

# Adding recursion to DPI (Extended abstract)

Samuel Hym

*PPS, Université Paris 7 & CNRS*

Matthew Hennessy

*Department of Informatics, University of Sussex*

---

## Abstract

DPI is a distributed version of the  $\pi$ -CALCULUS, in which processes are explicitly located, and a migration construct may be used for moving between locations. We argue that adding a recursion operator to the language increases significantly its descriptive power. But typing recursive processes requires the use of potentially infinite types.

We show that the capability-based typing system of DPI can be extended to *co-inductive types* so that recursive processes can be successfully supported. We also show that, as in the  $\pi$ -CALCULUS, recursion can be implemented via iteration. This translation improves on the standard ones by being compositional but still comes with a significant migration overhead in our distributed setting.

---

## 1 Introduction

The  $\pi$ -CALCULUS, [7,8], is a well-known formal calculus for describing, and reasoning about, the behaviour of concurrent processes which interact via two-way communication channels. DPI, [4], is one of a number of extensions in which processes are *located*, and may migrate between locations, or sites, by executing an explicit migrate command; the agent `goto  $k$ .P`, executing at a site  $l$ , will continue with the execution of  $P$  at the site  $k$ . This extension comes equipped with a sophisticated capability-based type system, and a co-inductive behavioural theory which takes into account the constraints imposed by these types, [3,4]. The types informally correspond to sets of *capabilities*, and the use a process may make of an entity, such as a location or a channel, depends on the current type at which the process owns the entity. Moreover this type may change over time, reflecting the fact that processes may gradually accumulate capabilities over entities.

*This is a preliminary version. The final version will be published in  
Electronic Notes in Theoretical Computer Science  
URL: [www.elsevier.nl/locate/entcs](http://www.elsevier.nl/locate/entcs)*

The most common formulations of the PI-CALCULUS use iteration (written  $* P$  for the iteration of  $P$ ) to describe repetitive processes. Thus

$$* c?(x) d!\langle x \rangle$$

represents a process which repeatedly inputs a value on channel  $c$  and outputs it at  $d$ . An alternative would be to use an explicit recursion operator, leading to definitions such as

$$\text{rec } Z. c?(x) d!\langle x \rangle Z$$

But it has been argued that explicit recursion is unnecessary, because it offers no extra convenience over iteration; indeed it is well-known that such a recursion operator can easily be implemented using iteration; see pages 132–138 in [8].

However the situation changes when we move to the distributed world of DPI. In Section 2 we demonstrate that the addition of explicit recursion leads to powerful programming techniques; in particular it leads to simple natural descriptions of processes for searching the underlying network for sites with particular properties.

Unfortunately this increase in descriptive power is obtained at a price. In order for these recursive processes to be accommodated within the typed framework of DPI, we need to extend the type system with *co-inductive types*, that is types of potentially infinite depth.

The purpose of this paper is to

- demonstrate the descriptive power of recursion when added to DPI;
- develop a system of co-inductive types which support recursive processes;
- prove that at the cost of significant migration costs recursion in DPI can still be implemented by purely iterative processes, in the absence of network failures.

In Section 2 we describe the extension to DPI, called RECDPI, and demonstrate the power of recursion by a series of prototypical examples. This is followed in Section 3 with an outline of how the co-inductive types are defined, and how the typing system for DPI can be easily extended to handle these new types. The translation of recursive processes into iterative processes is explained in Section 4, and we outline the proof of correctness in Section 5. This requires the use of a *typed* bisimulation equivalence to accommodate the typed labelled transition system for RECDPI.

The paper relies heavily on existing work on DPI, and the reader is referred to papers such as [3,4] for detailed explanations of both the semantics of DPI and its typing system.

**Fig. 1** Syntax of RECDPI

---

$M, N ::=$	<i>Systems</i>
$l[[P]]$	Located Process
$M   N$	Composition
$(\text{new } e : E) M$	Name Creation
<b><math>\mathbf{0}</math></b>	Termination
$P, Q ::=$	<i>Processes</i>
$u!\langle V \rangle P$	Output
$u?(X : T) P$	Input
$\text{goto } v.P$	Migration
$\text{if } u_1 = u_2 \text{ then } P \text{ else } Q$	Matching
$(\text{new } c : C) P$	Channel creation
$(\text{newreg } n : N) P$	Global name creation
$(\text{newloc } k : K) P$	Location creation
$P   Q$	Composition
<b>stop</b>	Termination
$* P$	Iteration
<b>here</b> $[x] P$	<b>Location look up</b>
<b>rec</b> $(Z : R). P$	<b>Recursion</b>
$Z$	<b>Recursion variable</b>

---

## 2 The language recDpi

The syntax of RECDPI is given in Figure 1, and is a simple extension of that of DPI; the new constructs are highlighted in bold font. As usual it assumes a set of *names*, ranged over by letters such as  $a, b, c, k, l, \dots$ , and a separate set of *variables*, ranged over by  $x, y, z, \dots$ ; to handle recursive processes we use another set of *recursion variables*, ranged over by  $X, Y, Z, \dots$ . The values in the language include *identifiers*, that is names or variables, and *addresses*, of the form  $u \circ w$ ; intuitively  $w$  stands for a location and  $u$  a channel located there. In the paper we will consider only closed terms, where all variables (recursion included) are bound.

The most important new construct is that for typed recursive processes, **rec**  $(Z : R). P$ ; as we shall see the type  $R$  dictates the requirements on any site wishing to host this process. We also have a new construct **here**  $[x] P$ , which allows a process to know its current location.

