Smullyan, in his beautiful fable about fpc's figuring as birds in an enchanted forest: "An extremely interesting bird is the owl  $O$  defined by the following condition:  $Oxy = y(xy)$ ." [28]. We will return to the Owl in Remark 14 below.

Thus the term  $Y\delta$  is an fpc whenever  $Y$  is. It follows that starting with  $Y_0$ , Curry's fpc, we have an infinite sequence of fpc's  $Y_0, Y_0\delta, Y_0\delta\delta, \dots, Y_0\delta^{\sim n}, \dots$ ; we call this sequence the 'Böhm sequence'. Note that  $Y_0\delta =_\beta \eta\eta$ , justifying the overloaded notation  $Y_1$ . Now

the question is whether all these ‘derived’ fpc’s are really new, in other words, whether the sequence is free of duplicates. This is \*Exercise 6.8.9 in [3].

Note that we could also have started the sequence from another fpc than Curry’s. Now for the sequence starting from an *arbitrary* fpc  $Y$ , it is actually an open problem whether that sequence of fpc’s  $Y, Y\delta, Y\delta\delta, \dots, Y\delta^{\sim n}, \dots$  is free of repetitions. All we know, applying Intrigila’s theorem, Theorem 13 below, is that no two consecutive fpc’s in this sequence are convertible. But let us first consider the Böhm sequence.

**Definition 11.** The *Böhm sequence* is the sequence  $(Y_n)_{n \geq 0}$  where  $Y_n$  is defined by

$$Y_0 = \lambda f. \omega_f \omega_f \qquad Y_n = \eta \eta \delta^{\sim(n-1)} \quad (n > 0)$$

We show that the Böhm sequence contains no duplicates by determining the set of reducts of every  $Y_n$ . For  $Y_3$ <sup>1</sup> the head reduction is displayed in Figure 2, but this is by no means the whole reduction graph. For future reference we note that the head reduction diagram suggests a ‘clock behaviour’.

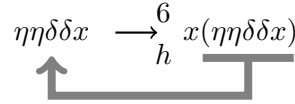


Figure 2: Head reduction of  $Y_3 x$ .

**Theorem 12.** The Böhm sequence  $(Y_n)_{n \geq 0}$  contains no duplicates.

*Proof.* We define languages  $\mathcal{L}_n \subseteq \text{Ter}(\lambda)$  as follows:

$$\begin{aligned} \mathcal{L}_0 &::= \lambda f. f^k(\omega_f \omega_f) & (k \geq 0) \\ \mathcal{L}_1 &::= \eta \eta \mid \lambda f. f^k(\mathcal{L}_1 f) & (k > 0) \\ \mathcal{L}_n &::= \mathcal{L}_{n-1} \delta \mid \lambda b. b^k(\mathcal{L}_n b) \mid \delta \mathcal{L}_n & (n > 1, k > 0) \end{aligned}$$

We show that:

- (i)  $Y_n \in \mathcal{L}_n$ ;
- (ii)  $\mathcal{L}_n$  is closed under  $\beta$ -reduction; and
- (iii)  $\mathcal{L}_n$  and  $\mathcal{L}_m$  are disjoint, for  $n \neq m$ .

Then it follows that  $\mathcal{L}_n$  contains the set of  $\rightarrow_\beta$ -reducts of  $Y_n$ , and using (iii) this implies that  $Y_n \neq_\beta Y_m$  for all  $n \neq m$ .

For (i) note that  $Y_0 \in \mathcal{L}_0$ , and  $Y_n = \eta \eta \delta^{\sim(n-1)}$  which is in  $\mathcal{L}_n$  by induction on  $n > 0$ .

We show (ii): if  $M \in \mathcal{L}_n$  and  $M \rightarrow N$ , then  $N \in \mathcal{L}_n$ . Using induction, we do not need to consider cases where the rewrite step is inside a variable of the grammar. We write  $\mathcal{L}_n$  in terms as shorthand for a term  $M \in \mathcal{L}_n$ .

- ( $\mathcal{L}_0$ ) We have  $\lambda f. f^k(\omega_f \omega_f) \rightarrow \lambda f. f^{k+1}(\omega_f \omega_f) \in \mathcal{L}_0$ .
- ( $\mathcal{L}_1$ ) We have  $\eta \eta \rightarrow \lambda f. f(\eta \eta f) \in \mathcal{L}_1$ ,  
and  $\lambda f. f^k(\lambda f. f^\ell(\mathcal{L}_1 f) f) \rightarrow \lambda f. f^{k+\ell}(\mathcal{L}_1 f) \in \mathcal{L}_1$ .

<sup>1</sup>Actually Figure 2 displays  $Y_3 x$ . We will frequently consider  $Yx$  instead of  $Y$  as then the repetition is immediate, and because we have that if  $Mx \neq_\beta M'x$  then also  $M \neq_\beta M'$ .

- ( $\mathcal{L}_n$ ) Case 1:  $\lambda f.f^k(\mathcal{L}_1 f)\delta \rightarrow \delta^k(\mathcal{L}_1 \delta) \in \mathcal{L}_n$  for  $n = 2$ , and  $(\lambda b.b^k(\mathcal{L}_{n-1} b))\delta \rightarrow \delta^k(\mathcal{L}_{n-1} \delta) \in \mathcal{L}_n$  for  $n > 2$ .  
 Case 2:  $\lambda b.b^k(\lambda c.c^\ell(\mathcal{L}_n c)b) \rightarrow \lambda b.b^{k+\ell}(\mathcal{L}_n b) \in \mathcal{L}_n$ .  
 Case 3:  $\delta \mathcal{L}_n \rightarrow \lambda b.b(\mathcal{L}_n b) \in \mathcal{L}_n$ .

For  $n \neq m$ ,  $n > 1$ , (iii) follows by counting the number of passive  $\delta$ 's, that is, the number of occurrences of the form  $P\delta$  for some  $P$ . To see that  $\mathcal{L}_0 \cap \mathcal{L}_1 = \emptyset$ , note that if  $M \in \mathcal{L}_1$  is an abstraction, then  $M \equiv \lambda f.f^k(Pf)$  containing a subterm  $Pf$  which is never the case in  $\mathcal{L}_0$ .  $\square$

A very interesting theorem involving  $\delta$  was proved by B. Intrigila, affirming a conjecture by R. Statman.

**Theorem 13** (Intrigila [16]). *There is no ‘double’ fixed point combinator. That is, for no fpc  $Y$  we have  $Y\delta =_\beta Y$ .*

**Remark 14** (Smullyan’s Owl  $\text{Sl} =_\beta \delta = \lambda ab.b(ab)$ ).

We collect some salient facts and questions.

- (i) If  $Z$  is a wfpc, both  $\delta Z$  and  $Z\delta$  are wfpc’s [28].
- (ii) Call an applicative combination of  $\delta$ ’s a  $\delta$ -term, that is, a term solely built from  $\delta$  and application. In spite of  $\delta$ ’s simplicity, not all  $\delta$ -terms are strongly normalizing (SN). An example of a  $\delta$ -term with infinite reduction is  $\delta\delta(\delta\delta)$  (Johannes Waldman, Hans Zantema, personal communication, 2007).
- (iii) Let  $M$  be a non-trivial  $\delta$ -term, i.e., not a single  $\delta$ . Then  $M$  is SN iff  $M$  contains exactly one occurrence of  $\delta\delta$ . Furthermore, if  $\delta$ -terms  $M, M'$  are SN, then they are convertible if and only if  $M, M'$  have the same length [20]. Here the length of a  $\delta$ -term is the number  $\delta$ ’s in the term.
- (iv) Convertibility is decidable for  $\delta$ -terms [29].
- (v) We define  $\Delta = \delta^\omega$ , so  $\Delta \equiv \delta\Delta$ . Then, the infinite  $\lambda$ -term  $\Delta$  is an fpc:  $\Delta x \equiv \delta\Delta x \rightarrow x(\Delta x)$ . The term  $\Delta$  can be normalized:  $\Delta \rightarrow_\omega \lambda f.f^\omega$ . There are many more infinitary fpc’s, e.g., for every  $n$ , the infinite term  $(\text{SS})^\omega \text{S}^{\sim n} \text{I}$  is one, as will be clear from the sequel.
- (vi) The term  $\delta\delta(\delta\delta)$  has no hnf, and hence its Böhm Tree is trivial,  $\text{BT}(\delta\delta(\delta\delta)) \equiv \perp$ . However, its Berarducci Tree is not trivial. Zantema remarked that  $\delta$ -terms, even infinite ones, such as  $\Delta\Delta$ , are “top-terminating” (Zantema considered the applicative rule for  $\delta$  only — we expect that his observation remains valid for the  $\lambda\beta$ -version).
- (vii) Is Intrigila’s theorem also valid for wfpc’s: for no wfpc  $Z$  we have  $Z\delta =_\beta Z$ ?

**3.2. The Scott Sequence.** In [24, p. 360] the equation  $\text{BY}_0 = \text{BY}_0\text{S}$  is mentioned as an interesting example of an equation not provable in  $\lambda\beta^2$ , while easily provable with Scott’s Induction Rule. Scott mentions that he expects that using ‘methods of Böhm’ the non-convertibility in  $\lambda\beta$  can be established, but that he did not attempt a proof. On the other hand, with the induction rule the equality is easily established. We will not consider Scott’s Induction Rule, but we will be working in the infinitary lambda calculus,  $\lambda^\infty\beta$ . It is readily verified that in  $\lambda^\infty\beta$  we have:

$$\text{BY}_0 =_{\lambda^\infty\beta} \text{BY}_0\text{S} =_{\lambda^\infty\beta} \lambda ab.(ab)^\omega$$

<sup>2</sup>This equation is also discussed in [11].



**Proposition 15.**  $\text{BY}_0 \neq_{\beta} \text{BY}_0\text{S}$ 

*Proof.* Postfixing the combinator  $\mathsf{l}$  yields  $\text{BY}_0\mathsf{l}$  and  $\text{BY}_0\text{S}\mathsf{l}$ . Now  $\text{BY}_0\mathsf{l} =_{\beta} \text{Y}_0$  and  $\text{BY}_0\text{S}\mathsf{l} =_{\beta} \text{Y}_0(\text{S}\mathsf{l}) = \text{Y}_1$ . Because  $\text{Y}_0 \neq_{\beta} \text{Y}_1$  (Theorem 12), the result follows.  $\square$

In the same way we can strengthen this non-equation to all fpc's  $Y$ , using Theorem 13.

**Remark 16.**

- (i) The idea of postfixing an  $\mathsf{l}$  is suggested by the Böhm Tree  $\lambda ab.(ab)^{\omega}$  of  $\text{BY}$  and  $\text{BYS}$ . Namely, in  $\lambda^{\infty}\beta$  we calculate:  $(\lambda ab.(ab)^{\omega})\mathsf{l} = \lambda b.(\mathsf{l}b)^{\omega} = \lambda b.b^{\omega}$  which is the Böhm Tree of any fpc.
- (ii) Interestingly, Scott's equation  $\text{BY} = \text{BYS}$  implies the equation of Statman and Intrigila,  $Y = Y\delta$  as one readily verifies, as in the proof of Proposition 15.

Actually, the comparison between the terms  $\text{BY}$  and  $\text{BYS}$  has more in store for us than just providing an example that the extension from finitary lambda calculus  $\lambda\beta$  to infinitary lambda calculus  $\lambda^{\infty}\beta$  is not conservative. The BT-equality of  $\text{BY}$  and  $\text{BYS}$  suggests looking at the whole sequence  $\text{BY}, \text{BYS}, \text{BYSS}, \dots, \text{BYS}^{\sim n}, \dots$ . All these terms have the Böhm Tree  $\lambda ab.(ab)^{\omega}$ , and hence they are not fpc's. But postfixing an  $\mathsf{l}$  turns them into fpc's.

**Definition 17.** The *Scott sequence* is defined by:

$$\text{BY}_0\mathsf{l}, \text{BY}_0\text{S}\mathsf{l}, \text{BY}_0\text{SS}\mathsf{l}, \dots, \text{BY}_0\text{S}^{\sim n}\mathsf{l}, \dots$$

We write  $\text{U}_n = \text{BY}_0\text{S}^{\sim n}\mathsf{l}$  for the  $n$ -th term in this sequence.

The Scott sequence concurs with the Böhm sequence of fpc's only for the first two elements, and then splits off with different fpc's. But there is a second surprise. In showing that  $\text{U}_n$  is an fpc, we find as a bonus the fpc-generating vector  $\square(\text{SS})\text{S}^{\sim n}\mathsf{l}$  (which does preserve the property of fpc's to be reducing).

