

# Taming Modal Impredicativity: Superlazy Reduction

Ugo Dal Lago<sup>1</sup>, Luca Roversi<sup>2</sup>, and Luca Vercelli<sup>3</sup>

<sup>1</sup> Dipartimento di Informatica, Università di Bologna  
<http://www.cs.unibo.it/~dallago/>

<sup>2</sup> Dipartimento di Informatica, Università di Torino  
<http://www.di.unito.it/~rover/>

<sup>3</sup> Dipartimento di Matematica, Università di Torino  
<http://www.di.unito.it/~vercelli/>

**Abstract.** Pure, or type-free, Linear Logic proof nets are Turing complete once cut-elimination is considered as computation. We introduce *modal impredicativity* as a new form of impredicativity causing the complexity of cut-elimination to be problematic from a complexity point of view. Modal impredicativity occurs when, during reduction, the conclusion of a residual of a box  $b$  interacts with a node that belongs to the proof net *inside* another residual of  $b$ . Technically speaking, *superlazy reduction* is a new notion of reduction that allows to control modal impredicativity. More specifically, superlazy reduction replicates a box only when all its copies are opened. This makes the overall cost of reducing a proof net finite and predictable. Specifically, superlazy reduction applied to any pure proof nets takes primitive recursive time. Moreover, any primitive recursive function can be computed by a pure proof net via superlazy reduction.

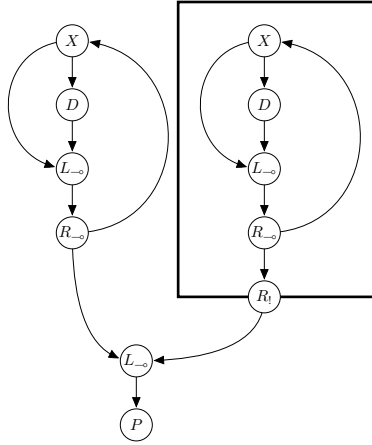
**Keywords.** Linear logic, implicit computational complexity, proof theory.

## 1 Introduction

Predicativity is a logical concept known from a century, starting from Russel's work. It has various technical meanings. All of them, however, refer implicitly or explicitly to some form of aciclicity (see [3] for an excellent survey). Impredicative definitions or logical rules in a logical system may lead to logical paradoxes. On the other hand, if logical systems are interpreted as programming languages (via the Curry-Howard correspondence), impredicativity may lead to type systems and programming languages with high expressive power.

In this paper, we introduce the notion *modal impredicativity*. We start from Linear Logic which gives first-order status to structural rules (on the logical side) and to duplication and erasure (on the computational side). The very definition of modal impredicativity refers to *boxes*, i.e., those portions of proof nets, related to the modal meaning of formulae, that may be duplicated. During cut-elimination, a duplication occurs when a box interacts with a contraction node,

which corresponds to an instance of the structural rule contraction in a logical derivation. Boxes allow to structure proofs into *layers*: any rule instance has a *level*, the number of boxes into which it is contained. Focusing our attention to boxes is the reason why our notion of impredicativity is dubbed as modal. Specifically, modal impredicativity occurs when, during reduction, the *root* of a residual of a box  $b$  interacts with a node that belongs to the residual of the proof net lying *inside*  $b$ . The paradigmatic example where an interaction of this kind occurs is in the (pure) proof net in Figure 1, which encodes the prototypical *non normalizing* lambda-term  $(\lambda x.xx)(\lambda x.xx)$ . Call  $b$  the (unique) box in Figure 1. After two reduction steps we get two copies  $b'$  and  $b''$  of  $b$ . Two further reduction steps plug the root of  $b'$  as premise of the node  $X$  belonging to the second copy  $b''$  of  $b$ . This is a basic form of one-step-long cycle, since the content of  $b$  interacts with the root of  $b$  itself. Compared to what happens classically, *self-copying* plays the rôle of *self-application* or *self-definition*. Notice that the cycles we are speak-



**Fig. 1.** The pure proof net  $\Delta\Delta$ .

ing about can have length greater than one. As an example, consider the proof net corresponding to the lambda term  $(\lambda x.\lambda y.xy x)(\lambda x.\lambda y.xy x)(\lambda x.\lambda y.xy x)$ : it includes two boxes  $b_1$  and  $b_2$  where  $b_1$  copies  $b_2$  and  $b_2$  copies  $b_1$ .

Our long-term goal is to define proper restrictions on Linear Logic allowing to control modal impredicativity. This paper is just the first step towards this goal. What we define here is a new notion of *reduction* for Linear Logic proof nets which rules out the previously described cyclic phenomenon dynamically, i.e. at the level of the graph-theoretic rewriting relation which governs proof net reduction. This way, we break impredicative cycles, while keeping the freedom of *statically* compose pure proof nets.

*Light Logics and Modal Impredicativity.* We now recall how the known subsystems of Linear Logic, introduced as characterizations of certain complexity classes, work. Let us call them light logics, for short. Proof-theoretically, they parsimoniously use the contraction rule. On the computational side, they control the duplication of structure. Technically, we currently know two ways of controlling the use of duplication. One is *stratification*. The other one is what we like to call *boundedness*.

*Stratification* is a structural constraint that, at the dynamic level, has the following meaning: one reduction step at level  $n$  can only increase the complexity of the underlying proof at levels (strictly) higher than  $n$ . This is achieved by dropping dereliction and digging as logical rules. The consequence is the control over the dimension of every single reduct, that reflects on the overall control of reduction time. The mechanism is implicit in the structural and combinatorial properties of proofs and is totally independent from its logical soundness. In stratified systems, the level of any node *cannot* change during reduction. As a consequence, any stratified system, by definition, cannot be modally impredicative because the nodes inside a box  $b$  cannot interact with the root of any copy of  $b$ . Elementary Linear Logic, Light Linear Logic [6] and their affine versions use *stratification*.