**Example 2.1** [Searching a network] Consider the following recursive process, which searches a network for certain values satisfying some unspecified predicate  $p$ :

$$\text{Search} \triangleq \text{rec } Z : S. \text{test?}(x) \text{if } p(x) \text{ then goto home.report!}\langle x \rangle \\ \text{else neigh?}(y) \text{goto } y.Z$$

When placed at a specific site such as  $k$ , giving the system

$$k\llbracket\text{Search}\rrbracket,$$

the process first gets the local value from the channel  $test$ . If it satisfies the test the search is over; the process returns **home**, and *reports* the value. Otherwise it uses the local channel  $neigh$  to find a neighbour to the current site, migrates there and launches a recursive call at this new site. ■

We refrain from burdening the reader with a formal reduction semantics for RECDPI, as it is a minor extension of that of DPI. However in Section 4 we give a typed labelled transition system for the language, the  $\tau$ -moves of which provides our reduction semantics (see Figure 3). For the current discussion we can focus on the following rules:

$$\begin{array}{l} \text{(LTS-HERE)} \\ k\llbracket\text{here } [x] P\rrbracket \xrightarrow{\tau} k\llbracket P[k/x]\rrbracket \\ \text{(LTS-ITER)} \\ k\llbracket * P\rrbracket \xrightarrow{\tau} k\llbracket * P\rrbracket \mid k\llbracket P\rrbracket \\ \text{(LTS-REC)} \\ k\llbracket \text{rec } (Z : R). P\rrbracket \xrightarrow{\tau} k\llbracket P\{\text{rec } (Z:R). P/Z\}\rrbracket \end{array}$$

The first simply implements the capture of the current location by the construct **here**. The second states that the iterative process at  $k$ ,  $k\llbracket * P\rrbracket$  can spawn a new copy  $k\llbracket P\rrbracket$ , while retaining the iterated process. This means that every new copy of this process will be located in  $k$ . The final one, (LTS-REC), implements recursion in the standard manner by unwinding the body, which is done by replacing every free occurrence of the recursion variable  $Z$  in  $P$  by the recursive process itself. This takes an explicit  $\tau$ -reduction to match the rule (LTS-ITER).

**Example 2.2** [Self-locating processes] We give an example to show why the construct **here** is particularly interesting for recursive processes. Consider the system  $k\llbracket\text{Quest}\rrbracket$  where

$$\text{Quest} \triangleq \text{rec } Z : R. \text{here } [x] (\text{newc } ans) \text{neigh }?(y : R) \\ (\text{ans }?(news) \dots \mid \text{goto } y.\text{req}!\langle \text{data}, \text{ans}@x \rangle Z)$$

After determining its current location  $x$ , this process generates a new local channel  $ans$  at the current site  $k$ , and sets up a listener on this channel to await news. Concurrently it finds a neighbour, via the local channel  $neigh$ . It then migrates to this neighbour and poses a question there, via the channel  $req$ , and fires a new recursive call, this time at the neighbouring site. The neighbour's request channel  $req$  requires some data, **data**, and a return address, which in this case is given via the value  $\text{ans}@x$ .

Note that at runtime the occurrence of  $x$  in the value proffered to the channel  $req$  is substituted by the originating site  $k$ . After the first three steps

**Fig. 2** Recursive pre-types

---

Base Types:	$B ::= \text{int} \mid \text{bool} \mid \text{unit} \dots$
Local Channel Types:	$A ::= R\langle U \rangle \mid W\langle T \rangle \mid RW\langle U, T \rangle$
Capability Types:	$C ::= u : A$
Location Types:	$K ::= \text{LOC}[C_1, \dots, C_n], n \geq 0 \mid \mu Y.K \mid Y$
Registered Name Types:	$G ::= RC\langle A \rangle$
Value Types:	$V ::= B \mid A \mid (\tilde{A})_{@u} \mid (\tilde{A})_{@K}$
Transmission Types:	$T, U ::= (V_1, \dots, V_n), n \geq 0$

---

in the reduction of the system  $k[\text{Quest}]$ , we get to

$$(\text{new } ans) k[\text{neigh?}(y : R) (\text{goto } y.\text{req}!\langle \text{data}, ans_{@k} \rangle \text{Quest} \mid ans?(news) \dots)]$$

If  $k$ 's neighbour is  $l$ , this further reduces to (up to some reorganisation)

$$(\text{new } ans) k[\text{ans?}(news) \dots] \mid l[Q] \\ \mid (\text{new } ans') l[\text{neigh?}(y : R) (\text{goto } y.\text{req}!\langle \text{data}, ans'_{@l} \rangle \text{Quest} \mid ans'?(news) \dots)]$$

with  $Q$  some code running at  $l$  to answer the request brought by **Quest**.

The `here` construct can also be used to write a process initialising a doubly linked list starting from a simply linked one. We assume for this that the cells are locations containing two specific channels:  $n$  to get the name of the next cell in the list,  $p$  for the previous. The initial state of our system is

$$l_0[n!\langle l_1 \rangle] \mid l_1[n!\langle l_2 \rangle] \mid \dots$$

and we run the following code in the first cell of this network to initialise the list:

$$\text{rec } Z : R. n?(n') \text{ here } [p'] (n!\langle n' \rangle \mid \text{goto } n'.(p!\langle p' \rangle \mid Z))$$

■

Now we need to look more closely at the types, like  $R$ , involved in the recursive construct.

### 3 Co-inductive types for `recDpi`

There is a well-established capability-based type system for `DPI`, [4], which we can adapt to `RECDPI`.

#### 3.1 The Types

In this type system local channels have read/write types of the form  $R\langle U \rangle$ ,  $W\langle T \rangle$ , or  $RW\langle U, T \rangle$  (meaning that values are written at type  $T$  and read at

type  $U$  on a channel of that type), provided the object types  $U$  and  $T$  “agree”, as will be explained later. Locations have record types, of the form

$$\text{LOC}[u_1 : A_1, \dots, u_n : A_n]$$

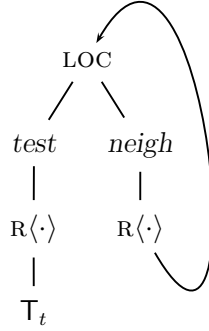
indicating that the local channels  $u_i$  may be used at the corresponding type  $A_i$ .

However with recursive processes it turns out that we need to consider *infinite* location types. To see this consider again the searching process **Search** from Example 2.1. Any site, such as  $k$ , which can support this process needs to have a local channel called *neigh* from which values can be read. These values must be locations, and let us consider their type, that is the object type of *neigh*. These locations must have a local channel called *test*, of an appropriate type, and a local channel called *neigh*; the object type of this local channel must be in turn the same as the type we are trying to describe. Using a recursion operator  $\mu$ , this type can be described as

$$\mu \mathbf{Y} . \text{LOC}[\text{test} : \mathbf{R}\langle \mathbf{T}_t \rangle, \text{neigh} : \mathbf{R}\langle \mathbf{Y} \rangle]$$

which will be used as the type  $S$  in the definition of **Search**; it describes precisely the requirements on any site wishing to host this process.