**Theorem 18.** *Let  $Y$  be a  $k$ -reducing fpc and  $n \geq 0$ . Then:*

- (i)  $\text{BYS}^{\sim n}\mathsf{l}$  is a (non-reducing) fpc;
- (ii)  $Y(\text{SS})\text{S}^{\sim n}\mathsf{l}$  is a  $(k + 3n + 7)$ -reducing fpc.

The proof of Theorem 18 is easy: see the next example.

**Example 19.** Let  $Y$  be a  $k$ -reducing fpc. Then:

$$\begin{aligned} \text{BYSSS}\mathsf{l}x &\rightarrow_h Y(\text{SS})\mathsf{l}x \xrightarrow{^k_h} \text{SS}(Y(\text{SS}))\mathsf{l}x \\ &\rightarrow_h^3 \text{SS}(Y(\text{SS})\text{S})\mathsf{l}x \rightarrow_h^3 \text{S}\mathsf{l}(Y(\text{SS})\mathsf{l})x \\ &\rightarrow_h^3 \mathsf{l}x(Y(\text{SS})\mathsf{l}x) \rightarrow_h^1 x(Y(\text{SS})\mathsf{l}x) \end{aligned}$$

This shows that  $\text{BYS}^{\sim 3}\mathsf{l}$  is a non-reducing fpc, and at the same time that  $Y(\text{SS})\mathsf{l}x$  is reducing.

**3.3. Generalized Generation Schemes.** The schemes mentioned in Theorem 18 for generating new fixed points from old, are by no means the only ones. There are in fact infinitely many of such schemes. They can be obtained analogously to the ones that we extracted above from the equation  $\text{BY} = \text{BYS} = \lambda ab.(ab)^{\omega}$ , or the equation  $Mab = ab(Mab)$ . We only treat the case for  $n = 3$ : consider the equation  $Nabc = abc(Nabc)$ . Then every solution  $N$  is again a 'pre-fpc', namely  $N\mathsf{l}\mathsf{l}$  is an fpc:  $N\mathsf{l}\mathsf{l}x =_{\beta} \mathsf{l}x(N\mathsf{l}\mathsf{l}x) =_{\beta} x(N\mathsf{l}\mathsf{l}x)$ .

- (i)  $Nabc = Y(abc)$ , which yields  $N = (\lambda yabc.y(abc))Y = (\lambda yabc.BBBYabc)Y$ . We obtain  $N = BBBY$ .
- (ii)  $N = Y\xi$  with  $\xi = \lambda nabc.abc(nabc)$ , yielding the fpc-generating vector  $\square\xi\mathbb{I}$ .
- (iii)  $Nabc = abc(Nabc) = S(ab)(Nab)c$ . So we take  $Nab = S(ab)(Nab)$ , which yields  $Nab = Y(S(ab)) = BBBY(BS)ab$ . So  $N = BBBY(BS)$ , and thus we find the equation  $BBBY = BBBY(BS)$ , in analogy with the equation  $BY = BYS$  above.

Also this equation spawns lots of fpc's as well as fpc-generating vectors. Let's abbreviate BS by A. First one forms the sequence

$$BBBY, BBBYA, BBBYAA, BBBYAAA, \dots$$

These terms all have the BT  $\lambda abc.abc(abc)^\omega$ . They are not yet fpc's, but only 'pre-fpc's'. But after postfixing this time  $\dots\mathbb{I}$  we do again obtain a sequence of fpc's:

$$BBBY\mathbb{I}, BBBY\mathbb{A}\mathbb{I}, BBBY\mathbb{A}\mathbb{A}\mathbb{I}, \dots$$

Again the first two coincide with  $Y_0, Y_1$ , but the series deviates not only from the Böhm sequence but also from the Scott sequence above. As above, the proof that a term in this sequence is indeed an fpc, yields an fpc-generating vector. In this way we find as a new fpc-generating scheme

$$Y \Rightarrow Y(\mathbb{A}\mathbb{A}\mathbb{A})\mathbb{A}^{\sim n}\mathbb{I}$$

We can derive many more of these schemes by proceeding with solving the general equation  $Na_1a_2\dots a_n = a_1a_2\dots a_n(Na_1a_2\dots a_n)$ , bearing in mind the following proposition.

**Proposition 20.** *If  $N$  is a term satisfying  $Na_1a_2\dots a_n = a_1a_2\dots a_n(Na_1a_2\dots a_n)$ , then  $N\mathbb{I}^{\sim(n-1)}$  is an fpc.*

We finally mention an fpc-generating scheme with 'dummy parameters':

$$Y \Rightarrow YQP_1\dots P_n$$

where  $P_1, \dots, P_n$  are arbitrary (dummy) terms, and  $Q = \lambda yp_1\dots p_nx.x(yp_1\dots p_nx)$ .

#### 4. CLOCK BEHAVIOUR OF LAMBDA TERMS

As we have seen, there is vast space of fpc's and there are many ways to derive new fpc's. The question is whether all these fpc's are indeed new. So we have to prove that they are not  $\beta$ -convertible.

For the Böhm sequence we did this by an ad hoc argument based on a syntactic invariant; and this method works fine to establish lots of non-equations between the alleged 'new' fpc's that we constructed above. Still, the question remains whether there are not more 'strategic' ways of proving such inequalities.

In this section we propose a more strategic way to discriminate terms with respect to  $\beta$ -conversion. The idea is to extract from a  $\lambda$ -term more than just its BT, but also how the BT was formed; one could say, in what tempo, or in what rhythm. A BT is formed from static pieces of information, but these are rendered in a clock-wise fashion, where the ticks of the internal clock are head reduction steps.

In the sequel we write  $[k]M$  for the term  $M$  where the root is *annotated with*  $k \in \mathbb{N}$ . Here, term formation binds stronger than annotation  $[k]$ . For example  $[k]MN$  stands for the term  $[k](MN)$  (that is, annotating the (non-displayed) application symbol in-between  $M$  and  $N$ , in contrast to  $([k]M)N$ ). Moreover, for an annotated term  $M$  we use  $\lfloor M \rfloor$  to

denote the term obtained from  $M$  by dropping all annotations (including annotations of subterms).

**Definition 21** (Clocked Böhm Trees). The *clocked Böhm Tree*  $\text{BT}_{\odot}(M)$  of a  $\lambda$ -term  $M$  is an annotated (potentially infinite) term coinductively defined as follows. If  $M$  has no hnf, then define  $\text{BT}_{\odot}(M)$  as  $\perp$ . Otherwise, there is a head reduction  $M \rightarrow_h^k \lambda x_1 \dots \lambda x_n. y M_1 \dots M_m$  to hnf. Then we define  $\text{BT}_{\odot}(M)$  as the term  $[k]\lambda x_1 \dots \lambda x_n. y \text{BT}_{\odot}(M_1) \dots \text{BT}_{\odot}(M_m)$ .

The notions of subterms and positions (see Definition 6) carry over to annotated terms, in particular clocked Böhm Trees, in a straightforward way. The (non-clocked) Böhm Tree of a  $\lambda$ -term  $M$  can be obtained by dropping the annotations:  $\text{BT}(M) = \lfloor \text{BT}_{\odot}(M) \rfloor$ .

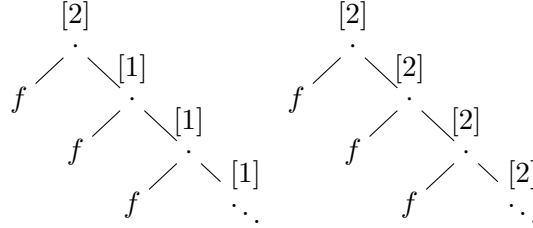


Figure 3: Clocked Böhm Trees of  $Y_0 f$  and  $Y_1 f$ .

Let us consider the fpc's  $Y_0$  of Curry and  $Y_1$  of Turing. We have  $Y_0 \equiv \lambda f. \omega_f \omega_f$  where  $\omega_f \equiv \lambda x. f(xx)$ , and

$$\omega_f \omega_f \rightarrow_h^1 f(\omega_f \omega_f)$$

Therefore we obtain  $\text{BT}_{\odot}(Y_0 f) = [2]f\text{BT}_{\odot}(\omega_f \omega_f)$ , and  $\text{BT}_{\odot}(\omega_f \omega_f) = [1]f\text{BT}_{\odot}(\omega_f \omega_f)$ .

For  $Y_1 \equiv \eta \eta$  where  $\eta \equiv \lambda x. \lambda f. f(xxf)$  we get:

$$Y_1 f \equiv \eta \eta f \rightarrow_h^2 f(\eta \eta f)$$

Hence,  $\text{BT}_{\odot}(Y_1 f) = [2]f\text{BT}_{\odot}(Y_1 f)$ . Figure 3 displays the clocked Böhm Trees of  $Y_0 f$  (left) and  $Y_1 f$  (right).

The following definition captures the well-known Böhm equality of  $\lambda$ -terms.

**Definition 22.**  $\lambda$ -terms  $M$  and  $N$  are *BT-equal*, denoted by  $M =_{\text{BT}} N$ , if  $\text{BT}(M) \equiv \text{BT}(N)$ .

If  $M$  and  $N$  are not BT-equal, then  $M \neq_{\beta} N$ . Consequently, if for some  $\lambda$ -term  $F$ , we have  $\text{BT}(MF) \neq \text{BT}(NF)$ , then  $M \neq_{\beta} N$ .

Below, we refine this approach by comparing the clocked Böhm Trees  $\text{BT}_{\odot}(M)$  and  $\text{BT}_{\odot}(N)$  instead of the ordinary (non-clocked) Böhm Trees. In general,  $\text{BT}_{\odot}(M) \neq \text{BT}_{\odot}(N)$  does not always imply that  $M \neq_{\beta} N$ . Nevertheless, for a large class of  $\lambda$ -terms, called ‘simple’ below, this implication will turn out to be true.

In the following definition, we lift relations over natural numbers to relations over clocked Böhm Trees.

**Definition 23.** Let  $T_1$  and  $T_2$  be clocked Böhm Trees, and  $R \subseteq \mathbb{N} \times \mathbb{N}$ . We define the following notations:

- (i) For  $p \in \text{Pos}(T_1) \cap \text{Pos}(T_2)$  we let  $T_1 R_p T_2$  denote that either both  $T_1|_p$  and  $T_2|_p$  are not annotated, or both are annotated and then  $T_1|_p \equiv [k_1]T'_1$  and  $T_2|_p \equiv [k_2]T'_2$  with  $k_1 R k_2$ .

- (ii) We write  $T_1 R T_2$  if  $[T_1] \equiv [T_2]$  and  $T_1 R_p T_2$  for every  $p \in \mathcal{P}os(T_1)$ .
- (iii) We write  $T_1 R_{\exists} T_2$ , and say that  $R$  holds eventually, if  $[T_1] \equiv [T_2]$  and there exists a depth level  $\ell \in \mathbb{N}$  such that  $T_1 R_p T_2$  for all positions  $p \in \mathcal{P}os(T_1)$  with  $|p| \geq \ell$ .

**Definition 24.** For  $\lambda$ -terms  $M$  and  $N$  we say:

- (i)  $M$  improves  $N$  globally if  $\mathbf{BT}_{\circlearrowleft}(M) \leq \mathbf{BT}_{\circlearrowleft}(N)$ ;
- (ii)  $M$  improves  $N$  eventually if  $\mathbf{BT}_{\circlearrowleft}(M) \leq_{\exists} \mathbf{BT}_{\circlearrowleft}(N)$ ;
- (iii)  $M$  matches  $N$  eventually if  $\mathbf{BT}_{\circlearrowleft}(M) =_{\exists} \mathbf{BT}_{\circlearrowleft}(N)$ .

The following proposition states that the ordering  $>_{\circlearrowleft}$  on  $\lambda$ -terms defined by  $M >_{\circlearrowleft} N$  if and only if  $\mathbf{BT}_{\circlearrowleft}(M) \geq \mathbf{BT}_{\circlearrowleft}(N)$  is a ‘semi-model’ of  $\beta$ -reduction [21]. We leave this for future research.

**Proposition 25.** *Clocks are accelerated under reduction, that is, if  $M \rightarrow N$ , then the reduct  $N$  improves  $M$  globally ( $\mathbf{BT}_{\circlearrowleft}(M) \geq \mathbf{BT}_{\circlearrowleft}(N)$ ). Dually, clocks slow down under expansion (the reverse of reduction).*

*Proof.* We proceed by an elementary diagram construction. Whenever we have co-initial steps  $M \rightarrow_h M_1$  and  $M \rightarrow_{\bar{h}} M_2$ , then by orthogonal projection [30] there exist joining steps  $M_1 \rightarrow_{\bar{h}} M'$  and  $M_2 \rightarrow_h M'$ . Note that the head step  $M \rightarrow_h M_1$  cannot be duplicated, only erased in case of an overlap. This leads to the elementary diagram displayed in Figure 4.

$$\begin{array}{ccc}
 M & \xrightarrow{\quad} & M_1 \\
 \downarrow & & \downarrow \\
 M_2 & \xrightarrow[\bar{h}]{} & M'
 \end{array}$$

Figure 4: Elementary diagram.