Concerning boundedness, recall that in ordinary Linear Logic,  $!A$  is semantically equivalent to  $A^* = \bigcup_{n \in \mathbb{N}} \underbrace{(A \otimes \dots \otimes A)}_n$ . *Boundedness* refers to various

methodologies that, informally, put  $!A$  in correspondence to a *finite* subset of  $A^*$ . Computationally, this means the number of copies of each box in a proof can somehow be statically or dynamically predicted, i.e. bounded. This way, we automatically get a system that cannot be modally impredicative, since in bounded systems  $!A$  cannot be equal to  $!A \otimes A$  (and this principle seems necessary to have self-copying). Soft Linear Logic [8] and Bounded Linear Logic [4] use *boundedness*.

*Superlazy Reduction and Modal Impredicativity.* *Superlazy reduction* is a new notion of reduction for Linear Logic (pure) proof nets. It is specifically designed to control modal impredicativity. Under superlazy reduction, any box  $b$  can interact with a tree of contraction, dereliction, digging and weakening nodes only if the global result of this interaction somehow reduces the overall complexity of the proof net, namely when it produces (possibly many) “open” copies of  $b$ . If this is not the case, reduction is blocked and cannot be performed. This way modal impredicativity is automatically ruled out, since whenever the *content* of  $b$  is copied,  $b$  as a box is destroyed and no residuals of  $b$  are produced. Technically, this is ensured by prescribing that reduction can happen only when the box is faced with a *derelicting tree of nodes*, a key notion introduced in Section 3.

*Superlazy Reduction and Primitive Recursion.* The calculus we obtain by adopting superlazy reduction over pure proof nets is still powerful enough to characterize the class PR of primitive recursive functions. We show the characterization under a standard pattern. As for *soundness*, we prove that every pure proof net

$G$  can be rewritten to its normal form in time bounded by a primitive recursive function in the size of  $G$ . This is remarkable by itself, since the mere fact that superlazy reduction *computes something* is interesting by itself, considered the strong requirements superlazy reduction must satisfy. As for *completeness*, every function in PR can be represented as a pure proof net, even under superlazy reduction.

*Superlazy Reduction and Expressive Power.* We here want to make some observations about pure proof nets and superlazy reduction as a paradigmatic programming language. The set of terms we can program with are pure proof nets coming from Linear Logic. Namely, we can use every lambda-term as a program,  $\Delta\Delta$  (Figure 1) included. However, we do not have the standard unconstrained reduction steps, which, by simulating the usual beta-reduction, allow to embed pure, i.e. untyped, lambda-calculus into pure proof nets. In particular, the proof net  $\Delta\Delta$  is normal in our setting, the reason being that we can never reach a situation where the “amount” of open copies of the box it contains is known in advance.

*Related Work.* Several authors have used some notion of predicativity as a way to restrict the expressive power of programming languages or logical systems. We here recall some of them, without any hope of being exhaustive. In [11], as a first example, Danner and Leivant presented a variant of second order  $\lambda$ -calculus obtained imposing a restriction on second order quantification. Such a restriction has semantic flavor: all the *types* have a *rank*, an ordinal number and universal quantification can only be instantiated if the witness has a proper rank. This way they get a characterization of primitive recursion. Another, earlier, example is Leivant’s predicative recursion [9]: if predicativity is imposed on ordinary primitive recursion (on any word algebra), one gets a characterization of polynomial time computable functions. Further work shows how other classes can be characterized with similar tools [10, 12]. A final example is Simmons’ fine analysis of tiering [14]. All the cited proposals, however, share a property which make them fundamentally different from superlazy reduction: predicativity is enforced through static constraints, i.e., constraints on programs rather than on the underlying reduction relation. The first author has recently proposed a characterization of primitive recursion by a fragment of Gödel’s T [1].

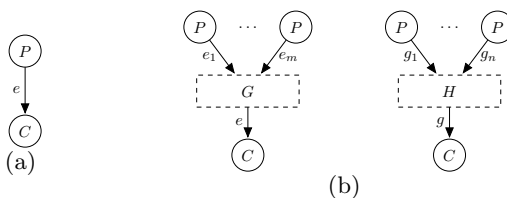
On the other hand, restricted notions of reduction on Linear Logic proof nets have appeared in the literature. This includes, for example, Girard’s closed reduction [5] or head linear reduction [13]. None of them, however, decreases the expressive power of the logical systems on top of which they are applied, as superlazy reduction does.

*Paper Outline.* Section 2 recalls pure, i.e., untyped, proof nets and defines derelecting trees and superlazy reduction. Section 3 proves the primitive recursive soundness of superlazy reduction on pure proof nets. Section 4 shows that even under superlazy reduction, pure proof nets remain expressive enough to represent all the primitive recursive functions. Section 5 presents some further devel-

opments on the ideas presented in the paper. An extended version of this paper including all proofs is available [2].

## 2 Pure Proof Nets

Pure proof nets are graph-like structures corresponding to proofs. The set of labels for the nodes is  $\{P, C, W, X, R_{\rightarrow}, L_{\rightarrow}, R_{\otimes}, L_{\otimes}, R_!, L_!, D, N\}$ . All nodes, but  $P$  and  $C$ , correspond to the usual proof net labels. We use  $P$  and  $C$  for the sake of uniformity, getting graphs without dangling edges. Figure 2(a) says that a wire is a proof net. Given the two proof nets in 2(b) we can build those in Figure 3. The inductive rule at the end of Figure 2(a) introduces (modal) *boxes*.