The set of recursive pre-types is given in Figure 2, and is obtained by adding the operator  $\mu \mathbf{Y} . K$  as a constructor to the type formation rules for DPI. Following [8] we can associate with each recursive pre-type  $T$  a co-inductive pre-type denoted  $\text{Tree}(T)$ , which takes the form of a finite-branching, but possibly infinite, tree whose nodes are labelled by the type constructors. For example  $\text{Tree}(S)$  is the infinite tree represented by the following graph:



**Definition 3.1** [Contractive and Tree pre-type] We call a recursive pre-type  $S$  *contractive* if for every  $\mu \mathbf{Y} . S'$  it contains,  $\mathbf{Y}$  can only appear in  $S'$  under an occurrence of  $\text{LOC}$ . In the paper we will only consider contractive pre-types.

For every contractive  $S$  we can define  $\text{Tree}(S)$ , the unique tree satisfying by the following equations:

- unwinding recursive pre-types  $\text{Tree}(\mu \mathbf{Y} . S') = \text{Tree}(S' \{\mu \mathbf{Y} . S' / \mathbf{Y}\})$
- not modifying any other construct; for instance  $\text{Tree}(\mathbf{R}\langle U \rangle) = \mathbf{R}\langle \text{Tree}(U) \rangle$

We call  $\text{Tree}(S)$  the *tree pre-type* associated with the recursive pre-type  $S$ . ■

Note that  $\text{Tree}(\mathbf{S})$  might not be defined when the recursive pre-type  $\mathbf{S}$  is not contractive.

To go from pre-types to types, we need to get rid of meaningless pre-types like  $\text{RW}\langle \mathbf{R}\langle \cdot \rangle, \text{int} \rangle$ , which would be the type of a channel on which integers are written but channels are read. This is achieved using a notion of subtype, and demanding that, in types of the form  $\text{RW}\langle \mathbf{U}, \mathbf{T} \rangle$ ,  $\mathbf{T}$  must be a subtype of  $\mathbf{U}$ .

In Figure A.1 (in the appendix) we give the standard set of rules which define the subtyping relation used in DPI; a typical rule, an instance of (SUB-CHAN), takes the form

$$\frac{\mathbf{T} <: \mathbf{U} <: \mathbf{U}'}{\text{RW}\langle \mathbf{U}, \mathbf{T} \rangle <: \mathbf{R}\langle \mathbf{U}' \rangle}$$

However here we interpret these rules co-inductively, [2]. Formally they give rise to a transformation on relations over tree pre-types. If  $\mathcal{R}$  is such a relation, then  $\text{Sub}(\mathcal{R})$  is the relation given by:

$$\begin{aligned} \text{Sub}(\mathcal{R}) = & \{(\mathbf{base}, \mathbf{base})\} \\ & \cup \{(u : \mathbf{A}, u : \mathbf{B}) \text{ if } (\mathbf{A}, \mathbf{B}) \text{ is in } \mathcal{R}\} \\ & \cup \{((\tilde{\mathbf{C}}), (\tilde{\mathbf{C}}')) \text{ if } (\mathbf{C}_i, \mathbf{C}'_i) \text{ is in } \mathcal{R} \text{ for all } i\} \\ & \cup \dots \end{aligned}$$

Intuitively, if the hypotheses of any rule of Figure A.1 are in  $\mathcal{R}$  the conclusion is in  $\text{Sub}(\mathcal{R})$ .

**Definition 3.2** [Subtyping and types] We define the *subtyping* relation between tree pre-types to be the greatest fixpoint of the function  $\text{Sub}$ , written  $\nu\text{Sub}$ . For convenience we often write  $\mathbf{T} <: \mathbf{T}'$  to mean that  $(\mathbf{T}, \mathbf{T}')$  is in  $\nu\text{Sub}$ .

Then a tree pre-type is called a *tree type* if every occurrence of  $\text{RW}\langle \mathbf{U}, \mathbf{T} \rangle$  it contains satisfies  $\mathbf{T} <: \mathbf{U}$ .

Finally this is lifted to recursive pre-types. A pre-type  $\mathbf{T}$  from Figure 2 is called a *recursive type* if  $\text{Tree}(\mathbf{T})$  is a tree type. ■

The co-inductive definition of subtyping gives rise to a natural co-inductive proof method, the dual of the usual inductive proof method used for sub-typing in DPI. This can be employed to derive many of the required properties of sub-typing in RECDPI. Here is a typical example. Let us write  $\mathbf{T}_1 \downarrow \mathbf{T}_2$  to mean that there is some  $\mathbf{T}$  such that  $\mathbf{T} <: \mathbf{T}_1$  and  $\mathbf{T} <: \mathbf{T}_2$ , that is  $\mathbf{T}_1$  and  $\mathbf{T}_2$  are compatible.

**Lemma 3.3** *The set of tree types, ordered by  $<:$ , has partial meets. That is  $\mathbf{T}_1 \downarrow \mathbf{T}_2$  implies  $\mathbf{T}_1$  and  $\mathbf{T}_2$  have a meet, denoted  $\mathbf{T}_1 \sqcap \mathbf{T}_2$ .*

In the full version of the paper, [5], we go on to show that this result also applies to recursive types, and moreover give a procedure for calculating the meet of any two compatible recursive types.