We have  $\rightarrow \subseteq \rightarrow^*$ . By induction on the length of the rewrite sequence  $\rightarrow^*$  it suffices to show that  $M \rightarrow N$  implies  $\mathbf{BT}_{\circlearrowleft}(M) \geq \mathbf{BT}_{\circlearrowleft}(N)$ . Let  $M \rightarrow N$ . If  $M$  has no hnf, then the same holds for  $N$ , and hence  $\mathbf{BT}_{\circlearrowleft}(M) = \perp = \mathbf{BT}_{\circlearrowleft}(N)$ . Assume that there exists a head rewrite sequence  $M \rightarrow_h^k H \equiv \lambda x_1 \dots \lambda x_n. y M_1 \dots M_m$  to hnf. We have  $\mathbf{BT}_{\circlearrowleft}(M) \equiv [k] \lambda x_1 \dots \lambda x_n. y \mathbf{BT}_{\circlearrowleft}(M_1) \dots \mathbf{BT}_{\circlearrowleft}(M_m)$ .

Using the elementary diagram above ( $k$  times), we can project  $M \rightarrow N$  over  $M \rightarrow_h^k H$ , and obtain  $H \rightarrow H'$ ,  $N \rightarrow_h^{\ell} H' \equiv \lambda x_1 \dots \lambda x_n. y M'_1 \dots M'_m$  with  $\ell \leq k$ . Then  $\mathbf{BT}_{\circlearrowleft}(N) \equiv [\ell] \lambda x_1 \dots \lambda x_n. y \mathbf{BT}_{\circlearrowleft}(M'_1) \dots \mathbf{BT}_{\circlearrowleft}(M'_m)$  and  $\ell \leq k$ . Since  $H \rightarrow H'$  and  $H$  is in hnf, we get  $M_i \rightarrow M'_i$  for every  $i = 1, \dots, m$ . We then can apply the same argument to  $M_i \rightarrow M'_i$  and by coinduction (or induction on the depth), we obtain  $\mathbf{BT}_{\circlearrowleft}(M) \geq \mathbf{BT}_{\circlearrowleft}(N)$ .  $\square$

While  $\mathbf{BT}_{\circlearrowleft}(M) \not\equiv \mathbf{BT}_{\circlearrowleft}(N)$  does not imply  $M \neq_{\beta} N$ , the following theorem allows us to use clocked Böhm Trees for discriminating  $\lambda$ -terms:

**Theorem 26.** *Let  $M$  and  $N$  be  $\lambda$ -terms. If  $N$  cannot be improved globally by any reduct of  $M$ , then  $M \neq_{\beta} N$ .*

*Proof.* If  $M =_{\beta} N$ , then  $M \rightarrow M' \leftarrow N$  for some  $M'$  by confluence. Hence  $\mathbf{BT}_{\circlearrowleft}(M') \leq \mathbf{BT}_{\circlearrowleft}(N)$  by Proposition 25.  $\square$

Note that for distinguishing  $M$  and  $N$  we can always consider  $\beta$ -equivalent terms  $M' =_{\beta} M$  and  $N' =_{\beta} N$  instead. For Theorem 26 we have to show  $\neg(\mathbf{BT}_{\circlearrowleft}(M') \leq \mathbf{BT}_{\circlearrowleft}(N))$  for all

reduces  $M'$  of  $M$ . This condition is in general difficult to prove. However, the theorem is of use if one of the terms has a manageable set of reducts, and this term happens to have slower clocks. A striking example will be given below in solving a question of Plotkin in Section 5.

For a large class of  $\lambda$ -terms it turns out that clocks are invariant under reduction. We call these terms ‘simple’.

**Definition 27.** A redex  $(\lambda x.M)N$  is called:

- (i) *linear* if  $x$  has at most one occurrence in  $M$ ;
- (ii) *call-by-value* if  $N$  is a normal form; and
- (iii) *simple* if it is linear or call-by-value.

The definition of simple redexes generalizes the well-known notions of call-by-value and linear redexes. Next, we define simple *terms*. Intuitively, we call a term  $M$  ‘simple’ if every reduction admitted by  $M$  only contracts simple redexes. The following definition further generalizes this intuition by considering only reductions computing the Böhm Tree.

**Definition 28** (Simple terms). A  $\lambda$ -term  $M$  is *simple* if either  $M$  has no hnf, or the head reduction to hnf  $M \rightarrow_h \lambda x_1 \dots \lambda x_n. y M_1 \dots M_m$  contracts only simple redexes, and  $M_1, \dots, M_m$  are simple terms.

Note that this definition is essentially coinductive: the set of simple terms is the largest set  $X$  such that if  $M \in X$  then either  $M$  has no hnf or the reduction to hnf  $M \rightarrow_h \lambda x_1 \dots \lambda x_n. y M_1 \dots M_m$  contracts only simple redexes and  $M_1, \dots, M_m \in X$  again.

All the fpc’s in this paper are either simple or have simple reducts. The clock of simple  $\lambda$ -terms is invariant under reduction, that is, reduction of a simple term affects only finitely many annotations in the clocked Böhm Tree. For example, by reducing a term we can always make the clock values in a finite prefix equal to 0.

**Proposition 29.** *Let  $N$  be a reduct of a simple term  $M$ . Then  $N$  matches  $M$  eventually ( $\text{BT}_{\circlearrowleft}^{\circ}(M) =_{\exists} \text{BT}_{\circlearrowleft}^{\circ}(N)$ ).*

*Proof.* The proof is a straightforward extension of the proof of Proposition 25 with the observation that for simple terms  $M$ , rewriting  $M \rightarrow_h^k H \equiv \lambda x_1 \dots \lambda x_n. y M_1 \dots M_m$  to hnf does not duplicate redexes. For simple terms  $M$ , the elementary diagrams are of the form displayed in Figure 5. We use  $s \xrightarrow{\emptyset} t$  to denote empty steps, that is,  $s \equiv t$ .

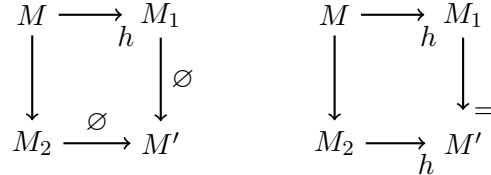


Figure 5: Elementary diagrams for simple  $M$ .

We briefly explain the two elementary diagrams. They depict the possible scenarios for joining co-initial steps  $M \rightarrow_h M_1$  and  $M \rightarrow M_2$ , that is, a head reduction step set out against an arbitrary reduction step. If  $M$  is a simple term, then either

- (i) both are the same step, and can be joined with empty steps  $\xrightarrow{\emptyset}$ , or

- (ii) they can be joined by  $M_2 \rightarrow_h M' \leftarrow^= M_1$  since a head step in a simple term with hnf cannot duplicate a redex (but deletion is possible).

We show the following implication by induction on  $n \in \mathbb{N}$  (employing the diagrams in Figure 5):

$$M_1 \leftarrow_h M \rightarrow^n N \quad \Longrightarrow \quad M_1 \rightarrow^{\leq n} O \leftarrow_h N \vee M_1 \rightarrow^{< n} N \quad (4.1)$$

For  $n = 0$ , there is nothing to be shown. Let  $n > 0$ , and consider  $M_1 \leftarrow_h M \rightarrow M_2 \rightarrow^{n-1} N$ . Then either (i)  $M_1 \equiv M_2$  and consequently  $M_1 \rightarrow^{n-1} N$ , or (ii)  $M_1 \rightarrow^= M' \leftarrow_h M_2$  and by induction hypothesis  $M' \rightarrow^{\leq n-1} O \leftarrow_h M_2$  or  $M' \rightarrow^{< n-1} N$ , yielding  $M_1 \rightarrow^{\leq n} O \leftarrow_h M_2$  or  $M_1 \rightarrow^{< n} N$ , respectively.

For  $\lambda$ -terms  $M$  and integers  $n \in \mathbb{N}$ , we define finite approximations  $\text{BT}_{\circlearrowleft, n}(M)$  of the clocked Böhm Tree  $\text{BT}_{\circlearrowleft}(M)$  of  $M$ . We let  $\text{BT}_{\circlearrowleft, 0}(M) = M$ , and for  $n > 0$  define

$$\text{BT}_{\circlearrowleft, n}(M) = [k]\lambda x_1 \dots \lambda x_n . y \text{BT}_{\circlearrowleft, n-1}(M_1) \dots \text{BT}_{\circlearrowleft, n-1}(M_m)$$

if  $M \rightarrow_h^k \lambda x_1 \dots \lambda x_n . y M_1 \dots M_m$  to hnf, and  $\text{BT}_{\circlearrowleft, n}(M) = \perp$  if  $M$  has no hnf. Note that  $\text{BT}_{\circlearrowleft}(M) = \lim_{n \rightarrow \infty} \text{BT}_{\circlearrowleft, n}(M)$ .

Let  $M, N$  be terms,  $M$  simple and  $M \rightarrow^n N$ . Then we claim

$$[\text{BT}_{\circlearrowleft, 1}(M)] \rightarrow^{\leq n'} [\text{BT}_{\circlearrowleft, 1}(N)] \quad (*)$$

where  $n' = n$  if  $\text{BT}_{\circlearrowleft, 1}(M)$  and  $\text{BT}_{\circlearrowleft, 1}(N)$  have the same annotation, and  $n' = n-1$  otherwise. If  $M$  has no hnf, then  $N$  has no hnf, and  $\text{BT}_{\circlearrowleft, 1}(M) = \text{BT}_{\circlearrowleft, 1}(N) = \perp$ . Thus assume that  $M$  admits a head reduction  $M \rightarrow_h^k M' \equiv \lambda x_1 \dots \lambda x_n . y M_1 \dots M_m$  to hnf for some  $k \in \mathbb{N}$ . Then by induction using (4.1) we obtain that either

$$(a) \ N \rightarrow_h^k N' \leftarrow^{\leq n} M' \quad \text{or} \quad (b) \ N \rightarrow_h^{< k} N' \leftarrow^{< n} M'$$

for some  $N' \equiv \lambda x_1 \dots \lambda x_n . y N_1 \dots N_m$ . This proves the claim.

Note that every reduction  $\text{BT}_{\circlearrowleft, m}(M) \rightarrow^{\leq n} \text{BT}_{\circlearrowleft, m}(N)$  is contained within the subterms  $M'$  of  $\text{BT}_{\circlearrowleft, m}(M)$  that are left unreduced by the clause  $\text{BT}_{\circlearrowleft, 0}(M') = M'$ ; everything outside these subterms is in normal form. Using the previous observation and (\*), we obtain by induction that for all  $m \in \mathbb{N}$ :  $[\text{BT}_{\circlearrowleft, m}(M)] \rightarrow^{\leq n-d_m} [\text{BT}_{\circlearrowleft, m}(N)]$  where  $d_m$  is the number of positions  $p$  where the annotation of  $\text{BT}_{\circlearrowleft, m}(M)|_p$  differs from the annotation of  $\text{BT}_{\circlearrowleft, m}(N)|_p$ . As a consequence, the annotations of  $\text{BT}_{\circlearrowleft}(M)$  and  $\text{BT}_{\circlearrowleft}(N)$  differ at most at  $n$  positions. Hence  $\text{BT}_{\circlearrowleft}(M) =_{\exists} \text{BT}_{\circlearrowleft}(N)$ .  $\square$

Reduction accelerates clocks, and for simple terms the clock is invariant under reduction, see Proposition 29. Hence if a term  $M$  has a simple reduct  $N$ , then  $N$  has the fastest clock reachable from  $M$  modulo a finite prefix. This justifies the following convention.

**Convention 30.** The (*minimal*) *clock* of a  $\lambda$ -term  $M$  with a simple reduct  $N$  is  $\text{BT}_{\circlearrowleft}(N)$ , the clocked BT of  $N$ .

For simple terms we obtain the following theorem:

**Theorem 31.** *Let  $M$  and  $N$  be  $\lambda$ -terms such that  $M$  is simple. If  $M$  does not improve eventually on  $N$ , then  $M \neq_{\beta} N$ .*

*Proof.* Assume  $M =_{\beta} N$ . Then  $M \twoheadrightarrow M' \leftarrow N$  for some  $M'$ . We have  $\text{BT}_{\circlearrowleft}(M) =_{\exists} \text{BT}_{\circlearrowleft}(M')$  by Proposition 29, and  $\text{BT}_{\circlearrowleft}(M') \leq \text{BT}_{\circlearrowleft}(N)$  by Proposition 25. Hence we obtain  $\text{BT}_{\circlearrowleft}(M) \leq_{\exists} \text{BT}_{\circlearrowleft}(N)$ .  $\square$

Theorem 31 significantly reduces the proof obligation in comparison to Theorem 26. We only consider the clocked BT's of  $M$  and  $N$ , instead of all reducts  $M'$  of  $M$ .