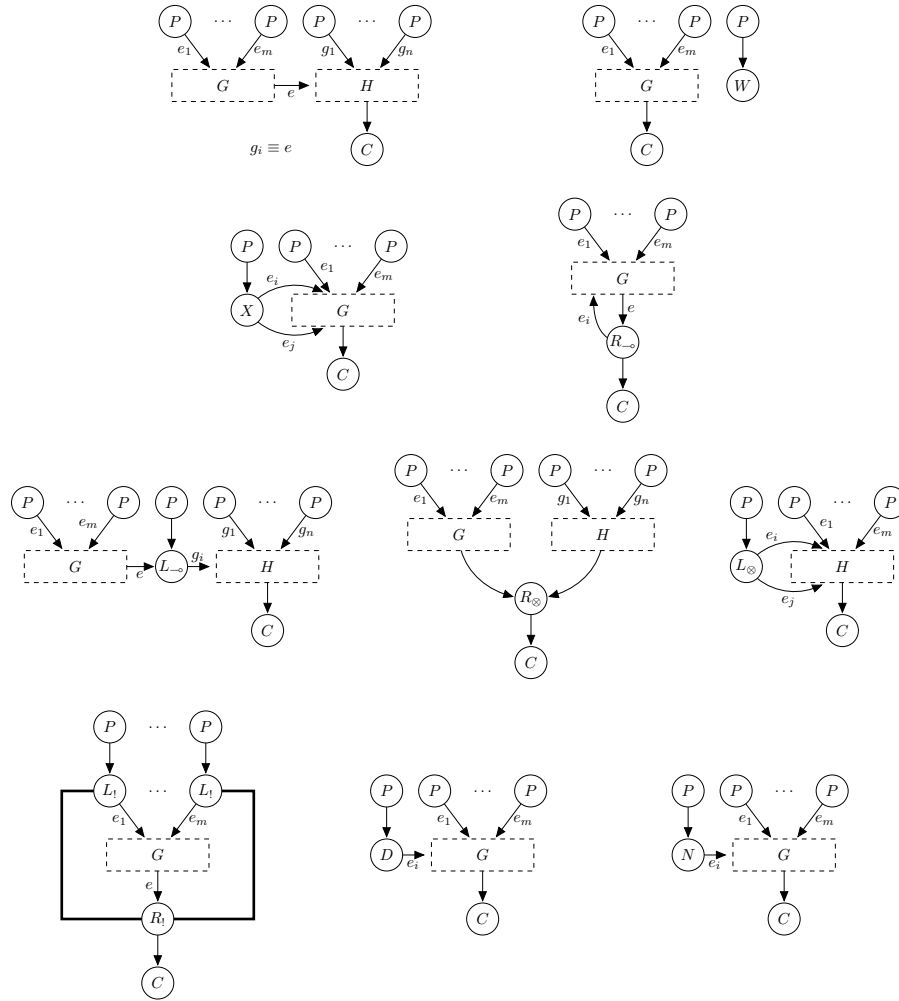


**Fig. 2.** Base cases.

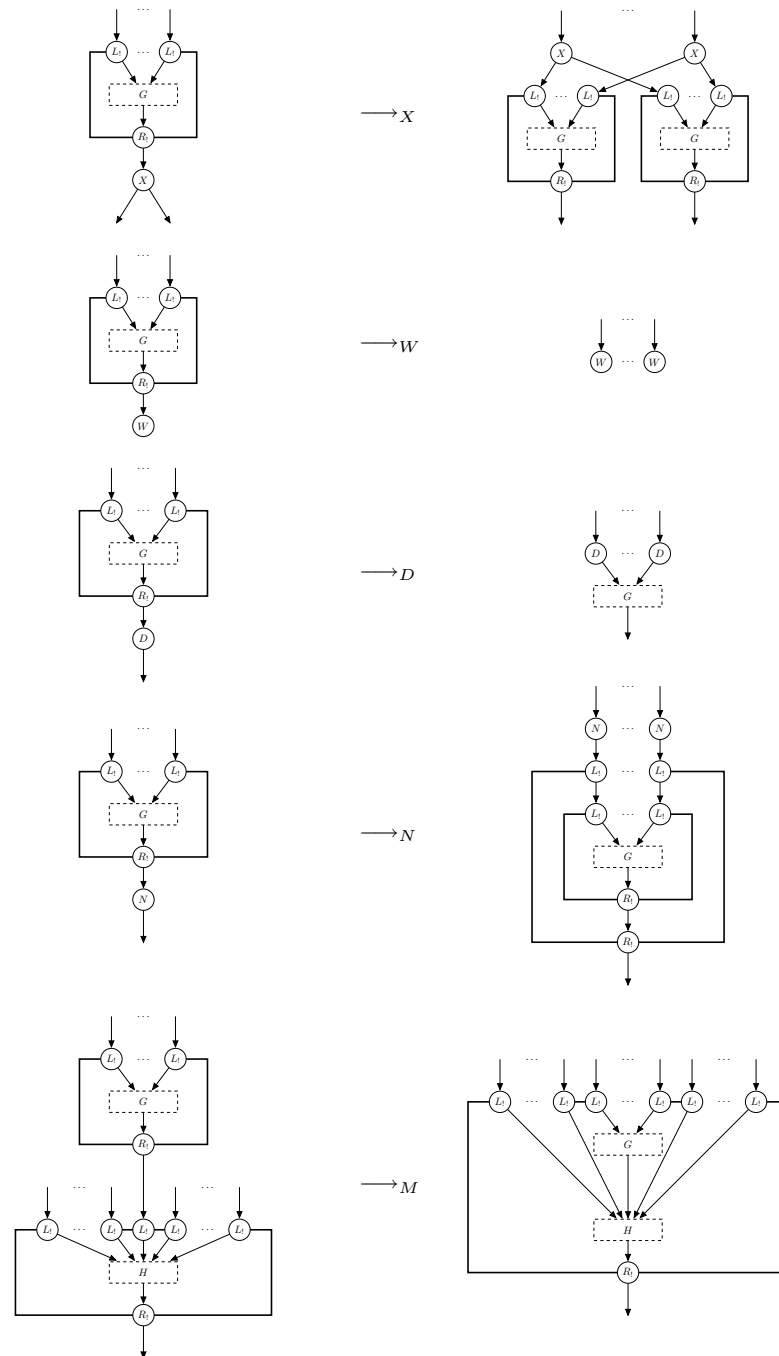
Please notice that the proof nets introduced here are slightly different from the usual ones. In particular, there is not any explicit node playing the rôle of the cut rule or of axioms. Moreover, proof net conclusions are partitioned into one proper conclusion and some premises. This way, proof nets get an intuitionistic flavor which makes the correspondence with lambda-terms more evident.

