

On Static Location in Links

(A progress report)

Kwanghoon Choi

Chonnam National University, Gwangju, Republic of Korea
kwanghoon.choi@jnu.ac.kr

Abstract

This progress report describes an attempt to recover static location contexts in the Links-RPC calculus, which captures the RPC aspect of Links, by compiling to the polymorphic RPC calculus and back to itself. Currently, the proposed idea is limited to the rank-1 subset of the Links-RPC calculus. The limitation is caused by a technical problem involving polymorphic constructs in the adjustment of locations to resolve the location discrepancies resulting from our heuristic type compilation. We will continue to study to extend the applicability to the whole calculus by solving this problem.

Keywords: Location inference, Links, RPC calculus

1 Introduction

A motivation is this. The feature of location types would be good for multi-tier programming languages like Links in local computation optimization and static location inconsistency checking.

However, a naive integration of this feature in Links should be challenging because there might be a need to change several advanced features in a significant way. An alternative direction would pursue a way to recover static location contexts in Links equally useful for the aforementioned benefits but in the hope of no change or minimal changes that do not affect the Links language system so much.

Note that this paper refers to Links as its intermediate representation, the Links IR.

Figure 1 shows a general methodology to approach this alternative direction by analyzing static location contexts via compiling to the polymorphic RPC calculus and then by representing the analyzed context information in Links via compiling back.



Figure 1: Recovering static location contexts by compiling to the polymorphic RPC calculus and back

Let us discuss location inference and representation by compiling to the polymorphic RPC calculus and subsequently compiling back. In Links, a client map function would be written as follows.

```
1  sig map : (a -> b, [a]) -> [b]
2  fun map(f, xs) client {
3      switch( xs ) {
4          case [] -> []
5          case y::ys -> f(y) :: map (f, ys)
6      }
7  }
```

As seen in the example, having a client location annotation enforces the map function to run at the client, but the location information is disregarded by types. Programmers may view the argument type as a client function type of $a \rightarrow b$ but the client map function could be applied to a server function of the same type. Therefore, the implementation of $f(y)$ should always check the location of the function f before its application to y .

A heuristic could be made up to reflect the programmers' view: by the presence of the client location annotation, all function types occurring in the map function type would be assumed to have the same location annotation. The following signature would describes the idea of this heuristic if Links did allow such a function type. For notation, $a \xrightarrow{client} b$ intends to mean a type of functions that are guaranteed to run in the client as in the polymorphic RPC calculus [2].

```
1  sig map : (a -client-> b, [a]) -client-> [b]
```

A client map function of this imaginary signature might not have to check the location of the function f in $f(y)$ anymore as long as f is guaranteed to be a client function as the signature. This is a local computation optimization.

When the map function is applied to a server function g of type $a \xrightarrow{server} b$, the location of g is required to be adjusted in some way to run this function inside the client map function that does not runtime checking. Choi and Chang [1] proposed an idea of using an eta conversion for the location adjustment as

$$fun(x) \text{ client } g(x)$$

in the Links syntax¹. This anonymous function is of type $a \xrightarrow{client} b$ as the signature. It would be responsible for moving to the server to perform the application of $g(x)$ while the map function is still running in the client without runtime location checking.

Thus our type inference strategy will consist of a heuristic of interpreting location annotation as explained and an adjustment of locations as the eta conversion in the example.

Generally, the map function is written as a location-neutral one where it should be able to run both of in the client and in the server. In Links, the map function is written without location annotation as follows.

```
1  sig map : (a -> b, [a]) -> [b]
2  fun map(f, xs) {
3    switch( xs ) {
4      case [] -> []
5      case y::ys -> f(y) :: map (f, ys)
6    }
7  }
```

Programmers could view this unannotated function as a polymorphic location function. This perception would be made use of when compiling into the polymorphic RPC calculus where programmers can write location abstractions and applications with location variables. Combining the previous heuristic with a location variable, a map function could be made up as a location polymorphic version where $map[client]$ is a client one and $map[server]$ is a server one.

```
1  sig map : forall l. (a -l-> b, [a]) -l-> [b]
```

Unfortunately, Links does not support the notion of location variables nor location abstractions and applications. So, we need to encode this location polymorphism in a way.

One way of encoding locations in Links would make use of GADTs as proposed in Choi et al [2]. Firstly, let *Location* α be a GADT whose data constructors are *Client* and *Server*. They have type *Location ClientType* and *Location ServerType*, respectively. The polymorphic map function of the signature would be written in Links, as follows.

```
1  sig map : Location l -> (a -> b, [a]) -> [b]
2  fun map(z) {
3    fun map'(f, xs) {
4      switch( xs ) {
5        case [] -> []
6        case y::ys -> f(y) :: map' (f, ys)
7      }
8    }
```

¹Actually, Links does not allow anonymous functions to have location annotations.