### 3.2 Typing Systems

With these types we can now adapt the typing system for DPI to RECDPI. At the system level the judgements take the form

$$\Gamma \vdash M$$

and the rules used are identical to those for DPI, see Figure 11 in the full version of this paper, [5], where they are restated. The main rule, there, is

$$\frac{\text{(T-PROC)} \quad \Gamma \vdash_k P}{\Gamma \vdash k[[P]]}$$

which in turn requires a set of inference rules for the judgements

$$\Gamma \vdash_k P$$

indicating that the process  $P$  is well-typed to run at location  $k$ . Once more most of these rules are inherited from DPI, see Figure 12 in [5], and we concentrate here on explaining the three new rules required for recursion and the `here` construct. The latter is straightforward:

$$\frac{\text{(T-HERE)} \quad \Gamma \vdash_w P[w/x]}{\Gamma \vdash_w \text{here}[x] P}$$

However in order to derive judgements about recursive processes, such as

$$\Gamma \vdash_k \text{rec}(Z : \mathbf{R}). P \tag{1}$$

we need to augment the type environments used in DPI with entries for recursion variables. Recall that here the type  $\mathbf{R}$  is a location type, such as  $\text{LOC}[u_1 : \mathbf{A}_1, \dots, u_n : \mathbf{A}_n]$ , indicating the minimal requirements on any location wishing to host a call to the recursive procedure. So in some way we want to consider recursion variables in the same manner as locations. But we must be careful as subtyping can not be allowed on these variables, unlike locations. Therefore we only allow unique entries of the form  $Z : \mathbf{K}$ , where  $\mathbf{K}$  is a location type, in type environments. Then the natural rule for typechecking a recursive call, that is an occurrence of a recursion variable, is given by:

$$\frac{\text{(T-RECVAR)} \quad \Gamma \vdash w : \Gamma(Z)}{\Gamma \vdash_w Z}$$

In order to typecheck a recursive definition, such as (1) above, we need to

- check that  $k$  has at least the capabilities required in  $\mathbf{R}$ , that is  $\Gamma \vdash k : \mathbf{R}$ ;



- ensure that the body  $P$  only uses the resources given in  $R$ .

To check this second point we again look at recursion variables as *locations*, and check that  $P$  is well-typed to run “in the location  $Z$ ”, which has all the resources mentioned in the type  $R$ . The final rule is

$$\begin{array}{c} \text{(T-REC)} \\ \Gamma \vdash w : R \\ \Gamma, \langle\langle Z : R \rangle\rangle \vdash_Z P \\ \hline \Gamma \vdash_w \text{rec } (Z : R). P \end{array}$$

where  $\Gamma, \langle\langle Z : R \rangle\rangle$  is a notation extending  $\Gamma$  with the information that  $Z$  has all the capabilities in  $R$ .

**Example 3.4** Referring back to Example 2.1 let us see how these rules can be used to infer  $\Gamma \vdash_k \mathbf{Search}$ , assuming that  $\Gamma$  knows about locations  $\mathbf{home}$ ,  $k$ , etc. and their channels. So, by (T-REC), this will amount to:

$$\Gamma, \langle\langle Z : S \rangle\rangle \vdash_Z \text{test?}(x) \text{if } p(x) \text{ then goto } \mathbf{home}.\text{report!}\langle x \rangle \\ \text{else } \text{neigh?}(y) \text{ goto } y.Z$$

which, in turn, will mainly consist of proving:

$$\begin{array}{l} \Gamma, \langle\langle Z : S \rangle\rangle \vdash \text{test} : R\langle T_t \rangle_{@Z} \\ \Gamma, \langle\langle Z : S \rangle\rangle, x : T_t \vdash \text{neigh} : R\langle S \rangle_{@Z} \\ \Gamma, \langle\langle Z : S \rangle\rangle, x : T_t, y : S \vdash_y Z \end{array}$$

These statements are provable since  $S$  is  $\mu \mathbf{Y}.\text{LOC}[\text{test} : R\langle T_t \rangle, \text{neigh} : R\langle \mathbf{Y} \rangle]$ ; the first one is obvious, while the second will be true by unfolding once the recursive type, which means in turn that the third will also be true.  $\blacksquare$

The partial view of recursion variables as locations complicates somewhat the formal rules for the construction of valid environments. In this extended abstract we do not go into the details. But for completeness sake we give the formation rules in the appendix, in Figure A.2, together with these for value typing, in Figure A.3. Notice that value typing rules allow statements of the form  $\Gamma \vdash Z : \text{LOC}$ , required when typing a process “at  $Z$ ”, even if, syntactically, recursion variables cannot be used as values.

The main new technical property of the type inference system is given by:

**Lemma 3.5 (Recursion Variable Substitution)** *Let us suppose that  $\Gamma \vdash_w \text{rec } Z : R. P$ . Then  $\Gamma \vdash_w P\{\text{rec } Z : R. P/Z\}$ .*

This in turn leads to:

**Lemma 3.6 (Subject Reduction)**  $\Gamma \vdash M$  and  $M \xrightarrow{\tau} M'$  implies that  $\Gamma \vdash M'$ .

## 4 Implementing recursion using iteration

The problem of implementing recursion using iteration in DPI, contrary to the  $\text{PI-CALCULUS}$ , is that any code of the form  $k\llbracket * P \rrbracket$  will force every instance of  $P$  to be launched at the originating site  $k$ ; this is in contrast to  $k\llbracket \text{rec } (Z : \mathbf{R}). P \rrbracket$  where the initial instance of the body  $P$  is launched at  $k$  but subsequent instances may be launched at arbitrary sites, provided they are appropriately typed.

Nevertheless, at the expense of repeated migrations, we can mimic the behaviour of a recursive process using iteration by designating a *home base* to which the process must return before a new instance is launched. For example if *home* is deemed to be the home base then we can implement our example  $k\llbracket \text{Search} \rrbracket$  using

$$\text{home}\llbracket * \text{IterSearch} \rrbracket \mid k\llbracket \text{FireOne} \rrbracket$$

where

$$\begin{aligned} \text{IterSearch} &\triangleq \text{ping}?(l) \text{ goto } l.\text{test}?(x) \text{ if } p(x) \text{ then goto } \text{home}.\text{report}!\langle x \rangle \\ &\quad \text{else } \text{neigh}?(y) \text{ goto } y.\text{FireOne} \\ \text{FireOne} &\triangleq \text{here } [l] \text{ goto } \text{home}.\text{ping}!\langle l \rangle \end{aligned}$$

With this example, we can easily see how the translation will mimic the original process step by step: the body of the process is left unmodified, only the recursion parts are changed, by implementing the recursive call with a few reductions. **FireOne** is the “translation” for the recursive calls, which means going to the home base and firing a new instance. This shows why the construct **here** is necessary: the translation for recursive calls needs to detect its current location to indeed trigger the new instance in the “proper” context. Then the replicated **IterSearch** starts off by migrating to the actual location where it will run.

This approach underlies our general translation of recursive processes into iterative processes, which we now explain.