*Reduction Rules.* The reduction rules for the proof nets are the usual ones. We omit the obvious linear rules  $\rightarrow_{\rightarrow}$ ,  $\rightarrow_{\otimes}$ , and we just recall the *modal* reduction rules in Figure 4. Call  $\rightarrow$  the contextual closure of the rewriting steps  $\rightarrow_{\rightarrow}$ ,  $\rightarrow_{\otimes}$ ,  $\rightarrow_D$ ,  $\rightarrow_X$ ,  $\rightarrow_W$ ,  $\rightarrow_N$ ,  $\rightarrow_M$ . The reflexive and transitive closure of  $\rightarrow$  is  $\rightarrow^*$ .

The reduction rules  $\rightarrow_X$  and  $\rightarrow_N$  are the only ones somehow increasing the size of the underlying proof-net: the first one copies a box, while the second one puts a box inside another box. Superlazy reduction, as we will see shortly, does not simply eliminate those rules, but rather forces them to be applicable only in certain contexts, i.e., only when those rules are part of a sequence of modal rewriting rules which have a globally predictable behavior.



**Fig. 3.** Inductive cases



**Fig. 4.** Modal graph rewriting rules.

## 2.1 Superlazy Reduction

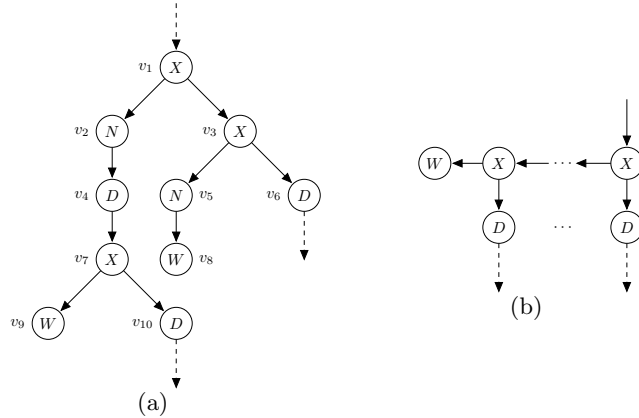
We shall be able to prove a soundness result about the cost of the reduction of the proof nets relatively to a *superlazy* version of  $\longrightarrow$  that requires the notion of *derelicting tree*.

*Derelicting trees.* For every proof net  $G$ , let us assume that the *cost* of traversing any  $X$ -node of  $G$  is 0, any  $D$ -node is  $-1$ , any  $W$ -node is 0 and any  $N$ -node is  $+1$ . The cost of a path from node  $u$  to node  $v$  is the sum of the costs of nodes in the path *including*  $u$  and  $v$ . A *derelicting tree* in  $G$  is a subgraph  $t$  of  $G$  that satisfies the following four conditions:

1.  $t$  only contains nodes  $X, D, N, W$ ; so it must be a tree, and we call  $w$  its root;
2. the leaves of  $t$  are labelled either with  $D$  or with  $W$ ;
3. for every leaf  $v$  labelled with  $D$  in  $t$ , the cost of the path from  $w$  to  $v$  in  $t$  is  $-1$ ;
4. the cost of any other path in  $t$  starting from  $w$  is nonnegative.

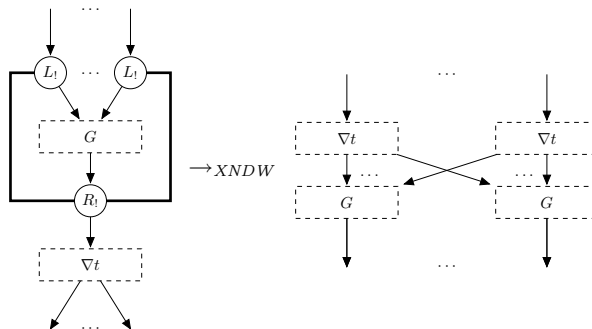
Figure 5(a) shows an example of a derelicting tree. Conditions 1. and 2. are trivially satisfied. The cost of  $v_1v_2v_4v_7v_{10}$  is  $-1$  and the same for  $v_1v_3v_6$ . Finally, any other path starting from  $v_1$  has nonnegative cost. For example,  $v_1v_3v_5v_8$  has cost 1.

Figure 5(b) depicts a remarkable instance of derelicting tree: an *n-bounded spine* with  $n$  occurrences of  $X$  nodes that we shall represent as a dashed box with name  $nX$ .