The Links signature should be interpreted similarly as explained with the heuristic previously. That is, it would be viewed as $(a \xrightarrow{L} b, [a]) \xrightarrow{L} [b]$.

This polymorphic map function can run wherever it chooses to run by *map Client* or *map Server* by applying it to one of the data constructors that represent client and server locations. Interestingly, the polymorphic map function can still run as a local computation. So, there is no need to refer to z in the map example.

However, in other examples, there is sometimes a need to check locations in runtime. For this purpose, the polymorphic map function is written to take a value for representing locations as an argument. When runtime location checking is necessary, the argument z could be examined.

This is a glimpse of an idea of location representation when compiling back to Links.

Section 2 defines a Links-RPC calculus. Section 3 discusses a compilation method of λ_{rpc}^\forall into λ_{rpc}^{Links} . Section 4 discusses a reverse direction compilation method. Section 5 concludes.

2 A Links-RPC Calculus

This section describes a RPC calculus that extends the RPC calculus [5] with unannotated λ -abstractions whose locations are not specified. The extended RPC calculus intends to capture the RPC feature of the Links programming language [4] in a simple way. For our setting, the calculus is also extended with polymorphic types with type abstractions and applications. Let us call it a Links-RPC calculus, λ_{rpc}^{Links} .

Syntax

Location	a, b	$::=$	$\mathbf{c} \mid \mathbf{s}$
	$?a$	$::=$	$a \mid (unknown)$
Term	L, M, N	$::=$	$V \mid L M \mid M[A] \mid (L, M) \mid \pi_i(M)$
Value	V, W	$::=$	$x \mid \lambda^{?a}x.M \mid \Lambda\alpha.V \mid (V, W)$

Types

Type	A, B, C	$::=$	$base \mid A \rightarrow B \mid \alpha \mid A \times B \mid \forall\alpha.A$
------	-----------	-------	--

Figure 2: The Links-RPC calculus λ_{rpc}^{Links}

The terms and types of the Links-RPC calculus are shown in Figure 2. In the calculus, it is not mandatory to annotate locations to λ abstractions. This is for programmers' convenience. An optional location $?a$ means that every annotation is either a location constant a or unknown. Regardless of the presence of location annotations, function types are $A \rightarrow B$ with no locations in types. No notion of location variables is available in the calculus, and so there are no location abstractions nor applications. Other than these differences, the terms and types are almost the same as those for the polymorphic RPC calculus, which is in the appendix for reference.

The semantics for λ_{rpc}^{Links} includes an evaluation rule for unannotated λ -abstractions, which once was proposed by the original RPC calculus [5], that evaluate to λ -abstractions annotated with the location of evaluation as

$$(\text{Unknown-Abs}) \frac{}{\lambda x.M \Downarrow_a \lambda^a x.M}$$

The typing rules are actually the same as for the System-F calculus except the appearance of annotated λ -abstractions whose locations are disregarded by types.

$$(\text{T-?a-Abs}) \frac{\Gamma, x : A \vdash M : B}{\Gamma \vdash \lambda^{?a}x.M : A \rightarrow B}$$

For brevity, we omit writing the full definitions of the semantics and the typing rules for λ_{rpc}^{Links} .

3 Compiling λ_{rpc}^{Links} into λ_{rpc}^\forall

The purpose of compiling the Links-RPC calculus into the polymorphic RPC calculus is to recover location information on where the evaluation of λ_{rpc}^{Links} terms are and to express it by located types of λ_{rpc}^\forall explicitly. Such statically recovered location information would be made use of later for transformations and optimizations, such as avoiding unnecessary location checking at runtime for solely local computation.

Basically, this compilation would translate normal function types $A_1 \rightarrow A_2$ into located function types $B_1 \xrightarrow{Loc} B_2$ for some B_i s recoving the static location information Loc . Then an obvious question would be where such location information comes from. There is a need to introduce a heuristic to recover location information lack in the Links-RPC terms and types.

A heuristic approach is to view λ_{rpc}^{Links} types A as λ_{rpc}^\forall types B where

- the *skeleton* of B obtained from erasing all locations from it is the same as A , and
- all the erased locations are assumed to be the same as a given location Loc .

For example, the type of a client map function, $(Int \rightarrow Int) \times [Int] \rightarrow [Int]$, can be viewed as $(Int \xrightarrow{c} Int) \xrightarrow{c} [Int] \xrightarrow{c} [Int]$. As is seen, not only a function type on the spine gets client locations but an argument function type is also annotated with it in the same way. This is a heuristic that we are going to adopt for our location inference during compiling λ_{rpc}^{Links} to λ_{rpc}^\forall .