As we want to ensure that our translation will be compositional, we will have to dynamically generate the home bases for iterative processes where, in the example **IterSearch**, the home base and the replicated process were already set up. We will also dynamically generate the registered channel *ping* used to provide to a new instance of the process the name of the location where the recursive call took place. The last thing to do when the recursion is unwound for the first time is to start the iterative process, which means two things: move the code that will be replicated to its home base and fire the first instance. As we explained with the example, the replicated code will just have to wait for the name of a location when the recursion is unwound, go

there and behave as the recursive process. So our translation looks like this:

$$\begin{aligned} \text{UNREC}(\text{rec } Z : R. P) &= (\text{newreg } ping_Z : \text{RC}\langle \text{RW}\langle R \rangle \rangle) \\ &\quad (\text{newloc } home_Z : \text{LOC}[ping_Z : \text{RC}\langle \text{RW}\langle R \rangle \rangle]) \\ &\quad (\text{UNREC}(Z) \mid \\ &\quad \quad \text{goto } home_Z. * ping_Z ?(l : R) \text{ goto } l. \text{UNREC}(P)) \\ \text{UNREC}(Z) &= \text{here } [x] \text{ goto } home_Z. ping_Z !\langle x \rangle \end{aligned}$$

Of course, any construct other than recursion is left unmodified by this translation; for example  $\text{UNREC}(u!\langle V \rangle P) = u!\langle V \rangle \text{UNREC}(P)$ .

We stress the fact that this translation heavily relies on migration to mimic the original process. We conjecture that in a DPI setting where locations or links can fail, like in [1], it would not be possible to get a reasonable encoding of recursion into iteration.

We could also give another translation, which would be closer to the one proposed for the PI-CALCULUS in [8] by:

- closing the free names of recursive processes, and then communicating their actual values through the channel  $ping$ , at the same time as the location;
- creating all the home bases at the top-level of the process, once and for all.

But such an approach would not be compositional.

Now that we have described our translation, we want to prove that the translation and the original process are “equivalent”, in some sense. Since we are in a typed setting, the first property we need to check is the following.

**Lemma 4.1**  $\Gamma \vdash M$  if and only if  $\Gamma \vdash \text{UNREC}(M)$

We can also show that the behaviour of  $M$  and that of its translation  $\text{UNREC}(M)$  are closely related. Intuitively we want to show that whenever  $\Gamma \vdash M$  then any observer, or indeed other system, which uses names according to the type constraints given in  $\Gamma$  can not differentiate between  $M$  and  $\text{UNREC}(M)$ . This idea has been formalised in [3] as a typed version of *reduction barbed congruence*, giving rise to the judgements

$$\Gamma \models M \cong_{rbc} N$$

The reader is referred to [3] for the formal details.

**Theorem 4.2** Suppose  $\Gamma \vdash M$ . Then  $\Gamma \models M \cong_{rbc} \text{UNREC}(M)$ .

The proof uses a characterisation of this relation as a bisimulation equivalence in a labelled transition system in which:

- the states are *configurations* of the form  $\Gamma \triangleright M$ ;
- the actions take the form  $\Gamma \triangleright M \xrightarrow{-\mu} \Gamma' \triangleright M'$ ; these are based on the labelled transitions system given in Figure 3 and 4.

**Fig. 3** Labelled transition semantics. Internal actions.

$$\begin{array}{l}
 \text{(LTS-GO)} \\
 \Gamma \triangleright k[\text{goto } l.P] \xrightarrow{\tau}_\beta \Gamma \triangleright l[P] \\
 \text{(LTS-SPLIT)} \\
 \Gamma \triangleright k[P \mid Q] \xrightarrow{\tau}_\beta \Gamma \triangleright k[P] \mid k[Q] \\
 \text{(LTS-ITER)} \\
 \Gamma \triangleright k[* P] \xrightarrow{\tau}_\beta \Gamma \triangleright k[* P] \mid k[P] \\
 \text{(LTS-HERE)} \\
 \Gamma \triangleright k[\text{here } [x] P] \xrightarrow{\tau}_\beta \Gamma \triangleright k[P^{k/x}] \\
 \text{(LTS-REC)} \\
 \Gamma \triangleright k[\text{rec } (Z : R). P] \xrightarrow{\tau}_\beta \Gamma \triangleright k[P\{\text{rec } (Z:R). P/Z\}] \\
 \text{(LTS-L-CREATE)} \\
 \Gamma \triangleright k[(\text{newloc } l : L) P] \xrightarrow{\tau}_\beta \Gamma \triangleright (\text{new } l : L) k[P] \mid l[C] \\
 \text{(LTS-N-CREATE)} \\
 \Gamma \triangleright k[(\text{newreg } n : N) P] \xrightarrow{\tau}_\beta \Gamma \triangleright (\text{new } n : N) k[P] \\
 \text{(LTS-C.CREATE)} \\
 \Gamma \triangleright k[(\text{newc } c : C) P] \xrightarrow{\tau}_\beta \Gamma \triangleright (\text{new } c : C_{\text{@}k}) k[P] \\
 \text{(LTS-EQ)} \\
 \Gamma \triangleright k[\text{if } u = u \text{ then } P \text{ else } Q] \xrightarrow{\tau}_\beta \Gamma \triangleright k[P] \\
 \text{(LTS-NEQ)} \\
 \Gamma \triangleright k[\text{if } u = v \text{ then } P \text{ else } Q] \xrightarrow{\tau}_\beta \Gamma \triangleright k[Q] \quad \text{when } u \neq v \\
 \text{(LTS-COMM)} \\
 \frac{\Gamma_M \triangleright M \xrightarrow{(\tilde{n}:\tilde{T})k.a!V} \Gamma'_M \triangleright M' \quad \Gamma_N \triangleright N \xrightarrow{(\tilde{n}:\tilde{U})k.a?V} \Gamma'_N \triangleright N'}{\Gamma \triangleright M \mid N \xrightarrow{\tau} \Gamma \triangleright (\text{new } \tilde{n} : \tilde{T}) M' \mid N' \quad \tilde{n} \cap \text{fn}(N) = \emptyset} \\
 \Gamma \triangleright N \mid M \xrightarrow{\tau} \Gamma \triangleright (\text{new } \tilde{n} : \tilde{T}) N' \mid M'
 \end{array}$$