**Fig. 5.** A generic derelicting tree (a) and a bounded spine (b)

*Superlazy Reduction Step and Rewriting System on Pure Proof Nets.* The *superlazy normalization step* is  $\rightarrow_{XNDW}$ , defined in Figure 6,  $\nabla t$  being any derelicting tree.  $\rightsquigarrow_{\rightarrow}, \rightsquigarrow_{\otimes}, \rightsquigarrow_M, \rightsquigarrow_{XNDW}$  denote the *surface* contextual closure of the



**Fig. 6.** Superlazy cut elimination step

rewriting steps  $\longrightarrow_{\rightarrow}, \longrightarrow_{\otimes}, \longrightarrow_M, \longrightarrow_{XNDW}$ . “Surface” means that we never apply a reduction inside a box.  $\sim$  is the union of  $\sim_{\rightarrow}, \sim_{\otimes}, \sim_M, \sim_{XNDW}$ . The reflexive and transitive closure of  $\sim$  is  $\sim^*$ .

Superlazy reduction is a *very* restricted notion of reduction. In particular, it is almost useless when applied to proof nets obtained from ordinary lambda-terms via the usual, uniform encoding. In particular:

- If terms are encoded via the so-called call-by-name encoding (i.e., the one induced by Girard’s correspondence  $(A \rightarrow B)^\circ \equiv !A^\circ \multimap B^\circ$ ), then any redex  $(\lambda x.M)N$  where  $M$  consists of an application  $LP$  and  $x$  appears free in  $P$ , cannot be reduced in the corresponding proof net: a box would be faced with something different from a derelicting tree.
- On the other hand, if terms are encoded via the call-by-value encoding (i.e. via the correspondence  $(A \rightarrow B)^\circ \equiv !(A^\circ \multimap B^\circ)$ ), then any redex  $(\lambda x.M)N$  where  $M$  consists itself of an abstraction  $\lambda y.L$  and  $x$  appears free in  $L$  cannot be reduced in the corresponding proof net.

Unfortunately, we do not even know any criteria allowing to guarantee that certain proof nets can be reduced to normal form (w.r.t. ordinary reduction) by way of superlazy reduction. Moreover, there currently isn’t any result characterizing the class of normal forms w.r.t. superlazy reduction; this is in contrast, for example, to lambda calculus and call-by-value reduction, where the cbv normal form of any (closed) term  $M$  (if any) is always a value. This is why proving that proof nets are complete w.r.t. some given class of functions, under the superlazy reduction, is non-trivial. Nonetheless, we explicitly prove the completeness of superlazy reduction (see Section 4). The next section shows why both  $(A \rightarrow B)^\circ \equiv !A^\circ \multimap B^\circ$  and  $(A \rightarrow B)^\circ \equiv !(A^\circ \multimap B^\circ)$  do not work well here: superlazy reduction of any proof net always terminates in a time bounded by suitable primitive recursive functions.

Linear Logic with superlazy reduction can be seen as a generalization of the principles of Soft Linear Logic. Every time a box is replicated, that box is opened. According to this vision, a derelicting tree with  $m$  leaves is similar to a multiplexor node of rank  $m$ . However, the structure of SLL gives a further

restriction: if  $k$  is the rank of the proof net  $G$ , that is the maximum rank of the multiplexers in  $G$ , we are sure that every box of  $G$  will be copied at most  $k$  times. Such a restriction leads to a polytime bound (see [8]).

### 3 Soundness

We prove the result starting with a restriction  $\hookrightarrow$  on the relation  $\rightsquigarrow$ .  $\hookrightarrow$  is simply the union of  $\rightsquigarrow_{\rightarrow}$ ,  $\rightsquigarrow_{\otimes}$ ,  $\rightsquigarrow_M$  and  $\rightsquigarrow_{XNDW}$ , where any  $\rightsquigarrow_M$  can only be applied to nets which only contain  $\rightsquigarrow_M$  redexes. In other words,  $\rightsquigarrow_M$  is postponed as much as possible. This makes our arguments simpler without loss of generality, as we now show.  $\hookrightarrow^k$  will denote a reduction with  $k$  steps of  $\hookrightarrow$ . Given a proof  $G$  and any reduction relation  $\rightarrow$ ,  $[G]_{\rightarrow}$  and  $\|G\|_{\rightarrow}$  denote the maximum length of a reduction sequence starting in  $G$  (under  $\rightarrow$ ) and the maximum size of any reduct of  $G$  (under  $\rightarrow$ ), respectively.  $|G|$  is the size of the proof net  $G$ .

**Lemma 1.** *For every proof  $G$ ,  $[G]_{\rightsquigarrow} = [G]_{\hookrightarrow}$  and  $\|G\|_{\rightsquigarrow} = \|G\|_{\hookrightarrow}$ .*

*Proof.* Whenever  $G \rightsquigarrow_M F \rightsquigarrow_x H$  and  $x \neq M$ , there are  $F_1, \dots, F_n$  (where  $n \geq 1$ ) such that  $G \rightsquigarrow_{x_1} F_1 \rightsquigarrow_{x_2} \dots \rightsquigarrow_{x_n} F_n \rightsquigarrow_{x_{n+1}} H$ , and  $x_{i+1} = M$  whenever  $x_i = M$ . For example, if  $G \rightsquigarrow_M F \rightsquigarrow_{XNDW} H$  and the box copied in the second step is exactly the one created by the first step, then clearly

$$G \rightsquigarrow_{XNDW} J \rightsquigarrow_{XNDW} H.$$

As a consequence, for any sequence  $G_1 \rightsquigarrow \dots \rightsquigarrow G_n$  there is another sequence  $F_1 \hookrightarrow \dots \hookrightarrow F_m$  such that  $G_1 = F_1$ ,  $G_n = F_m$  and  $m \geq n$ . This proves the first claim. Now, observe that for any  $1 \leq i \leq n$  there is  $j$  such that  $|F_j| \geq |G_i|$ : a simple case analysis suffices. This concludes the proof.  $\square$

The following definition is the main ingredient since it allows to structure the proof of soundness inductively. A reduction sequence  $\sigma \equiv G \rightsquigarrow_{x_1} F \rightsquigarrow_{x_2} \dots \rightsquigarrow_{x_k} H$  is said to be a  $(n, d)$ -box reduction when  $x_i \neq M$  for every  $1 \leq i \leq k$  and there are  $r \leq n$  boxes  $b_1, \dots, b_r$  between those at level 0 in  $G$  such that the following conditions hold (let  $J_1, \dots, J_r$  be the proof nets inside  $b_1, \dots, b_r$ , respectively):

- $\partial(J_1), \dots, \partial(J_r) < d$ .
- The box involved in any step  $\rightsquigarrow_{XNDW}$  of  $\sigma$  is either a residual of  $b_1, \dots, b_r$  or a residual of a box appearing in one of  $J_1, \dots, J_r$ .