Note that Theorem 31 can also be employed for discriminating non-simple  $\lambda$ -terms if one of the terms has a simple reduct. For the case that both  $M$  and  $N$  are simple, there is no need to look for reducts since the eventual clocks are invariant under reduction, see Proposition 29:

**Corollary 32.** *If simple terms  $M, N$  do not match eventually, then they are not  $\beta$ -convertible, that is,  $M \neq_\beta N$ .*

*Proof.* Assume  $M = N$ , then  $M \twoheadrightarrow O \leftarrow N$  for a common reduct  $O$ . Then  $\text{BT}_{\circlearrowleft}^{\circ}(M) =_{\exists} \text{BT}_{\circlearrowleft}^{\circ}(O) =_{\exists} \text{BT}_{\circlearrowleft}^{\circ}(N)$  by Proposition 29. Hence  $\text{BT}_{\circlearrowleft}^{\circ}(M) =_{\exists} \text{BT}_{\circlearrowleft}^{\circ}(N)$ , that is,  $M$  and  $N$  match eventually.  $\square$

**Remark 33.** The reason for the qualifier ‘eventually’ in the notions above, in other words, working modulo a finite prefix of the BT, is that by some preliminary reduction we can always make the clock values in any finite prefix equal to 0. So we are interested exclusively in the ‘tail behaviour’, or the behaviour ‘at infinity’, and not in the initial behaviour of the development to the BT.

To give a concrete example:  $Y_0$  and  $Y_1$ , the fpc's of Curry and Turing, can be reduced to reducts  $M_0, M_1$  respectively, that have an initial segment of arbitrary length  $n$  of their BT's with clock labels 0 (just reduce first to  $\lambda f.f^n(\dots)$ ). However, the infinite remainders of their BT's, their tails as it were, will reveal the difference in clock values, witnessing the fact that  $Y_0$  eventually improves  $Y_1$ . And this situation is stable under reduction; indeed, for any two reducts  $M_0, M_1$  as above, the first eventually improves the second.

**Remark 34.** Take  $\lambda$ -terms  $M, N$  with finite BT's. Then the clock comparison of  $M$  and  $N$  amounts to the comparison of their non-clocked BT's. In case their BT's are equal and  $\perp$ -free, then  $M =_\beta N$ . If their BT's coincide but are not  $\perp$ -free, then we can fine-tune the analysis by comparing their clocked Lévy–Longo or Berarducci Trees (a  $\perp$  in a BT may give rise to an infinite Berarducci Tree), see Section 8.

**Example 35.** We compute the clocks of  $Y_n x$  with  $Y_n$  the  $n$ -th term of the Böhm sequence. The clocks of  $Y_0 x$  and  $Y_1 x$  have been computed before, see Figure 3. We now compute the clock of  $Y_n x$  for  $n \geq 2$ :

$$\begin{aligned} Y_n &\equiv \eta\eta\delta^{\sim n-1}x \\ &\rightarrow_h^2 \delta(\eta\eta\delta)\delta^{\sim n-2}x \\ &\rightarrow_h^{2(n-2)} \delta(\eta\eta\delta^{\sim n-1})x \\ &\rightarrow_h^2 x(\eta\eta\delta^{\sim n-1}x) \end{aligned}$$

Notice that none of these steps duplicate a redex, hence  $Y_n$  is a simple term. We find  $\text{BT}_{\circlearrowleft}^{\circ}(Y_n x) = [2n](x\text{BT}_{\circlearrowleft}^{\circ}(Y_n x))$ . Hence, for  $n \geq 2$  the clock of  $Y_n$  is  $2n$ . Consequently, all terms in the Böhm sequence have distinct clocks. Thus we have an alternative proof of Theorem 12: the Böhm sequence contains no duplicates.

**Example 36.** Let  $n \geq 2$ . We compute the clocks of the fpc's  $U_n = \text{BY}_0 S^{\sim n} \mathbf{1}$  of the Scott sequence. We first reduce  $U_n x$  to a simple term:

$$\begin{aligned} U_n x &\rightarrow Y_0(SS)S^{\sim(n-2)} \mathbf{1}x \\ &\rightarrow \omega_{SS} \omega_{SS} S^{\sim(n-2)} \mathbf{1}x \\ &\rightarrow \underline{\omega_{SS}} \underline{\omega_{SS}} S^{\sim(n-2)} \mathbf{1}x \end{aligned}$$

where  $\underline{\omega_{SS}} \equiv \lambda abc.bc(aabc)$ . We abbreviate  $\theta = \underline{\omega_{SS}}$ . Then we compute the clocks for  $n = 2$ ,  $n = 3$ , and  $n > 3$ :

$$\begin{aligned} \theta \theta \mathbf{1}x &\rightarrow_h^3 \mathbf{1}x(\theta \theta \mathbf{1}x) \rightarrow_h^1 x(\theta \theta \mathbf{1}x) \\ \theta \theta S \mathbf{1}x &\rightarrow_h^3 S \mathbf{1}(\theta \theta S \mathbf{1}x) \rightarrow_h^4 x(\theta \theta S \mathbf{1}x) \\ \theta \theta S^{\sim(n-2)} \mathbf{1}x &\rightarrow_h^3 SS(\theta \theta SS)S^{\sim(n-4)} \mathbf{1}x \\ &\rightarrow_h^{3(n-4)} SS(\theta \theta SSS^{\sim(n-4)}) \mathbf{1}x \\ &\rightarrow_h^3 S \mathbf{1}(\theta \theta S^{\sim(n-2)} \mathbf{1}x) \\ &\rightarrow_h^4 x(\theta \theta S^{\sim(n-2)} \mathbf{1}x) \end{aligned}$$

respectively. For all three cases, we find:

$$\text{BT}_{\odot}(\theta \theta S^{\sim(n-2)} \mathbf{1}x) = [3n - 2](x \text{BT}_{\odot}(\theta \theta S^{\sim(n-2)} \mathbf{1}x)).$$

Using Theorem 31 we infer from Example 36 and Figure 3 (recall that  $U_0 =_{\beta} Y_0$  and  $U_1 =_{\beta} Y_1$ ):

**Corollary 37.** *The Scott sequence contains no duplicates.*

## 5. AN ANSWER TO A QUESTION OF PLOTKIN

Plotkin [22] asked: *Is there a fixed point combinator  $Y$  such that*

$$A_Y \equiv Y(\lambda z.fzz) =_{\beta} Y(\lambda x.Y(\lambda y.fxy)) \equiv B_Y \quad (5.1)$$

or in other notation:

$$\mu z.fzz =_{\beta} \mu x.\mu y.fxy ,$$

with the usual definition  $\mu x.M(x) = Y(\lambda x.M(x))$ . The terms  $A_Y$  and  $B_Y$  have the same Böhm Trees, namely the solution of  $T = fTT$ .

Plotkin's question is pertinent to the question whether absolutely unorderable models of the  $\lambda$ -calculus exist, see Selinger's work [25, 26, 27]. The negative solution of Plotkin's question blocks an appeal to Lemma 3.6 in [27] to show that the generalized Mal'cev equations for all  $n$  are inconsistent with the  $\lambda\beta$ -calculus. (For  $n = 0, 1$  this is established, but the general case is a difficult open problem.)

We show that there is no fpc  $Y$  satisfying (5.1). We begin with an example.



**Example 38.** We consider Turing's fpc  $Y_1$ , and compute the clocked BT's of  $A_{Y_1}$  and  $B_{Y_1}$ . Recall that  $Y_1 \equiv \eta\eta$  with  $\eta \equiv \lambda x f.f(xxf)$ .

$$\begin{aligned} A_{Y_1} &\equiv \eta\eta(\lambda z.fzz) \rightarrow_h^2 (\lambda z.fzz)A_{Y_1} \rightarrow_h^1 fA_{Y_1}A_{Y_1} \\ B_{Y_1} &\equiv \eta\eta(\lambda x.\eta\eta(\lambda y.fxy)) \rightarrow_h^2 (\lambda x.\eta\eta(\lambda y.fxy))B_{Y_1} \\ &\rightarrow_h^1 \eta\eta(\lambda y.fB_{Y_1}y) \rightarrow_h^2 (\lambda y.fB_{Y_1}y)(\eta\eta(\lambda y.fB_{Y_1}y)) \\ &\rightarrow_h^1 fB_{Y_1}(\eta\eta(\lambda y.fB_{Y_1}y)) \end{aligned}$$

Note that for  $B_{Y_1}$  developing the left branch takes six steps, whereas the right branch only needs three. We remark that this is not sufficient to conclude that  $A_{Y_1} \not\equiv_\beta B_{Y_1}$  since  $A_Y$  and  $B_Y$  are non-simple even for simple fpc's  $Y$ . The clocked BT's for  $A_{Y_1}$  and  $B_{Y_1}$  are depicted in Figure 6 using hnf-notation, see [3] or [5].

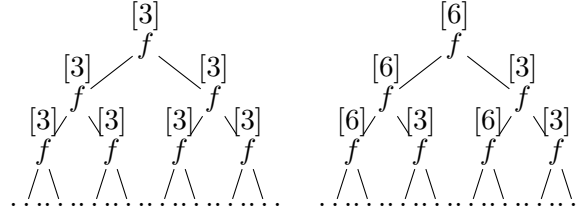


Figure 6: Clocked BT's for  $A_{Y_1}$  and  $B_{Y_1}$ , in hnf-notation.

To prove  $A_Y \not\equiv_\beta B_Y$  for every fpc  $Y$ , we construct a term  $B'_Y$  convertible with  $B_Y$  such that no reduct of  $A_Y$  improves globally on  $B'_Y$ . Then by Theorem 26 we have  $A_Y \not\equiv_\beta B'_Y$ , and thus  $A_Y \not\equiv_\beta B_Y$ . We have  $B_Y =_\beta B'_Y$  where  $B'_Y$  is defined by:

$$B'_Y \equiv Y(\lambda x.fx(fx(Y(\lambda y.fxy))))$$

While the annotations of the clocked Böhm Tree  $\text{BT}_{\text{c}}(B'_Y)$  of  $B'_Y$  of course depend on  $Y$ , we can derive partial knowledge as follows. We have  $B'_Y \equiv (Yx)[x := M]$ , where  $M \equiv \lambda x.fx(fx(Y(\lambda y.fxy)))$ . Then we have for every  $Y' =_\beta Yx$ :

$$\begin{aligned} Y'[x := M] &\rightarrow_h^* (xY'')[x := M] \equiv M(Y''[x := M]) \\ &\rightarrow_h f(Y''[x := M]) \underline{(f(Y''[x := M]) (Y(\lambda y.f(Y''[x := M])y)))} \end{aligned}$$

for some  $Y'' =_\beta Yx$ . The underlined term in head normal form is obtained without reducing, and hence will be annotated with [0] in the clocked Böhm tree of  $Y'[x := M]$ . The same argument applies for  $Y''[x := M]$ , and so on. Hence the second argument of every  $f$  on the leftmost spine of  $\text{BT}_{\text{c}}(Y'[x := M])$  has annotation [0], as in Figure 7, which depicts the clocked Böhm tree  $\text{BT}_{\text{c}}(B'_Y)$  of  $B'_Y$ . Therefore we obtain:

**Lemma 39.** *Every annotation at a position of the form  $(12)^*2$  in  $\text{BT}_{\text{c}}(B'_Y)$  is [0].*  $\square$

Note that in the statement of the above lemma, the positions  $(12)^n2$  refer to trees in applicative notation, which is equivalent to  $1^n2$  in the hnf-notation, as seen in Figure 7.

We continue by proving that no reduct of  $A_Y$  improves globally on  $B'_Y$ . More precisely, we show that for every reduct  $A'_Y$  of  $A_Y$  there exists a position  $p$  of the form  $(12)^*2$  such that the annotation of  $\text{BT}_{\text{c}}(A'_Y)$  at  $p$  is nonzero.

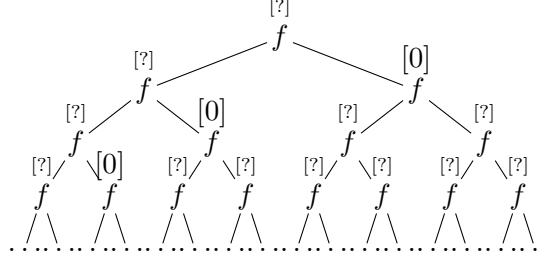


Figure 7: Clocked Böhm Tree for  $B'_Y$  in hnf-notation.