**Definition 4.3** [Actions] For configurations  $\mathcal{C}$  of the form  $(\Gamma \triangleright M)$ , we say that they can do the following actions:

- $\mathcal{C} \xrightarrow{\tau} \mathcal{C}'$  or  $\mathcal{C} \xrightarrow{(\tilde{n}:\tilde{T})k.a?V} \mathcal{C}'$  if we can prove so with a derivation in the LTS;
- $\mathcal{C} \xrightarrow{(\tilde{n})k.a!V} \mathcal{C}'$  if there exists some derivation proving  $\mathcal{C} \xrightarrow{(\tilde{m}:\tilde{T}')k.a!V} \mathcal{C}'$  in the LTS with  $(\tilde{n})$  the names that are both in  $V$  and  $(\tilde{m})$ .  $\blacksquare$

Again we refer the reader to [3] for further motivation; this paper also contains the result that

$$(\Gamma \triangleright M) \approx_{bis} (\Gamma \triangleright N) \text{ implies } \Gamma \models M \cong_{rbc} N$$

whenever  $\Gamma \vdash M$  and  $\Gamma \vdash N$ . So we establish Theorem 4.2 by showing

$$\Gamma \vdash M \text{ implies } (\Gamma \triangleright M) \approx_{bis} (\Gamma \triangleright \text{UNREC}(M)) \quad (2)$$

**Fig. 4** Labelled transition semantics. External actions.

$$\begin{array}{c}
 \text{(LTS-OUT)} \\
 \Gamma \vdash k : \text{LOC} \\
 a : \text{R}\langle \mathbb{T} \rangle_{\otimes k} \in \Gamma \\
 \Gamma, \langle V : \mathbb{T} \rangle_{\otimes k} \vdash \text{env} \\
 \hline
 \Gamma \triangleright k \llbracket a!(V) P \rrbracket \xrightarrow{k.a!V} \Gamma, \langle V : \mathbb{T} \rangle_{\otimes k} \triangleright k \llbracket P \rrbracket \\
 \\
 \text{(LTS-IN)} \\
 \Gamma \vdash k : \text{LOC} \\
 a : \text{W}\langle \mathbb{U} \rangle_{\otimes k} \in \Gamma \\
 \Gamma \vdash V : \mathbb{U}_{\otimes k} \\
 \hline
 \Gamma \triangleright k \llbracket a?(X : \mathbb{T}) P \rrbracket \xrightarrow{k.a?V} \Gamma \triangleright k \llbracket P \{V/X\} \rrbracket \\
 \\
 \text{(LTS-NEW)} \\
 \Gamma \triangleright M \xrightarrow{\mu} \Gamma' \triangleright M' \\
 \hline
 \Gamma \triangleright (\text{new } n : \mathbb{T}) M \xrightarrow{\mu} \Gamma' \triangleright (\text{new } n : \mathbb{T}) M' \quad n \notin \mu \\
 \\
 \text{(LTS-OPEN)} \\
 \Gamma \triangleright M \xrightarrow{(\tilde{n}:\tilde{\mathbb{T}})k.a!V} \Gamma' \triangleright M' \quad n \notin \{a, k\} \\
 \hline
 \Gamma \triangleright (\text{new } n : \mathbb{T}) M \xrightarrow{(\tilde{n}:\tilde{\mathbb{T}})k.a!V} \Gamma' \triangleright M' \quad n \in \text{fn}(V) \cup \text{n}(\tilde{\mathbb{T}}) \\
 \\
 \text{(LTS-WEAK)} \\
 \Gamma, \langle n : \mathbb{T} \rangle \triangleright M \xrightarrow{(\tilde{n}:\tilde{\mathbb{T}})k.a?V} \Gamma' \triangleright M' \\
 \hline
 \Gamma \triangleright M \xrightarrow{(n:\mathbb{T}, \tilde{n}:\tilde{\mathbb{T}})k.a?V} \Gamma' \triangleright M' \quad n \notin \{a, k\} \\
 \\
 \text{(LTS-PAR)} \\
 \Gamma \triangleright M \xrightarrow{\mu} \Gamma' \triangleright M' \\
 \hline
 \Gamma \triangleright M \mid N \xrightarrow{\mu} \Gamma' \triangleright M' \mid N \quad \text{bn}(\mu) \cap \text{fn}(N) = \emptyset
 \end{array}$$

## 5 Proof of recursion implementability

Let us see the problems encountered in trying to prove the equation (2) on an example. For this, let us consider a parameterised server version of our `Search` process that would be exploring a binary tree instead of a list:

$$\begin{aligned}
 \text{PSearch} &\triangleq \text{search\_req?}(x, \text{client}) \\
 &\quad \text{goto } k_0.\text{rec } Z : \text{S}. \text{test?}(y) \text{if } p(x, y) \text{ then goto } \text{client.report!}\langle y \rangle \\
 &\quad \quad \text{else } \text{neigh?}(n_1, n_2) \text{ goto } n_1.Z \mid \text{goto } n_2.Z
 \end{aligned}$$

used in the system `Server[* PSearch]`. So this sets up a search server, at `Server`; but the difference with `Search` from Example 2.1 is the fact that the data to search for in the network is given in the search request on `search_req`, and is subsequently used as a parameter by the testing predicate  $p$ .

Our translation of this process gives the following DPI code:

$$\begin{aligned} \text{IPSearch} &\triangleq \text{search\_req?}(x, \text{client}) \\ &\quad \text{goto } k_0. (\text{newreg } \text{ping}) (\text{newloc } \text{base}) \text{F} \mid \text{goto } \text{base}. * \text{Inst} \\ \text{Inst} &\triangleq \text{ping?}(k) \text{goto } k. \text{test?}(y) \text{ if } p(x, y) \text{ then goto } \text{client.report!}\langle y \rangle \\ &\quad \text{else } \text{neigh?}(n_1, n_2) \text{goto } n_1. \text{F} \mid \text{goto } n_2. \text{F} \\ \text{F} &\triangleq \text{here } [l] \text{goto } \text{base.ping!}\langle l \rangle \end{aligned}$$

with `Inst` an instance of the iterative process, and `F` the triggering process, written `FireOne` is the example in the previous section.