*Remark 1.* A few observations about the above definition:

1.  $G$  may contain more than  $n$  boxes at level 0, or it may contain boxes whose depth is greater than  $d$ ;
2.  $H$  is *not* necessarily a normal form. It may contain boxes at level  $d$ , or higher;
3. By definition, any node inside the boxes  $b_1, \dots, b_r$  *must have* depth at least 1. This means that if  $\sigma$  is a  $(n, 0)$ -box reduction it is, in fact, talking about boxes that contain proof nets with negative depth. They cannot exist, so  $\sigma$  only uses the linear rewriting steps  $\longrightarrow_{\rightarrow}$ ,  $\longrightarrow_{\otimes}$ ;

4. By definition, *every* reduction starting at any net  $G$  is a  $(|G|, \partial(G))$ -box reduction because  $|G|$  necessarily bounds the number  $n$  of boxes, and  $\partial(G)$  bounds the value  $d$ ;
5. The definition is a good one *only* because we are assuming to work without the rule  $\rightarrow_M$ . Otherwise (residuals of) boxes coming from  $b_1, \dots, b_r$  could be “merged” with (residuals of) boxes in  $G$  but not in the list.

Now we define a family of functions  $\{f_d : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}\}_{d \in \mathbb{N}}$  such that every  $f_d(n, m)$  bounds *both* the reduction cost *and* the size of the reducts, of every net  $G$  when performing a  $(n, d)$ -box reduction on it:

$$\begin{aligned} f_0(n, m) &= m \\ f_{d+1}(0, m) &= m \\ f_{d+1}(n+1, m) &= 1 + f_{d+1}(n, m) + f_d((2m+1)f_{d+1}(n, m), (2m+1)f_{d+1}(n, m)) \end{aligned}$$

By definition, all the functions  $f_d$  are primitive recursive.

**Lemma 2.** *For every  $d, n, m \in \mathbb{N}$ ,  $f_{d+1}(n, m) \geq f_d(n, m)$  and  $f_d(n+1, m) \geq f_d(n, m)$ .*

The proof is by induction on  $n$ . We can prove the following:

**Proposition 1.** *Let  $G$  be any proof net and let  $G \hookrightarrow^k H$  be a  $(n, d)$ -box reduction sequence with  $k$  steps. Then  $k, |H| \leq f_d(n, |G|)$ .*

*Proof.* We write  $m = |G|$ . We can proceed by induction on  $d$ :

- If  $d = 0$ , thanks to Remark 1.3 above, the reduction  $G \hookrightarrow^k H$  only involves multiplicative reduction steps. They strictly reduce the size of the underlying net. So,  $|H|, k \leq m = f_0(n, m)$ .
- Suppose the thesis holds for  $d$  and suppose  $G \hookrightarrow^k H$  is a  $(n, d+1)$ -box reduction. We proceed by another induction, this time on  $n$ :
  - If  $n = 0$ , then none of the boxes at level  $d+1$  can be reduced. As a consequence,  $G \hookrightarrow^k H$  only involves multiplicative reduction steps and, again,

$$|H|, k \leq m = f_{d+1}(0, m).$$

- Suppose the thesis holds for  $n$  and suppose  $G \hookrightarrow^k H$  is a  $n+1$ -box reduction sequence at level  $d+1$ . By definition, there are  $r \leq n+1$  boxes  $b_1, \dots, b_r$  in  $G$  satisfying the definition above. There is clearly one index  $t$ , where  $1 \leq t \leq r$  such that  $b_t$  is the last box being copied (among  $b_1, \dots, b_r$ ) in the sequence  $G \hookrightarrow^k H$ . Up to the point where  $b_t$  is copied, the reduction sequence can be considered as a  $(n, d+1)$ -box reduction sequence: the witness list of box is exactly  $b_1, \dots, b_{t-1}, b_{t+1}, \dots, b_r$ . After  $b_t$  is copied, on the other hand, the reduction sequence under consideration can be considered as a  $(|J|, d)$ -box reduction sequence, since all the boxes that will be copied are part of the residual of the content of boxes in the list  $b_1, \dots, b_r$ . The reduction under consideration can be decomposed as follows

$$G \hookrightarrow^i F \rightsquigarrow_{XNDW} J \hookrightarrow^j H$$

where  $G \hookrightarrow^i F$  is an  $(n, d+1)$ -box reduction sequence, the step  $F \rightsquigarrow_{XNDW} J$  involves *exactly*  $b_t$  and  $J \hookrightarrow^j H$  is a  $(|J|, d)$ -box reduction. Now, observe that  $|b_t| \leq m$ . In  $|J|$  we will find at most  $|F|$  copies of  $b_t$  and at most  $|b_t|$  copies of the underlying derelicting tree (which can itself contain at most  $|F|$  nodes). Applying both inductive hypothesis, we get

$$\begin{aligned} i, |F| &\leq f_{d+1}(n, m) \\ |J| &\leq 2mf_{d+1}(n, m) + |F| \\ j, |H| &\leq f_d(|J|, |J|) \end{aligned}$$