**Definition 40.** A reduct  $M$  of  $A_Y \equiv Y(\lambda z.fzz)$  is called *balanced* if  $s \equiv t$  for every subterm  $f s t$  in  $M$  where  $f$  is a descendant of the displayed  $f$  (see also Equation (5.1)).

Note that, in case  $Y$  is not closed, the term may contain occurrences  $f$ . In the definition of ‘balanced’ we are not interested in those occurrences of  $f$ , but only the residuals of  $f$  displayed in  $Y(\lambda z.fzz)$ , that is, the  $f$  in  $\lambda z.fzz$ . Let us label this  $f$  as  $f^*$ , i.e.  $Y(\lambda z.f^*zz)$ , such that  $f^*$  does not occur in  $Y$ . Then by ‘residuals of the displayed  $f$ ’ we mean the free occurrences of  $f^*$  in all reducts of  $Y(\lambda z.f^*zz)$ .

The following lemma states that every reduct of  $A_Y$  can be balanced:

**Lemma 41.** *Let  $A_Y \equiv Y(\lambda z.fzz) \twoheadrightarrow M$ . There exists a balanced  $N$  such that  $M \twoheadrightarrow N$ .*

*Proof.* Let  $\twoheadrightarrow$  denote complete developments of the set of all redex occurrences (also known as Gross–Knuth steps). We consider the sequence  $A_Y \equiv N_1 \twoheadrightarrow N_2 \twoheadrightarrow \dots$ . By cofinality of  $\twoheadrightarrow$  [3, 30] there exists  $i \in \mathbb{N}$  such that  $M \twoheadrightarrow N_i$ . We define  $N \equiv N_i$ .

It remains to be shown that the term  $N$  is balanced. This follows from the fact that  $A_Y$  is balanced, and that balancedness is preserved under  $\twoheadrightarrow$ . The latter can be seen as follows: consider a term  $C[f s s]$ . Obviously both displayed occurrences of  $s$  contain the same redexes, and the same variables bound from above  $f s s$  in  $C$ . Hence all descendants of both occurrences of  $s$  after  $\twoheadrightarrow$  will again be identical.  $\square$

**Lemma 42.** *Let  $A_Y \twoheadrightarrow A'_Y$ . There exists a position  $p$  of the form  $(12)^*2$  such that  $\text{BT}_{\text{clock}}(A'_Y)$  has a nonzero annotation at position  $p$ .*

*Proof.* Let  $A''_Y$  be a reduct of  $A'_Y$  such that  $A''_Y$  is in hnf. By Lemma 41 there exists a balanced reduct  $A'''_Y$  of  $A''_Y$ . Note that also  $A'''_Y$  is in hnf and its head symbol is  $f$ .

Let  $q$  be the shortest position of the form  $(12)^*$  such that  $A'''_Y|_{q12}$  is not in hnf. Then  $A'''_Y|_q$  is in hnf, and thus  $A'''_Y(q11) = f$ . Because  $A'''_Y$  is balanced we have that  $A'''_Y|_{q12} \equiv A'''_Y|_{q2}$ , and so  $A'''_Y|_{q2}$  is not in hnf, too. Hence the annotation of  $\text{BT}_{\text{clock}}(A'''_Y)$  at the position  $p = q2$  is nonzero. The claim follows since  $\text{BT}_{\text{clock}}(A'_Y) \leq \text{BT}_{\text{clock}}(A'''_Y)$ .  $\square$

The combination of Lemmas 39 and 42 implies that there exists no reduct of  $A_Y$  that improves globally on  $B'_Y$ . Therefore by Theorem 26 we obtain:

**Proposition 43.** *There exists no fpc  $Y$  such that  $Y(\lambda z.fzz) =_{\beta} Y(\lambda x.Y(\lambda y.fxy))$ .*  $\square$

## 6. CLOCKED BÖHM TREES AND PERIODIC TERMS

In the last two sections we saw that the concept of clocked Böhm Trees can be successfully used to discriminate between  $\lambda$ -terms with equal Böhm Trees or to analyze the existence of terms satisfying certain equations. In this section we will address the following question: to which classes of  $\lambda$ -terms is the clocks method most readily applicable?

A common feature of terms studied in Sections 4 and 5 is that their Böhm Trees are ‘periodic’: they have inherent cyclical content that allows us to analyze how their clocks are developed towards infinity. For example,  $Y$  is an fpc if the term  $Y_x \equiv Yx$  satisfies

$$Y_x = x(Y_x) \quad (6.1)$$

That is, the subterm of  $Y_x$  occurring at position 2 is equal to the term  $Y_x$  itself. If we use Böhm Tree equality instead of  $\beta$ -equality then the equation (6.1) would say that  $Y$  is a wfpc, or a looping combinator:

$$\text{BT}(Y_x) = \text{BT}(x(Y_x))$$

The next definition attempts to capture such situations in a general fashion.

In what follows,  $M_\sigma$  denotes the term at position  $\sigma$  in the Böhm Tree of  $M$ , with  $M_\sigma = \Omega = \mathbf{Y}_0!$  in the event that  $\sigma \notin \text{Pos}(\text{BT}(M))$ . By a *Böhm Tree context*  $C \square_1 \cdots \square_n$  we mean a finite Böhm Tree in which the leaves, in addition to being variables or  $\perp$ , are allowed to be either of the holes  $\square_i$ . (Equivalently, it is a  $\lambda\perp$ -context which is redex-free.)

**Definition 44.** Let  $M \in \text{Ter}(\lambda)$ , and  $\epsilon \neq \sigma \in \text{Pos}(\text{BT}(M))$ .

- (i)  $M$  is *periodic at  $\sigma$*  if  $M =_\beta M_\sigma$ .
- (ii)  $M$  is *weakly periodic at  $\sigma$*  if  $M =_{\text{BT}} M_\sigma$ .
- (iii)  $M$  is *(weakly) locally periodic* if for every infinite path  $x = \langle x_0, x_1, \dots \rangle$  through the Böhm Tree of  $M$  we have that  $M$  is (weakly) periodic at  $x_n$  for some  $n$ .
- (iv)  $M$  is *(weakly) globally periodic* if there is a Böhm Tree context  $C \square_1 \cdots \square_n$  such that  $M \equiv C[M_1] \cdots [M_n]$  and  $M =_\beta M_i$  ( $M =_{\text{BT}} M_i$ ).

**Example 45.**

- (i) If  $Y$  is an fpc, then  $Yx$  is periodic at 2.
- (ii) If  $Y$  is a weak fpc, then  $Yx$  is weakly periodic at 2.
- (iii) Using the fixed point theorem, let  $M$  be such that  $M = \lambda z.zMM$ . Then  $M$  is periodic at 02, 012. Since any infinite path through  $\text{BT}(M)$  passes through one of these positions,  $M$  is locally periodic. In fact,  $M$  is globally periodic with the context  $C \square_1 \square_2 \equiv \lambda z.z \square_1 \square_2$ .

Generally, any term which is defined by an application of the fixed point theorem is periodic by construction. If a looping combinator is used instead of an fpc, then the resulting term will be weakly periodic. This suggests that (w)fpc’s play a central role in the study of general periodic terms.

We prove the following elementary fact:

**Proposition 46.** *Let  $M \in \text{Ter}(\lambda)$ . Then*

$$M \text{ is (weakly) locally periodic} \iff M \text{ is (weakly) globally periodic}$$

*Proof.* The statement is proved simultaneously for weak and strong periodicity. For the latter case, the equality should be read as  $\beta$ -convertibility. For the former, it should be read as Böhm Tree equality.

( $\Rightarrow$ ) Suppose that  $M = C[M_1] \cdots [M_n]$ , with  $C \square_1 \cdots \square_n$  a Böhm Tree context. Let  $x$  be an infinite path through  $\text{BT}(M)$ . Since  $C \square_1 \cdots \square_n$  is finite,  $x$  must pass through one of the holes  $\square_i$ . Since  $M = M_i$  is periodic at this position,  $x$  passes through a periodic position. Since  $x$  was arbitrary,  $M$  is locally periodic.

( $\Leftarrow$ ) Suppose that  $M$  is locally periodic, and let  $\mathcal{X}$  denote the set of infinite paths through the Böhm Tree of  $M$  ( $\mathcal{X}$  could be called the *Böhm space* of  $M$ ). Since  $\text{BT}(M)$  is a finitely branching tree, the space  $\mathcal{X}$  is compact, when given the topology generated by the cylinders

$$X_\sigma = \{\theta \mid \sigma \sqsubseteq \theta\}$$

For  $x \in \mathcal{X}$ , let  $\sigma_x$  be a periodic position that  $x$  passes through (by local periodicity). Then the collection  $\Sigma = \{\sigma_x\}_{x \in \mathcal{X}}$  is an open cover of  $\mathcal{X}$ . By compactness, there exists a subcollection  $\sigma_1, \dots, \sigma_n \in \Sigma$  which covers  $\mathcal{X}$ . Let  $C \equiv C \square_1 \cdots \square_n$  be obtained from  $M$  by unfolding its Böhm Tree up to depth  $\max\{|\sigma_1|, \dots, |\sigma_n|\}$  and placing holes  $\square_i$  at positions  $\sigma_i$ . We claim that  $C$  is the desired Böhm Tree context.

By construction,  $C$  is a Böhm Tree context with  $M = C[M_{\sigma_1}] \cdots [M_{\sigma_n}]$ . We also have that  $M$  is periodic at  $\sigma_i$ , hence  $M = M_{\sigma_i}$ . Furthermore, if  $x \in \mathcal{X}$ , then  $x$  passes through  $\sigma_i$  for some  $i$ . Therefore, outside the holes  $\square_i$  the Böhm Tree of  $C$  is finite.

Indeed,  $M$  is globally periodic.  $\square$

In light of the proposition above, when  $M$  is either locally or globally periodic, we simply say that  $M$  is periodic.

The resulting notion of periodicity is somewhat restrictive, because it does not account for terms in which repetition begins later in the Böhm Tree. For example if  $Y$  is an fpc, then  $Yx$  is periodic, but  $Y$  itself is not, because the first node of its Böhm Tree has the abstraction  $\lambda f.f$ , which is never repeated in the infinite sequence of  $f$ 's that follows.

This situation motivates the following definition.

**Definition 47.** Suppose that  $\sigma \in \mathcal{P}os(\text{BT}(M))$  and  $\theta \sqsubseteq \sigma$ . Let  $\pi$  be such that  $\theta\pi = \sigma$ .

- (i)  $M$  is *periodic at  $\sigma$  with offset  $\theta$*  if  $M_\theta$  is periodic at  $\pi$ . In this case, we say that  $\sigma$  is *cyclic* in  $\text{BT}(M)$ , and call  $\pi$  the *period* and  $\theta$  the *phase* of the *cycle at  $\sigma$* .
- (ii)  $M$  is *weakly periodic at  $\sigma$  with offset  $\theta$*  if  $M_\theta$  is weakly periodic at  $\pi$ . We say that  $\sigma$  is *weakly cyclic*, and call  $\pi$  the period and  $\theta$  the phase of the *weak cycle at  $\sigma$* .
- (iii)  $M$  is (*weakly*) *fully periodic* if every infinite path through  $\text{BT}(M)$  passes through a (weak) cycle  $(\theta, \pi)$  (that is, eventually comes into a periodic subterm of  $M$ ).

As before, it is easy to see that  $M$  is fully periodic if and only if there are finitely many positions  $\sigma_1, \dots, \sigma_n \in \mathcal{P}os(\text{BT}(M))$ ,  $\sigma_i = \theta_i\pi_i$  such that  $M$  is periodic at  $\sigma_i$  with offset  $\theta_i$  and period  $\pi_i$ , and outside these positions the Böhm Tree is finite.

Periodic and fully periodic terms are the perfect targets for exercising the clocks method. We will now illustrate this with an example, where, instead of fpc's, we consider enumerators, also called evaluators, for Combinatory Logic. There are many possibilities for such an enumerator  $E$ , depending on how one does the coding. However, in most schemes  $E$  is a periodic term.