Since `IPSearch` is replicated, it will generate a new home base for `Inst` for every request on `search_req`. This means that, after servicing a number of such requests we will end up with a system of the form:

$$\begin{aligned} &(\text{new } \text{ping}_1) (\text{new } \text{base}_1) (\text{new } \text{ping}_2) (\text{new } \text{base}_2) \dots \\ &\quad \text{Server}[\dots] \mid \text{base}_1[\dots] \mid k_1^1[\dots \text{F}_1] \mid k_1^2[\dots \text{F}_1] \mid \dots \mid \text{base}_2[\dots] \mid k_2^1[\dots \text{F}_2] \dots \end{aligned} \quad (3)$$

Of course, this will correspond to the RECDPI system:

$$\text{Server}[\dots] \mid k_1^1[\dots \text{rec } Z. P] \mid k_1^2[\dots \text{rec } Z. P] \mid \dots \mid k_2^1[\dots \text{rec } Z. P] \dots$$

On this example, we can see quite clearly the main difference at runtime between our translation and the standard, but non-compositional, one used in the `PI-CALCULUS` we previously mentioned (see [8]), which arises because of the replication of `rec Z. P`. A translation following the lines of that in [8], would give rise to the following state, corresponding to (3) above:

$$\begin{aligned} &(\text{new } \text{ping}) (\text{new } \text{base}) \text{Server}[* \text{search\_req?}(x, \text{client}) \text{goto } k_0. \text{F}(x, \text{client})] \\ &\quad \mid \text{base}[* \text{ping?}(k, x, \text{client}) \text{goto } k. \text{test?}(y) \dots] \\ &\quad \mid k_1^1[\dots \text{F}(x_1, \text{client}_1)] \mid k_1^2[\dots \text{F}(x_1, \text{client}_1)] \mid \dots \mid k_2^1[\dots \text{F}(x_2, \text{client}_2)] \dots \\ &\text{F}(x, \text{client}) \triangleq \text{here } [l] \text{goto } \text{base.ping!}\langle l, x, \text{client} \rangle \end{aligned}$$

Note that here all the free names used in the recursive process are closed and the actual parameters are obtained when an instance is called via `ping`. But more importantly only one home base is ever created. Thus the loss of compositionality allows an easier proof of equivalence, since there is only one `base` per recursion variable.

To return to the discussion of our translation, we have here a RECDPI process containing a number of recursive constructs but the way they are to be translated to get the DPI system (3) depends on the system history. That is why our proof of (2) is based on an extended version of the translation in which we specify whether a given occurrence of `rec Z. P` has already been attributed a home base. If not, it should generate a new one; if it has, then the

actual home base needs to be recorded. In the example, we need to attribute the same home base to the  $\text{rec } Z. P$  in every  $k_1^i$ , and different ones for the other  $k_j^i$ .

Let us write  $\text{UNREC}_{\mathcal{P}}(M)$  for the translation of  $M$  parameterised by  $\mathcal{P}$ , with  $\mathcal{P}$  specifying how each  $\text{rec } Z. P$  should be translated in  $M$ . Then to prove the equation (2), we would like to exhibit a bisimulation containing the pair

$$(\Gamma \triangleright M, \Gamma \triangleright \text{UNREC}_{\mathcal{P}}(M))$$

for any term  $M$ , any environment  $\Gamma$  such that  $\Gamma \triangleright M$  is well-formed configuration and any parameter  $\mathcal{P}$ .

But we still have a major problem in trying to exhibit this bisimulation: there are a lot of possible states in the translation corresponding to only one state in the original process, since we introduce numerous extra reductions. To deal with those extra steps, we will resort to a proof technique given in [6], namely bisimulation up-to- $\beta$ . This is based on the remark that, among the reductions added by the translation, only the communication on the channel *ping* is “dangerous”, because it could fail if one of the two agents involved in the communication were absent. Every other step is a so-called  $\beta$ -move, written  $\xrightarrow{\tau}_{\beta}$  in the LTS, in Figure 3. Thanks to bisimulations up-to- $\beta$  we can focus only on the communication moves, by using a new variant of the translation, written  $\text{UNREC}_{\mathcal{P}}^{\beta}(M)$ , meaning that every  $\beta$ -move required for the bisimulation has been done. Then by considering that the *ping*-communication (which is a  $\tau$ -move) in the translation corresponds to the recursion unwinding in  $\text{RECDPI}$ , we can prove that the relation

$$(\Gamma \triangleright M, \Gamma \triangleright \text{UNREC}_{\mathcal{P}}^{\beta}(M))$$

is a bisimulation-up-to- $\beta$ .

## 6 Conclusion

In this paper we gave an extension of the DPI-calculus with recursive processes. In particular we described why this construct was more suited to programming in the distributed setting, by allowing the description of agents migrating through network, visiting and interrogating different locations. We also gave a typing system for this extended calculus, which involved recursive types, dealt with by using co-inductive proof techniques, and showed that Subject Reduction remains valid. Finally we showed how to encode our recursive processes into standard DPI which uses iteration, by resorting to the addition of extra migrations in the network. The encoding was proved to be sound and complete, in the sense that the original and translated processes are indistinguishable in a typed version of reduction barbed congruence.

It would now be interesting to study the behaviour of recursive processes in a setting where some parts of the network could fail (either locations or links),

since failures are of major importance in the study of distributed computations. We conjecture that in such a setting there is no translation of recursive processes into iterative ones, which preserve their behaviour.

## References

- [1] Francalanza, A. and M. Hennessy, *Location and link failure in a distributed pi-calculus*, Computer Science Report 2005:01, University of Sussex (2005).
- [2] Gapeyev, V., M. Levin and B. Pierce, *Recursive subtyping revealed*, Journal of Functional Programming **12** (2003), pp. 511–548, preliminary version in *International Conference on Functional Programming (ICFP)*, 2000. Also appears as Chapter 21 of *Types and Programming Languages* by Benjamin C. Pierce (MIT Press, 2002).
- [3] Hennessy, M., M. Merro and J. Rathke, *Towards a behavioural theory of access and mobility control in distributed systems*, Theoretical Computer Science **322** (2003), pp. 615–669.
- [4] Hennessy, M. and J. Riely, *Resource access control in systems of mobile agents*, Information and Computation **173** (2002), pp. 82–120.
- [5] Hym, S. and M. Hennessy, *Adding recursion to Dpi*, Computer science report, University of Sussex (2005).
- [6] Jeffrey, A. and J. Rathke, *A theory of bisimulation for a fragment of concurrent ml with local names*, Theoretical Computer Science **323** (2004), pp. 1–48.
- [7] Milner, R., “Communicating and Mobile Systems: the  $\pi$ -Calculus,” Cambridge University Press, 1999.
- [8] Sangiorgi, D. and D. Walker, “The  $\pi$ -calculus,” Cambridge University Press, 2001.