As a consequence:

$$\begin{aligned} k = i + 1 + j &\leq f_{d+1}(n, m) + 1 + f_d(|J|, |J|) \\ &\leq f_{d+1}(n, m) + 1 + f_d((2m+1)f_{d+1}(n, m), (2m+1)f_{d+1}(n, m)) \\ |H| &\leq f_d(|J|, |J|) \leq f_{d+1}(n, m) + 1 + f_d(|J|, |J|) \\ &\leq f_{d+1}(n, m) + 1 + f_d((2m+1)f_{d+1}(n, m), (2m+1)f_{d+1}(n, m)) \end{aligned}$$

This concludes the proof.  $\square$

**Corollary 1 (Soundness).** *For every  $n \in \mathbb{N}$  there is a primitive recursive function  $g_n : \mathbb{N} \rightarrow \mathbb{N}$  such that for every proof net  $G$ ,  $[G]_{\rightsquigarrow}, ||G||_{\rightsquigarrow} \leq g_n(|G|)$ .*

*Proof.* By Lemma 1, we can bound  $[G]_{\hookrightarrow}$  and  $||G||_{\hookrightarrow}$  rather than  $[G]_{\rightsquigarrow}$  and  $||G||_{\rightsquigarrow}$ . Now, observe that any reduction  $G \hookrightarrow^k H$  is a  $(|G|, \partial(G))$ -box reduction, followed by at most a linear number of  $M$ -reduction steps, which anyway shrinks the size of the underlying proof net. The thesis follows from Proposition 1.  $\square$

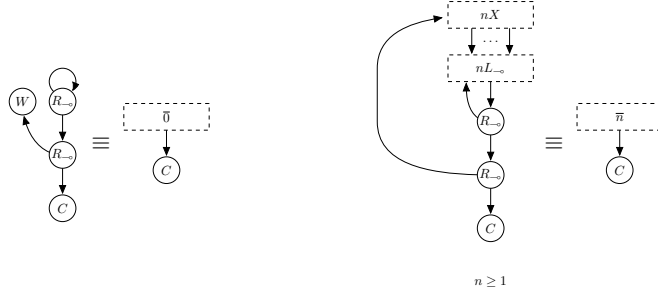
## 4 Completeness

The goal is to show the existence of an embedding of any primitive recursive function  $f$  into a pure proof net  $G_f$  such that  $G_f$  simulates  $f$  via superlazy reduction. The existence of such an embedding is what we mean by *completeness*.

Our comments at the end of Section 2 should have convinced the reader about the impossibility of proving completeness by the usual encoding of the untyped lambda-calculus into pure proof nets. We really need to tailor the encoding of data and programs in such a way that superlazy reduction *works* on them, i.e., we want to be sure that a program applied to an argument *reduces* to the intended result.

The details of the completeness proof do not fit here, but they can be found in [2], where we also develop a slightly more general proof of soundness. Only the major ingredients towards completeness are reported here.

The first ingredient is the representation of natural numbers. Figure 7 introduces the closed proof nets  $\bar{0}$  and  $\bar{n}$ , with  $n \geq 1$ . A basic observation to prove the completeness is that every  $\bar{n}$  contains a  $n$ -bounded spine  $nX$  (see Fig. 5(b))



**Fig. 7.** Church numerals as proof nets

). This implies that whenever  $\bar{n}$  is applied to a box containing a proof net  $G$ , superlazy reduction will copy the box  $n$  times and  $n$  copies of  $G$  will appear, one applied to the next one. Computing the successor of a natural number can be done by a proof net, quite similarly to what happens in ordinary lambda calculus. Now, at least, the key notion can be given. Let  $G$  be a proof net with a single premise and a single conclusion. We write  $G \smile \langle \bar{n}_1, \dots, \bar{n}_m \rangle$  to mean the closed proof net with a single conclusion obtained by plugging the conclusion of the net  $\langle \bar{n}_1, \dots, \bar{n}_m \rangle$  (obtained by “tensoring” together  $\bar{n}_1, \dots, \bar{n}_m$ ) into the unique assumption of  $G$ . For every function  $f : \mathbb{N}^m \rightarrow \mathbb{N}$ , with arity  $m \geq 0$ , we shall say that a proof net  $G_f$  with one premise *represents*  $f$  iff  $G_f \smile \langle \bar{n}_1, \dots, \bar{n}_m \rangle$  *superlazily* reduces to  $f(n_1, \dots, n_m)$ , for every  $n_1, \dots, n_m \in \mathbb{N}$ .

The second ingredient is the possibility of freely duplicate data, i.e., natural numbers. This is possible even if the natural number being copied (or erased) does not lie inside a box. Copying a natural number  $n$  involves applying a (boxed) pair of successors and a pair of  $\bar{0}$  to the net  $\bar{n}$ . Erasing  $n$ , on the other hand, can be performed by applying a boxed identity and another boxed identity to  $\bar{n}$ : this way  $\bar{n}$  can be reduced itself to a boxed identity, which can be erased by cutting it against a  $W$  node.

The first two ingredients allow to encode basic primitive recursive functions and composition. To get the most important construction, namely primitive recursion itself, a third ingredient is necessary, namely iteration. Iterating  $n$  times a given function  $f$ , where  $\bar{n}$  is a parameter, can be done by way of superlazy reduction: putting the proof net representing  $f$  inside a box and apply the box to  $\bar{n}$  suffices. However, if the proof net representing  $f$  has some premises, the resulting proof net would only accept boxed natural numbers as arguments, and this would break the scheme which makes superlazy reduction works. The solution consists in iterating only closed functions, exploiting the higher-order nature of proof nets. The usual primitive recursion scheme can be finally obtained by using the standard technique of simulating recursion by iteration.