**Example 48.** Let  $\ulcorner \cdot \urcorner : CL \rightarrow \text{Ter}(\lambda)$  be defined as follows:

$$\ulcorner K \urcorner = \lambda z.zKKI$$

$$\ulcorner S \urcorner = \lambda z.zKSI$$

$$\ulcorner MN \urcorner = \lambda z.z(KI)\ulcorner M \urcorner \ulcorner N \urcorner$$

Let  $\langle M \rangle = \lambda z.zM$ , with  $z \notin \text{FV}(M)$  and  $\omega = \lambda x.xx$ . Then some possible evaluators for  $\ulcorner \cdot \urcorner$  include:

$$\mathbf{E}_1 \equiv \omega(\lambda w.\langle \lambda abc.ab((wwb)(wvc)) \rangle) = \langle \lambda abc.ab((\mathbf{E}_1 b)(\mathbf{E}_1 c)) \rangle \quad (6.2)$$

$$\mathbf{E}_2 \equiv \omega(\lambda w.\langle \lambda abc.ab(\mathbf{S}\langle b \rangle\langle c \rangle(ww)) \rangle) = \langle \lambda abc.ab(\mathbf{S}\langle b \rangle\langle c \rangle\mathbf{E}_2) \rangle \quad (6.3)$$

$$\mathbf{E}_3 \equiv \langle \omega(\lambda wabc.ab(\mathbf{S}bc(ww))) \rangle = \langle \mathbf{E}' \rangle, \quad \text{where } \mathbf{E}' = \lambda abc.ab(\mathbf{S}bc\mathbf{E}') \quad (6.4)$$

It is straightforward to verify that, for  $i \in \{1, 2, 3\}$  we have  $\mathbf{E}_i \ulcorner M \urcorner = M$  for any  $\{\mathbf{K}, \mathbf{S}\}$ -term  $M$ . The evaluators look similar, and their Böhm Trees are indeed the same. However, it is not immediate whether the terms are  $\beta$ -convertible or if they are distinct.

Fortunately, since these terms are periodic, their clocked Böhm Trees can be computed easily:

- (i) For  $N \in \text{Ter}(\lambda)$ , let  $V_N \equiv \lambda abc.ab((NNb)(NNc))$ , and  $W = \lambda w.\langle V_w \rangle$ . Then we have

$$\begin{aligned} \mathbf{E}_1 &\rightarrow_h WW \rightarrow_h \lambda z.zV_W \\ &\equiv \lambda z.z(\lambda abc.ab((WWb)(WWc))) \\ &\rightarrow_h \lambda z.z(\lambda abc.ab((\lambda z'.z'V_W)b(WWc))) \\ &\rightarrow_h \lambda z.z(\lambda abc.ab(bV_W(WWc))) \\ &\rightarrow \lambda z.z(\lambda abc.ab(bV_W((\lambda z'.z'V_W)c))) \\ &\rightarrow \lambda z.z(\lambda abc.ab(bV_W(cV_W))) \\ &\rightarrow \dots \end{aligned}$$

Hence  $\mathbf{E}_1$  is periodic at positions 02000212 and 02000222 (in hnf-notation 010 and 0110 respectively) with offset 02 and periods 000212 and 000222, which allows us to construct the clocked Böhm Tree of  $\mathbf{E}_1$ .

- (ii) For  $N \in \text{Ter}(\lambda)$ , let  $V_N \equiv \lambda abc.ab(\mathbf{S}\langle b \rangle\langle c \rangle(ww))$ ,  $W = \lambda w.\langle V_w \rangle$ ,  $\mathbf{S}'_N = \lambda yz.Nz(yz)$ , and  $\mathbf{S}''_{N,M} = \lambda z.Nz(Mz)$ . Then we have

$$\begin{aligned} \mathbf{E}_2 &\rightarrow_h WW \rightarrow_h \lambda z.zV_W \\ &\equiv \lambda z.z(\lambda abc.ab(\mathbf{S}\langle b \rangle\langle c \rangle(WW))) \\ &\rightarrow_h \lambda z.z(\lambda abc.ab(\mathbf{S}'_{\langle b \rangle}\langle c \rangle(WW))) \\ &\rightarrow_h \lambda z.z(\lambda abc.ab(\mathbf{S}''_{\langle b \rangle, \langle c \rangle}(WW))) \\ &\rightarrow_h \lambda z.z(\lambda abc.ab(\langle b \rangle(WW)(\langle c \rangle(WW)))) \\ &\rightarrow_h \lambda z.z(\lambda abc.ab(WWb(WWc))) \\ &\rightarrow^2 \lambda z.z(\lambda abc.ab((\lambda z'.z'V_W)b((\lambda z'.z'V_W)c))) \\ &\rightarrow^2 \lambda z.z(\lambda abc.ab((bV_W)(cV_W))) \\ &\rightarrow \dots \end{aligned}$$

Again, we are at a point where every position lies on a cycle, giving us a full description of the clocked Böhm Tree of  $\mathbf{E}_2$ .

(iii) Let  $V_N = \lambda abc.ab(\text{Sbc}(NN))$ , and  $W, S'_N, S''_{N,M}$  as before. Then

$$\begin{aligned}
 E_3 &\equiv \langle \omega W \rangle \rightarrow_h \langle WW \rangle \rightarrow_h \lambda z.zV_W \\
 &\equiv \lambda z.z(\lambda abc.ab(\text{Sbc}(WW))) \\
 &\rightarrow_h \lambda z.z(\lambda abc.ab(S'_b c(WW))) \\
 &\rightarrow_h \lambda z.z(\lambda abc.ab(S''_{b,c}(WW))) \\
 &\rightarrow_h \lambda z.z(\lambda abc.ab(b(WW)(c(WW))))
 \end{aligned}$$

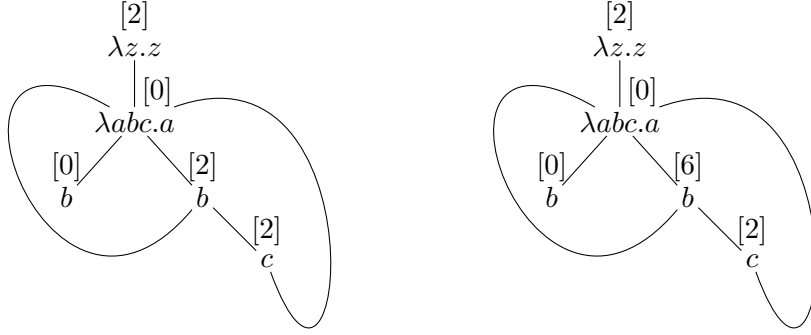


Figure 8: Clocked BT's for  $E_1$  and  $E_2$ , in hnf-notation.

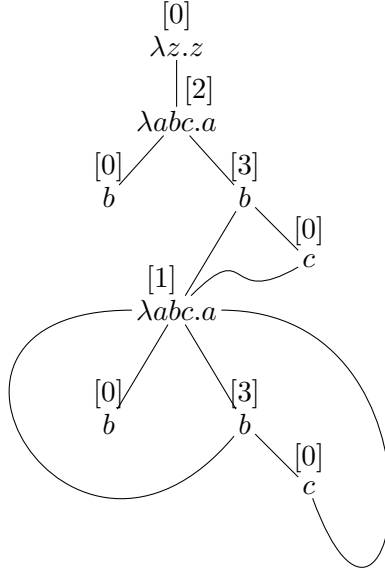


Figure 9: Clocked BT for  $E_3$ , in hnf-notation.

The Böhm Trees of  $E_i$  are displayed in Figures 8 and 9. This information, together with Theorem 31, allows us to discriminate between the terms. Note that  $E_1$  does not

eventually improve on  $E_3$ : the occurrences of the variable  $c$  in the clocked Böhm Tree of  $E_3$  are annotated by a 0, while the corresponding subterms of  $E_1$  take two head steps to converge. Furthermore, the head redexes contracted in this reduction  $WW$  and  $\langle V_W \rangle c$  are both call-by-value. By Theorem 31 we conclude that  $E_3 \not\equiv_{\beta} E_1$ . In contrast,  $E_2$  is improved by  $E_1$ . In fact,  $E_1$  and  $E_2$  can be reduced so that their clocked Böhm Trees match eventually. Upon further inspection, we see that the two terms are indeed convertible.

**Remark 49.** We note that (fully) periodic terms are the  $\lambda$ -calculus analogue of the usual ‘circular’ terms defined by mutually recursive definitions. The `let` binder in Haskell is an example of such a construct — one difference being that, because the semantics of Haskell is graph rewriting, such terms actually denote possibly cyclic graphs rather than their unfoldings as infinite trees. For example, the following is an example of a ‘periodic Haskell term’:

```
streamSys =
  let x = f 0 y
      where f n (b:ys) = n : f (b n) ys
          y = f id z
              where f g (c:cs) = g : f (\n -> g c + n) cs
          z = zz 0
              where zz n = n : zz (n+1)
  in x
```

```
*Main> take 20 streamSys
[0,0,0,1,4,10,20,35,56,84,120,165,220,286,364,455,560,680,816,969]
```

In discussion of such recursive systems, a well-known topic is the distinction between ‘strong equality’ and ‘weak equality’; the former being syntactic convertibility between two such systems, and the latter being bisimulation or ‘behavioral equivalence’. For this reason, the complexity of strong equality is usually  $\Sigma_1$ , since one only needs to check the existence of a proof object (the conversion sequence) for the equality between two terms, while weak equality is  $\Pi_2$ , since it asks whether for every number  $n$ , there is a conversion that makes the terms coincide up to depth  $n$ .

This discussion applies to our notions of fpc’s and wfpc’s, and also periodic and weakly periodic terms. From this point of view, the clocked Böhm Trees offer a refinement of these two types of equality, replacing dichotomy with a spectrum, or hierarchy, of equality, by augmenting the syntactic shape of the terms with the information of how quickly the computation they define is performed. For the case of  $\lambda$ -calculus, this spectrum was displayed in Figure 1.

Concluding this elaboration on cyclic terms, we note that also for the fragment of  $\lambda$ -calculus consisting of  $\mu$ -terms, with the definition of the corresponding  $\mu$ -rule  $\mu x.A(x) \rightarrow A(\mu x.A(x))$ , the notion of clocked Böhm Trees may be interesting, as it yields an equality strictly in between weak and strong equality, as it is called in [12, 4]. There  $\mu$ -terms are treated extensively, for their application as a representation of recursive types.

## 7. ATOMIC CLOCKS

We have introduced clocked Böhm Trees for discriminating  $\lambda$ -terms. In this section, we refine the clocks to measure not only the *number* of head steps, but, in addition, the

position of each of these steps. We call these clocks ‘atomic’. We write  $\rightarrow_{h,p}$  for the head reduction step at position  $p$ .

Before we give the formal definition (Definition 50), we consider a motivating example. We discriminate  $Y_2$  from  $U_2$ . First, we reduce both terms to simple reducts:

$$\begin{aligned} Y_2x &\equiv Y_0\delta\delta x \rightarrow \eta\eta\delta x && \text{where } \eta = \lambda ab.b(aab) \\ U_2x &\equiv Y_0(SS)lx \rightarrow \theta\theta lx && \text{where } \theta = \lambda abc.bc(aabc) \end{aligned}$$

Second, we compute the atomic clocks of these simple reducts, that is, the positions of the head steps, as follows:

$$\begin{aligned} \eta\eta\delta x &\rightarrow_{h,11} (\lambda b.b(\eta\eta b))\delta x \rightarrow_{h,1} \delta(\eta\eta\delta)x \\ &\rightarrow_{h,1} (\lambda b.b(\eta\eta\delta b))x \rightarrow_{h,\epsilon} x(\eta\eta\delta x) \\ \theta\theta lx &\rightarrow_{h,11} (\lambda bc.bc(\theta\theta bc))lx \rightarrow_{h,1} (\lambda c.lc(\theta\theta lc))x \\ &\rightarrow_{h,\epsilon} lx(\theta\theta lx) \rightarrow_{h,1} x(\theta\theta lx) \end{aligned}$$

Thus the atomic clocks of these terms are:

$$\begin{aligned} \text{BT}_{\bullet,\bullet}(\eta\eta\delta x) &= [11, 1, 1, \epsilon](x\text{BT}_{\bullet,\bullet}(\eta\eta\delta x)) \\ \text{BT}_{\bullet,\bullet}(\theta\theta lx) &= [11, 1, \epsilon, 1](x\text{BT}_{\bullet,\bullet}(\theta\theta lx)) \end{aligned}$$

Note that both terms have the (non-atomic) clocked Böhm Tree  $T \equiv [4](xT)$ . Hence the method from the previous section is not applicable. However, the atomic clocks do allow us to discriminate the terms. Hence  $Y_2 \neq_{\beta} U_2$  (by Corollary 32 which generalises to the setting of atomic BT’s). Note that the (non-atomic) clocked BT’s can be obtained by taking the length of the lists of positions.

For lists  $\vec{p}, \vec{q}$  of positions, we write  $\vec{p} \cdot \vec{q}$  for concatenating  $\vec{p}$  to  $\vec{q}$ . We write  $\rightarrow_{h,\langle p_1, \dots, p_n \rangle}$  for the rewrite sequence  $\rightarrow_{h,p_1} \cdots \rightarrow_{h,p_n}$  consisting of steps at position  $p_1, \dots, p_n$ .