## A Dpi subtyping and typing rules

---

**Fig. A.1** Subtyping rules
 

---

(SUB-BASE)

$$\frac{}{\mathbf{base} <: \mathbf{base}}$$

(SUB-CAP)

$$\frac{A <: B}{u : A <: u : B}$$

(SUB-TUPLE)

$$\frac{C_i <: C'_i}{\widetilde{C} <: \widetilde{C}'}$$

(SUB-HOM)

$$A_1 <: A_2$$

$$K_1 <: K_2$$

(SUB-CHAN)

$$\frac{T_2 <: T_1 <: U_1 <: U_2}{W\langle T_1 \rangle <: W\langle T_2 \rangle}$$

$$R\langle U_1 \rangle <: R\langle U_2 \rangle$$

$$RW\langle U_1, T_1 \rangle <: R\langle U_2 \rangle$$

$$RW\langle U_1, T_1 \rangle <: W\langle T_2 \rangle$$

$$RW\langle U_1, T_1 \rangle <: RW\langle U_2, T_2 \rangle$$

$$A_1 @ K_1 <: A_2 @ K_2$$

$$A_1 @ u <: A_2 @ u$$

$$RC\langle A_1 \rangle <: RC\langle A_2 \rangle$$

(SUB-LOC)

$$\frac{A_i <: A'_i, \quad 1 \leq i \leq n}{LOC[u_1 : A_1, \dots, u_n : A_n, \dots, u_{n+p} : A_{n+p}] <: LOC[u_1 : A'_1, \dots, u_n : A'_n]}$$

**Fig. A.2** Well-formed environments

<p>(E-EMPTY)  <math>\vdash \mathbf{env}</math></p> <p>(E-NEW-LCHAN)  <math>\Gamma \vdash \mathbf{env}</math>  <math>\Gamma \vdash w : \text{LOC}</math>  <math>\frac{\Gamma(u) = \{A_{i@v_i}\} \quad \{v_i, w\} \text{ contains at most one location}}{\Gamma, u : A_{i@w} \vdash \mathbf{env}} \quad \{A_i\} \downarrow A</math></p> <p>(E-RCHAN)  <math>\frac{\Gamma \vdash \mathbf{env}}{\Gamma, u : \text{RC}\langle A \rangle \vdash \mathbf{env}} \quad u \notin \Gamma</math></p> <p>(E-REC)  <math>\Gamma \vdash \mathbf{env}</math>  <math>\frac{\Gamma(u_i) = \{\text{RC}\langle A_{ij} \rangle, A_{ik@v_k}\} \quad Z \notin \Gamma}{\Gamma, Z : \text{LOC}[(u_i : A_i)] \vdash \mathbf{env}} \quad \{A_{ij}, A_{ik}\} \downarrow A_i</math></p>	<p>(E-BASE)  <math>\frac{\Gamma \vdash \mathbf{env}}{\Gamma, u : \mathbf{base} \vdash \mathbf{env}} \quad \Gamma(u) \downarrow \mathbf{base}</math></p> <p>(E-REF-LCHAN)  <math>\Gamma \vdash \mathbf{env}</math>  <math>\Gamma \vdash w : \text{LOC}</math>  <math>\frac{\Gamma \vdash u : \text{RC}\langle B \rangle, \quad B &lt;: A}{\Gamma, u : A_{i@w} \vdash \mathbf{env}}</math></p> <p>(E-LOC)  <math>\frac{\Gamma \vdash \mathbf{env}}{\Gamma, v : \text{LOC} \vdash \mathbf{env}} \quad \Gamma(v) \downarrow \text{LOC}</math></p> <p>(E-DEC-AT-REC)  <math>\Gamma \vdash \mathbf{env}</math>  <math>\frac{\Gamma(Z) = \text{LOC}[\dots, u : A, \dots]}{\Gamma, u : A_{i@Z} \vdash \mathbf{env}}</math></p>
--	--

**Fig. A.3** Typing values

<p>(V-MEET)  <math>\Gamma \vdash u : T_1</math>  <math>\Gamma \vdash u : T_2</math>  <math>\frac{}{\Gamma \vdash u : T_1 \sqcap T_2}</math></p> <p>(V-NAME)  <math>\frac{\Gamma, u : T, \Gamma' \vdash \mathbf{env}}{\Gamma, u : T, \Gamma' \vdash u : T'} \quad T &lt;: T'</math></p> <p>(V-LOCATED-CHANNEL)  <math>\Gamma \vdash u_i : A_{i@v}</math>  <math>\Gamma \vdash v : K</math>  <math>\frac{}{\Gamma \vdash (\tilde{u})_{@v} : (\tilde{A})_{@K}}</math></p> <p>(V-LOCATED-TUPLE)  <math>\Gamma \vdash u_i : A_{i@k}</math>  <math>\frac{}{\Gamma \vdash (\tilde{u}) : (\tilde{A})_{@k}}</math></p>	<p>(V-TUPLE)  <math>\frac{\Gamma \vdash u_i : T_i}{\Gamma \vdash (\tilde{u}) : (\tilde{T})}</math></p> <p>(V-LOC)  <math>\Gamma \vdash v : \text{LOC}</math>  <math>\Gamma \vdash u_i : A_{i@v}</math>  <math>\frac{}{\Gamma \vdash v : \text{LOC}[u_1 : A_1, \dots, u_n : A_n]}</math></p> <p>(V-DEC-LOC)  <math>\Gamma \vdash v : \text{LOC}</math>  <math>\Gamma \vdash u_i : A_{i@v}</math>  <math>\Gamma \vdash u_i : \text{RC}\langle D_i \rangle, D_i &lt;: A_i</math>  <math>\frac{}{\Gamma \vdash_{\text{dec}} v : \text{LOC}[u_1 : A_1, \dots, u_n : A_n]}</math></p>
---	---