**Theorem 1 (Completeness).** *Every  $f$  in PR is represented by a proof net  $G_f$ .*

## 5 Further Developments

As we explained in the Introduction, this paper is just the first step in a long-term study about how to control modal impredicativity.

There are at least two distinct research directions the authors are following at the time of writing. We recall the obtained results here, pointing to further work for additional details and proofs.

First of all, the way we proved completeness of proof net reduction w.r.t. primitive recursion suggests a way of capturing Hofmann’s non-size increasing computation [7] by way of a proper, further restriction to superlazy reduction. We basically need two constraints:

- Boxes can only be copied when they are closed, i.e., when they have no premises.
- Any box  $b$  can only be copied if the proof net contained in  $b$  contains at most *one* node  $X$  at level 0.

The obtained reduction relation is said to be the *non-size increasing superlazy reduction* and is denoted with  $\Rightarrow$ . With these constraints, non-size increasing polytime computation can be simulated by pure proof nets. Moreover:

**Theorem 2.** *For every  $n \in \mathbb{N}$  there is a polynomial  $p_n : \mathbb{N} \rightarrow \mathbb{N}$  such that for every proof net  $\Pi$ ,  $[G]_{\Rightarrow}, \|G\|_{\Rightarrow} \leq p_{\partial(G)}(|G|)$*

The second direction we are considering concerns methods to keep modal impredicativity under control by more traditional static methods in the spirit of light logics. Consider the equivalence between  $!A$  and  $!A \otimes A$ . As already pointed out, this equivalence is somehow necessary to get modal impredicativity. So, controlling it means controlling modal impredicativity. Now, suppose that the above equivalence holds *but*  $A$  only contains instances of the  $!$  operator which are *intrinsically different* from the top-level one in  $!A$ . Morally, this would imply that even if a contraction node is “in”  $A$ , it cannot communicate with the box in  $!A$ , because they are of a different nature. This way modal impredicativity would be under control. But the question is: how to distinguish different instances of  $!$  from each other? One (naive) answer is the following: consider a generalization of (multiplicative and exponential) Linear Logic where syntactically different copies  $!_{a_1}, !_{a_2}, \dots$  of the modal operator  $!$  are present. As an example, take the set  $\{!_n\}_{n \in \mathbb{N}}$ . Then, impose the following constraint:  $!_a A$  is a *legal* formula only if the operators  $!_{b_1}, \dots, !_{b_n}$  appearing in  $A$  are all different from  $a$ , i.e., if  $a \notin \{b_1, \dots, b_n\}$ . We strongly believe that this way a system enjoying properties similar to those of predicative recurrence schemes [9] can be obtained.

## 6 Conclusions

We described modal impredicativity and a concrete tool — superlazy reduction — that controls it. Superlazy reduction on pure proof nets greatly influences the expressive power of the underlying computational model: from a Turing complete model we go down to first-order primitive recursive functions.

In a sentence, we learnt that the expressive power of a programming language system can be controlled by acting on the dynamics (i.e., the underlying reduction relation) without touching the statics (i.e., the language into which programs are written). To get the complete picture, however, we still need tools to predict which set of programs will be useful from a computational point of view.

This could have potential applications in the field of implicit computational complexity, where one aims at designing programming languages and logical systems corresponding to complexity classes. Indeed, the impact of ICC in applications crucially depends on the intensional expressivity of the proposed systems: one should be able to write programs naturally.

## References

1. Ugo Dal Lago. The geometry of linear higher-order recursion. In *Proc. 20th Annual Symposium on Logic in Computer Science*, pages 366–375. IEEE Computer Society, 2005.
2. Ugo Dal Lago, Luca Roversi, and Luca Vercelli. Taming modal impredicativity: Superlazy reduction. Extended Version. Available at <http://arxiv.org/abs/0810.2891>, 2008.
3. Solomon Feferman. Predicativity. In S. Shapiro, editor, *The Oxford Handbook of the Philosophy of Mathematics and Logic*, pages 590–624. Oxford University Press, 2005.
4. J.-Y. Girard, A. Scedrov, and P. Scott. Bounded linear logic: A modular approach to polynomial time computability. *Theor. Comput. Sci.*, 97:1–66, 1992.
5. Jean-Yves Girard. Geometry of interaction 1: interpretation of system F. In *Proc. Logic Colloquium '88*, pages 221–260, 1989.
6. Jean-Yves Girard. Light linear logic. *Information and Computation*, 143(2):175–204, 1998.
7. Martin Hofmann. Linear types and non-size-increasing polynomial time computation. In *Proceedings of the 14th Symposium on Logic in Computer Science*, pages 464–473. IEEE Computer Society, 1999.
8. Yves Lafont. Soft linear logic and polynomial time. *Theor. Comput. Sci.*, 318:163–180, 2004.
9. Daniel Leivant. Stratified functional programs and computational complexity. In *Proceedings of 20th ACM Symposium on Principles of Programming Languages*, pages 325–333, 1993.
10. Daniel Leivant. Ramified recurrence and computational complexity III: Higher type recurrence and elementary complexity. *Ann. Pure Appl. Logic*, 96(1-3):209–229, 1999.
11. Daniel Leivant and Norman Danner. Stratified polymorphism and primitive recursion. *Math. Struct. in Comp. Science*, 9:507 — 522, 1999.
12. Daniel Leivant and Jean-Yves Marion. A characterization of alternating log time by ramified recurrence. *Theor. Comput. Sci.*, 236(1-2):193–208, 2000.
13. Gianfranco Mascari and Marco Pedicini. Head linear reduction and pure proof net extraction. *Theor. Comput. Sci.*, 135(1):111–137, 1994.
14. Harold Simmons. Tiering as a recursion technique. *Bulletin of Symbolic Logic*, 11(3):321–350, 2005.