**Definition 50** (Atomic clock Böhm Trees). Let  $M \in \text{Ter}(\lambda)$ . The *atomic clock Böhm Tree*  $\text{BT}_{\bullet,\bullet}(M)$  of  $M$  is an annotated infinite term defined as follows. If  $M$  has no hnf, then define  $\text{BT}_{\bullet,\bullet}(M)$  as  $\perp$ . Otherwise, there is a head reduction

$$M \rightarrow_{h,p_1} \cdots \rightarrow_{h,p_k} \lambda x_1. \dots \lambda x_n. y M_1 \dots M_m$$

of length  $k$  to hnf. Then we define  $\text{BT}_{\bullet,\bullet}(M)$  as the term:

$$\text{BT}_{\bullet,\bullet}(M) = [\langle p_1, \dots, p_k \rangle] \lambda x_1. \dots \lambda x_n. y \text{BT}_{\bullet,\bullet}(M_1) \dots \text{BT}_{\bullet,\bullet}(M_m)$$

The theory developed for (non-atomic) BT’s in Section 4 generalises to atomic trees, as follows.

**Theorem 51.** *For lists of positions  $\vec{p}, \vec{q}$  we define  $\vec{p} \geq \vec{q}$  whenever  $\vec{q}$  is a subsequence of  $\vec{p}$ , and  $\vec{p} > \vec{q}$  if additionally  $\vec{p} \neq \vec{q}$ . Here  $\langle a_1, \dots, a_n \rangle$  is a subsequence of  $\langle b_1, \dots, b_m \rangle$  if there exist indexes  $i_1 < i_2 < \dots < i_n$  such that  $\langle a_1, \dots, a_n \rangle = \langle b_{i_1}, \dots, b_{i_n} \rangle$ .*

*Using this notation for comparing the atomic annotations (lists of positions), Proposition 25, Theorem 26, Proposition 29, Theorem 31, and Corollary 32 remain valid.*

As an application of using atomic clocks to discriminate  $\lambda$ -terms, we show that every combination of the fpc-generating vectors  $\square(SS)S^{\sim n}I$  from Theorem 18 applied to Curry’s fpc  $Y_0$  gives rise to inconvertible fpc’s. We note that this cannot be proved using non-atomic clocks, as for example we have  $\text{BT}_{\bullet,\bullet}(Y_0(SS)S^{\sim n}I(SS)S^{\sim m}I) = \text{BT}_{\bullet,\bullet}(Y_0(SS)S^{\sim m}I(SS)S^{\sim n}I)$ .



**Proposition 52.** *Let  $\mathbf{G}_n = \square(\mathbf{SS})\mathbf{S}^{\sim n}\mathbf{I}$  the fpc-generating vectors from Theorem 18. For  $n_1, \dots, n_k \in \mathbb{N}$  we define*

$$\mathbf{Y}^{\langle n_1, \dots, n_k \rangle} = \mathbf{G}_{n_k}[\dots \mathbf{G}_{n_1}[\mathbf{Y}_0] \dots]$$

All these fpc's are inconvertible, that is,  $\vec{n} \neq \vec{m}$  implies  $\mathbf{Y}^{\vec{n}} \neq_{\beta} \mathbf{Y}^{\vec{m}}$ .

*Proof.* In this proof we abbreviate context filling by simply concatenating the contexts, so that for  $\mathbf{Y}^{\langle n_1, \dots, n_k \rangle}$  defined above we have  $\mathbf{Y}^{\langle n_1, \dots, n_k \rangle} = \mathbf{Y}_0 \mathbf{G}_{n_1} \dots \mathbf{G}_{n_k}$ .

For  $\vec{p} = \langle 1^{m_1}, \dots, 1^{m_k} \rangle$  a list of positions we define  $\vec{p} \times n$  as follows:

$$\begin{aligned} \vec{p} \times 0 &= \langle \rangle \\ \vec{p} \times n &= \vec{p} \cdot (\langle 1^{m_1-1}, \dots, 1^{m_k-1} \rangle \times (n-1)) \quad (n > 0) \end{aligned}$$

For example, we have  $\langle 1^5 \rangle \times 3 = \langle 1^5, 1^4, 1^3 \rangle$ .

We also use the following abbreviations, for  $n \in \mathbb{N}$ :

$$\begin{aligned} \mathbf{G}'_n &= \square \mathbf{S}^{\sim n} \mathbf{I} & \theta &= \lambda abc.bc(aabc) \\ \underline{\mathbf{G}}_n &= \square \zeta \mathbf{S}^{\sim n} \mathbf{I} & \zeta &= \underline{\mathbf{SS}} = \lambda abc.bc(abc) \end{aligned}$$

(Note that  $\omega_{\zeta} \rightarrow_{\beta} \theta$ .) For  $n_1, \dots, n_k \in \mathbb{N}$  we define an fpc  $\mathbf{Y}$  by

$$\mathbf{Y} = \theta \theta \mathbf{G}'_{n_1} \underline{\mathbf{G}}_{n_2} \dots \underline{\mathbf{G}}_{n_k}$$

so that we have the following reduction:

$$\mathbf{Y}^{\langle n_1, \dots, n_k \rangle} x \equiv \mathbf{Y}_0 \mathbf{G}_{n_1} \dots \mathbf{G}_{n_k} x \rightarrow \mathbf{Y} x \quad (k \geq 1)$$

and we observe that  $\mathbf{Y} x$  is a simple term (as can be inferred from the reductions below).

Let  $V_m = \square N_1 \dots N_m$  be a vector of length  $m$ . Then we have, for every  $n \geq 0$ :

$$\begin{aligned} \theta \theta \mathbf{G}'_n V_m &\equiv \theta \theta \mathbf{S}^{\sim n} \mathbf{I} V_m \\ &\rightarrow_{h, \langle (1^{n+m+1}) \times 3 \rangle \times n} \mathbf{SI}(\theta \theta \mathbf{S}^{\sim n} \mathbf{I}) V_m & (\dagger)_{n,m} \\ &\rightarrow_{h, \langle 1^{m+1} \rangle \times 2} (\lambda c.lc(\theta \theta \mathbf{S}^{\sim n} \mathbf{I} c)) V_m \end{aligned}$$

For  $M$  an arbitrary term, we have the following reductions  $(*)_{n,m}$  ( $n, m \geq 0$ ):

$$\begin{aligned} (\lambda c.lc(Mc)) \underline{\mathbf{G}}_n V_m &\equiv (\lambda c.lc(Mc)) \zeta \mathbf{S}^{\sim n} \mathbf{I} V_m \\ &\rightarrow_{h, \langle 1^{n+m+1} \rangle} \mathbf{I} \zeta (M \zeta) \mathbf{S}^{\sim n} \mathbf{I} V_m \\ &\rightarrow_{h, \langle 1^{n+m+2} \rangle} \zeta (M \zeta) \mathbf{S}^{\sim n} \mathbf{I} V_m & (*)_{n,m} \\ &\rightarrow_{h, \langle (1^{n+m+1}) \times 3 \rangle \times n} \mathbf{SI}(M \underline{\mathbf{G}}_n) V_m \\ &\rightarrow_{h, \langle 1^{m+1} \rangle \times 2} (\lambda c.lc(M \mathbf{G}_n c)) V_m \end{aligned}$$

Moreover, let  $(\ddagger)$  denote the following rewrite sequence:

$$(\lambda c.lc(\mathbf{Y}c))x \rightarrow_{h, \langle 1^0, 1^1 \rangle} x(\mathbf{Y}x) \quad (\ddagger)$$

The rewrite sequence  $\mathbf{Y}x \rightarrow_h x(\mathbf{Y}x)$  is composed of  $k$  subsequences as follows:

$$(\dagger)_{n_1, m_1}, (* )_{n_2, m_2}, (* )_{n_3, m_3}, \dots, (* )_{n_k, m_k}, (\ddagger) \quad (*)$$

where  $m_i$  is defined by  $m_k = 1$  and  $m_i = m_{i+1} + n_{i+1} + 2$  ( $1 \leq i < k$ ). For a rewrite sequence  $\sigma$ , let  $\langle \sigma \rangle$  denote the sequence of positions of the steps in  $\sigma$ . Note that the atomic clock Böhm Tree of  $\mathbf{Y}^{\vec{n}}x$  is of the form  $\mathbf{BT}_{\bullet}(\mathbf{Y}^{\vec{n}}x) = [\langle (*) \rangle] x([\langle (*) \rangle] x(\dots))$ .

As the goal is to prove inconvertibility, we may without loss of generality assume that  $n_k \neq 0$  since the fact  $MN \neq_{\beta} M'N \implies M \neq_{\beta} M'$  allows us to append  $\mathbf{G}_n$  with arbitrary

$n > 0$ . We argue that the starting positions of the subsequences  $\langle (*_{n_i, m_i}) \rangle$  in  $\langle (*_{n_i, m_i}) \rangle$  coincide with the occurrences of subsequences of the form  $(\blacklozenge) 1^l, 1^{l+1}, 1^l, 1^{l-1}, 1^{l-2}$  (for some  $l \in \mathbb{N}$ ), that is, an increment followed by four decrements. (In particular, we find exactly the patterns for  $l = n_i + m_i + 1$ .) This suffices to derive  $k, n_1, \dots, n_k$  since then the number of occurrences  $(\blacklozenge)$  is  $k - 1$ , and the  $n_i$ 's are a function of the length of the blocks; see also Example 53 below. This shows that  $\vec{n} \neq \vec{m}$  implies that  $\neg(\text{BT}_{\blacklozenge}(\mathbf{Y}^{\vec{n}}) \equiv \text{BT}_{\blacklozenge}(\mathbf{Y}^{\vec{m}}))$ , and hence we conclude  $\mathbf{Y}^{\vec{n}} \not\equiv_{\beta} \mathbf{Y}^{\vec{m}}$  by atomic version of Corollary 32, see Theorem 51.

For the sequences  $\langle (*_{n_i, m_i}) \rangle$  where  $n_i > 0$ , notice that  $\langle (*_{n_i, m_i}) \rangle$  starts with the positions  $1^{n_i+m_i+1}, 1^{n_i+m_i+2}, 1^{n_i+m_i+1}, 1^{n_i+m_i}, 1^{n_i+m_i-1}$ , a subsequence of the form  $(\blacklozenge)$ ; this is the only occurrence of four consecutive decrements in  $\langle (*_{n_i, m_i}) \rangle$ . For sequences  $\langle (*_{n_i, m_i}) \rangle$  with  $n_i = 0$ , we have  $i < k$  (since  $n_k \neq 0$ ), and  $\langle (*_{n_i, m_i}) \rangle$  is of the form  $1^{m_i+1}, 1^{m_i+2}, 1^{m_i+1}, 1^{m_i}$ . This combined with the first element  $m_i - 1 = m_{i+1} + n_{i+1} + 1$  of  $\langle (*_{n_{i+1}, m_{i+1}}) \rangle$  is an occurrence of the form  $(\blacklozenge)$ . Finally, we need to check that there are no other occurrences of  $(\blacklozenge)$  in  $\langle (*_{n_i, m_i}) \rangle$ . Note that other occurrences of four consecutive decrements can only occur as overlaps between  $\langle (*_{n_i, m_i}) \rangle$  and  $\langle (*_{n_{i+1}, m_{i+1}}) \rangle$ . Each of the sequences  $\langle (*_{n_{i+1}, m_{i+1}}) \rangle$  starts with an increment  $1^{n_{i+1}+m_{i+1}+1}, 1^{n_{i+1}+m_{i+1}+2}$ , thus only the first element can overlap. For  $n_i = 0$ , we have already analyzed the overlap, and for  $n_i > 1$ , only the last three elements of  $\langle (*_{n_i, m_i}) \rangle$  are not decreasing. This concludes the proof.  $\square$

**Example 53.** Let  $\mathbf{Y} = \theta\theta\mathbf{G}'_2\mathbf{G}_0\mathbf{G}_1$  so that  $\mathbf{Y}^{(2,0,1)}x \rightarrow \mathbf{Y}x$  as in the above proof. According to the calculations in the proof, the atomic clock of the simple term  $\mathbf{Y}x$  is

$$\langle 1^9, 1^8, 1^7, 1^8, 1^7, 1^6, 1^7, 1^6, \underbrace{1^5, 1^6, 1^5, 1^4}_{(\blacklozenge)}, \overbrace{1^3, 1^4, 1^3, 1^2, 1^1, 1^2, 1^1, 1^0, 1^1}^{(\blacklozenge)} \rangle$$

where we have indicated the occurrences of  $(\blacklozenge)$ , and have alternately used italics to separate the blocks  $(\dagger)_{2,6}$ ,  $(*)_{0,4}$ ,  $(*)_{1,1}$ , and  $(\ddagger)$ .

## 8. CLOCKED LÉVY–LONGO AND BERARDUCCI TREES

In fact, there are three main semantics for the  $\lambda$ -calculus:  $\text{BT}$ ,  $\text{LLT}$ , and  $\text{BeT}$ ; see [1, 6, 8, 18, 5]. In the Böhm Tree semantics, a term is meaningful only if it has a hnf. The Lévy–Longo semantics weakens this condition to whnf's, and thereby allows more terms to be distinguished. The Berarducci Tree semantics is a further weakening where only root-active terms are discarded as meaningless.

The notions from the Sections 4 and 7 generalize directly to  $\text{LLT}$  and  $\text{BeT}$  semantics. We only treat the non-atomic versions here.

**Definition 54** (Clocked Lévy–Longo Trees). Let  $M$  be a  $\lambda$ -term. The *clocked Lévy–Longo Tree*  $\text{LLT}_{\odot}(M)$  of  $M$  is an annotated potentially infinite term defined as follows. If  $M$  has no whnf, then define  $\text{LLT}_{\odot}(M)$  as  $\perp$ . Otherwise, there exists a head rewrite sequence  $M \rightarrow_h^k \lambda x.N$  or  $M \rightarrow_h^k xM_1 \dots M_m$  to whnf. In this case, we define  $\text{LLT}_{\odot}(M)$  as  $[k]\lambda x.\text{LLT}_{\odot}(N)$  or  $[k]x\text{LLT}_{\odot}(M_1) \dots \text{LLT}_{\odot}(M_m)$ , respectively.

**Definition 55** (Clocked Berarducci Trees). Let  $M$  be a  $\lambda$ -term. The *clocked Berarducci Tree*  $\text{BeT}_{\odot}(M)$  of  $M$  is an annotated potentially infinite term defined as follows. If  $M$  is root-active, let  $\text{BeT}_{\odot}(M) \equiv \perp$ . If  $M \rightarrow_h^k N$  rewrites to a root-stable term  $N \equiv x, N \equiv \lambda x.P$  or  $N \equiv PQ$ , then define  $\text{BeT}_{\odot}(M)$  as  $[k]x, [k]\lambda x.\text{BeT}_{\odot}(P)$  or  $[k]\text{BeT}_{\odot}(P)\text{BeT}_{\odot}(Q)$ , respectively.

**Example 56.** Consider the terms  $M = PP$  with  $P = \lambda x.\lambda y.xx$  and  $N = QQ$  with  $Q = \lambda x.\lambda y.\lambda z.xx$ . Then we find

$$\begin{aligned}\text{LLT}_{\circlearrowleft}(M) &\equiv [1]\lambda y.\text{LLT}_{\circlearrowleft}(M) \\ \text{LLT}_{\circlearrowleft}(N) &\equiv [1]\lambda y.[0]\lambda z.\text{LLT}_{\circlearrowleft}(N)\end{aligned}$$

Thus, in  $\text{LLT}_{\circlearrowleft}(M)$  every  $\lambda$  requires one head reduction step whereas in  $\text{LLT}_{\circlearrowleft}(N)$  every second  $\lambda$  is obtained for ‘free’ (that is, in 0 steps).

We remark that  $M$  and  $N$  cannot be distinguished in the Böhm Tree semantics since  $\text{BT}_{\circlearrowleft}(M) \equiv \text{BT}_{\circlearrowleft}(N) \equiv \perp$ .

## 9. CONCLUDING REMARKS

We conclude with an encompassing conjecture, and some further research questions.

**Conjecture.** *Building fpc’s with fpc-generating vectors is a free construction, that is, there are no non-trivial identifications.*

A first step is found in Intrigila’s theorem  $Y\delta \neq_{\beta} Y$ , for any fpc  $Y$ . A second step is that the Böhm sequence is duplicate-free. A third step is found in our proof that the Scott sequence is duplicate-free, and Proposition 52, which states that there are no identifications when starting the construction with  $Y_0$ .

Other parts of the conjecture are as follows. Let  $Y, Y'$  be fpc’s and  $B_1 \dots B_n, C_1 \dots C_k$  be fpc-generating vectors.

- (i)  $Y\delta =_{\beta} Y'\delta$  iff  $Y =_{\beta} Y'$ .
- (ii)  $YB_1 \dots B_n =_{\beta} Y'B_1 \dots B_n$  iff  $Y = Y'$ .
- (iii)  $YB_1 \dots B_n \neq_{\beta} YC_1 \dots C_k$  if  $B_1 \dots B_n \not\equiv C_1 \dots C_k$ .

For general fpc’s  $Y, Y'$  these conjectures may be beyond current techniques, but for the well-known fpc’s of Curry and Turing, and the fpc-generating vectors introduced here, including their versions for  $n > 3$ , these problems are tractable.

Other directions of research could be

- (i) For atomic clock Böhm Trees (Section 7) we have recorded the positions of head reduction steps building up the head normal form. What about a generalization to other spine reduction strategies [3]? Would this give rise to a stronger discrimination method?
- (ii) The notion of simple terms could be refined by focusing on an infinite path in the clocked Böhm Trees. Then duplication of redexes may be allowed along other paths, thereby making the method applicable to a larger class of terms.
- (iii) What general condition on a given term’s head reduction is sufficient to ensure that it possesses a ‘minimal clock’?
- (iv) Is it possible to characterize fully cyclic terms using a coinductive version of simple type theory?
- (v) Can the notion of a clock itself be given a type-theoretic interpretation?
- (vi) Can the clocks method be used to give a quantitative measure for optimization of functional programs?
- (vii) Is it possible, using the clocks method, to extend Intrigila’s result [16], and prove that there is no fixed point combinator  $Y$  and no  $n \in \mathbb{N}$  such that  $Y =_{\beta} Y\delta^{\sim n}$ ? This is an instance of the conjecture above.

- (viii) Also interesting is to systematically study all solutions of equations  $M\vec{x} = \vec{x}(M\vec{x})$  (we might call them *vector-fpc's*) where it is understood that  $\vec{x} = x_1 \dots x_n$  associates to the right when it occurs at an active position  $\vec{x}M = x_1(x_2(\dots(x_nM)\dots))$ , and to the left if it occurs in a passive position  $M\vec{x} = Mx_1x_2\dots x_n$ . Can these solutions  $M$  be discriminated by the methods presented here? Notice that the terms  $M$  lead to fpc's; for example  $MI^{\sim n-1}$  is an fpc (see Section 3.3), or more generally,  $M\vec{N}$  with  $\vec{N} = N_1 \dots N_{n-1}$  is an fpc when  $\vec{N}x \rightarrow_{\beta} x$ .
- (ix) Another interesting notion is that of *prime fpc's*, that is, fpc's  $Y$  not of the form  $Y = Y'\vec{P}$  where  $Y'$  is an fpc. In  $\lambda$ -calculus no prime fpc's exist, due to  $Y =_{\beta} Y(KY)$ , but in the  $\lambda I$ -calculus the notion is non-trivial.

**Acknowledgements.** We thank Raymond Smullyan for raising our awareness of the mystery and magic surrounding fixed point combinators, Hans Zantema and Johannes Waldmann for communicating some facts about Smullyan's Owl (SI), Henk Barendregt for intensive discussions which led us to simple terms, Rick Statman for his stimulating interest in this work and suggesting some future elaboration, and Gordon Plotkin for his question and subsequent communication pertinent to Selinger's work concerning the existence of unorderable models of the  $\lambda$ -calculus.

## REFERENCES

- [1] S. Abramsky and C.-H.L. Ong. Full Abstraction in the Lazy Lambda Calculus. *Information and Computation*, 105(2):159–267, 1993.
- [2] K. Aehlig and F. Joachimski. On Continuous Normalization. In *Proc. Workshop on Computer Science Logic (CSL 2002)*, volume 2471 of *LNCS*, pages 59–73. Springer, 2002.
- [3] H.P. Barendregt. *The Lambda Calculus. Its Syntax and Semantics*, volume 103 of *Studies in Logic and The Foundations of Mathematics*. North-Holland, revised edition, 1984.
- [4] H.P. Barendregt, W. Dekkers, and R. Statman. *Lambda Calculus with Types*. Perspectives in Logic. Cambridge University Press, 2011.
- [5] H.P. Barendregt and J.W. Klop. Applications of Infinitary Lambda Calculus. *Information and Computation*, 207(5):559–582, 2009.
- [6] A. Berarducci and B. Intrigila. Church–Rosser  $\lambda$ -theories, Infinite  $\lambda$ -calculus and Consistency Problems. *Logic: From Foundations to Applications*, pages 33–58, 1996.
- [7] I. Bethke. Lambda Calculus. Chapter 10 in [30].
- [8] I. Bethke, J.W. Klop, and R.C. de Vrijer. Descendants and Origins in Term Rewriting. *Information and Computation*, 159(1–2):59–124, 2000.
- [9] C. Böhm. The CUCH as a Formal and Description Language. *Annual Review in Automatic Programming*, 3:179–197, 1963.
- [10] Th. Coquand and H. Herbelin. A-Translation and Looping Combinators in Pure Type Systems. *Journal of Functional Programming*, 4(1):77–88, 1994.
- [11] M. Dezani-Ciancaglini, P. Severi, and F. J. de Vries. Infinitary Lambda Calculus and Discrimination of Berarducci Trees. *Theoretical Computer Science*, 2(298):275–302, 2003.
- [12] J. Endrullis, C. Grabmayer, J.W. Klop, and V. van Oostrom. On Equal  $\mu$ -Terms. *Theoretical Computer Science*, 412(28):3175–3202, 2011.
- [13] J. Endrullis, D. Hendriks, and J.W. Klop. Modular Construction of Fixed Point Combinators and Clocked Böhm Trees. In *Proc. Symp. on Logic in Computer Science (LICS 2010)*, pages 111–119, 2010.
- [14] H. Geuvers and B. Werner. On the Church–Rosser Property for Expressive Type Systems and its Consequences for their Metatheoretic Study. In *Proc. Symp. on Logic in Computer Science (LICS 1994)*, pages 320–329, 1994.
- [15] M. Goldberg. Constructing Fixed-Point Combinators Using Application Survival. Technical Report BRICS RS-95-35, Dept. of Computer Science, University of Aarhus, 1995.

- [16] B. Intrigila. Non-Existent Statman’s Double Fixed Point Combinator Does Not Exist, Indeed. *Information and Computation*, 137(1):35–40, 1997.
- [17] F. Joachimski. Confluence of the Coinductive  $\lambda$ -Calculus. *Theoretical Computer Science*, 311(1-3):105–119, 2004.
- [18] R. Kennaway and F.-J. de Vries. Infinitary Rewriting. Chapter 12 in [30].
- [19] R. Kennaway, J.W. Klop, M.R. Sleep, and F.-J. de Vries. Infinitary Lambda Calculus. *Theoretic Computer Science*, 175(1):93–125, 1997.
- [20] J.W. Klop. New Fixed Point Combinators from Old. In *Reflections on Type Theory,  $\lambda$ -Calculus, and the Mind. Essays Dedicated to Henk Barendregt on the Occasion of his 60th Birthday*, pages 197–210. 2007. Online version: <http://www.cs.ru.nl/barendregt60>.
- [21] G.D. Plotkin. A Semantics for Type Checking. In *Proc. Conf. on Theoretical Aspects of Computer Software (TACS 1991)*, volume 526 of *LNCS*, pages 1–17. Springer, 1991.
- [22] G.D. Plotkin, 2007. Personal communication at the symposium for H. Barendregt’s 60th birthday.
- [23] D. Sangiorgi and J.J.M.M. Rutten. *Advanced Topics in Bisimulation and Coinduction*, volume 52 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2012.
- [24] D.S. Scott. Some Philosophical Issues Concerning Theories of Combinators. In C. Böhm, editor, *Lambda Calculus and Computer Science Theory*, volume 37 of *LNCS*, pages 346–366, 1975.
- [25] P. Selinger. Order-Incompleteness and Finite Lambda Models. In *Proc. Symp. on Logic in Computer Science (LICS 1996)*, pages 432–439, 1996.
- [26] P. Selinger. *Functionality, Polymorphism, and Concurrency: a Mathematical Investigation of Programming Paradigms*. PhD thesis, University of Pennsylvania, 1997.
- [27] P. Selinger. Order-Incompleteness and Finite Lambda Reduction Models. *Theoretical Computer Science*, 309(1–3):43–63, 2003.
- [28] R. Smullyan. *To Mock a Mockingbird, and Other Logic Puzzles: Including an Amazing Adventure in Combinatory Logic*. Alfred A. Knopf, New York, 1985.
- [29] R. Statman. The Word problem for Smullyan’s Lark Combinator is Decidable. *Journal of Symbolic Computation*, 7:103–112, 1989.
- [30] Terese. *Term Rewriting Systems*, volume 55 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2003.