This heuristic idea can be formulated as a type translation $\llbracket A \rrbracket_{Loc} = B$:

$$\begin{array}{llll} \llbracket \alpha \rrbracket_{Loc} & = & \alpha & \llbracket A \rightarrow B \rrbracket_{Loc} = \llbracket A \rrbracket_{Loc} \xrightarrow{Loc} \llbracket B \rrbracket_{Loc} \\ \llbracket base \rrbracket_{Loc} & = & base & \llbracket A \times B \rrbracket_{Loc} = \llbracket A \rrbracket_{Loc} \times \llbracket B \rrbracket_{Loc} \quad \llbracket \forall \alpha. A \rrbracket_{Loc} = \forall \alpha. \llbracket A \rrbracket_{Loc} \end{array}$$

Figure 3: A type compilation of λ_{rpc}^{Links} into λ_{rpc}^\forall for our heuristic interpretation

Note that the explained client map function type in λ_{rpc}^\forall is obtained from $\llbracket (Int \rightarrow Int) \times [Int] \rightarrow [Int] \rrbracket_c$ by extending it over list types as $\llbracket [A] \rrbracket_{Loc} = \llbracket A \rrbracket_{Loc}$ in a straightforward way.

Obviously, this heuristic is not always satisfactory. For example, *map f list* is legitimate in λ_{rpc}^{Links} even when the client map function takes a function of type $Int \xrightarrow{c} Int$ as the first argument but a server function of type $Int \xrightarrow{s} Int$ is given for the argument in λ_{rpc}^\forall . This is because λ_{rpc}^{Links} can deal with this discrepancy by runtime location checking. However, there should be some adjustments to locations in compiled types and terms.

A type-directed location adjustment is formulated as the following form of judgments over terms and types in λ_{rpc}^\forall .

$$M : A \Rightarrow M' : A'$$

From a term M of type A , a new adjusted term M' is derived under the guidance of another type A' having the same skeleton as but possibly locational discrepancies with A . For example,

$$f : Int \xrightarrow{s} Int \Rightarrow (\lambda^c x. f x) : Int \xrightarrow{c} Int$$

where given three inputs, f , the type of f , and an adjusted type, an adjusted client function $\lambda^c x. f x$ is derived. Then the adjusted function could be fed into the client map function as its first argument in λ_{rpc}^\forall .

A type-directed location adjustment to terms is formulated as Figure 4. The rules are defined in terms of type structure, actually doing the eta conversion for function, pair, polymorphic, polymorphic location types of given terms. By the conversion, they will adjust all discrepant locations of two function types to produce new terms of the adjusted types.

$$\begin{array}{c} \frac{}{M : A \Rightarrow M : A} \quad \frac{x : C \Rightarrow N : A \quad M N : B \Rightarrow M' : D}{M : A \xrightarrow{Loc_1} B \Rightarrow \lambda^{Loc_2} x. M' : C \xrightarrow{Loc_2} D} \quad \frac{\pi_i M : A_i \Rightarrow M'_i : A'_i \quad (i = 1, 2)}{M : A_1 \times A_2 \Rightarrow (M_1, M_2) : A'_1 \times A'_2} \\ \frac{M[\alpha] : A \Rightarrow M' : B}{M : \forall \alpha. A \Rightarrow \Lambda \alpha. M' : \forall \alpha. B} \quad \frac{M[l] : A \Rightarrow M' : B}{M : \forall l. A \Rightarrow \Lambda l. M' : \forall l. B} \end{array}$$

Figure 4: A type-directed location adjustment to terms

Using the rules, one can derive the example adjustment judgment explained previously for $\lambda^c x. f x$.

Based on the heuristic type compilation with the type-directed location adjustment, we can define the compilation rules of λ_{rpc}^{Links} into λ_{rpc}^\forall as shown in Figure 5.

Basically, the compilation rules forms a transformation $\llbracket \mathcal{D} \rrbracket_{\Delta, Loc} = \mathcal{D}'$ of typing derivations \mathcal{D} in λ_{rpc}^{Links} into those \mathcal{D}' in λ_{rpc}^\forall . Additionally, as contextual information in λ_{rpc}^\forall , the compilation method is defined to take a type environment Δ and a location Loc . The type environment Δ provides how types

$$\begin{aligned}
\llbracket \Gamma \vdash x : A \rrbracket_{\Delta, Loc} &= \frac{x : B \in \Delta}{\Delta \vdash_{Loc} x : B} \\
\llbracket \Gamma \vdash \lambda^a x. M : A \rightarrow B \rrbracket_{\Delta, Loc} &= \text{let } \Delta, x : \llbracket A \rrbracket_a \vdash_a M' : D = \llbracket \Gamma, x : A \vdash M : B \rrbracket_{(\Delta, x : \llbracket A \rrbracket_a), a} \\
&\quad \Delta \vdash_{Loc} \lambda^a x. M' : \llbracket A \rrbracket_a \xrightarrow{a} D \\
\llbracket \Gamma \vdash \lambda x. M : A \rightarrow B \rrbracket_{\Delta, Loc} &= \text{let } l \text{ be fresh} \\
&\quad \text{let } \Delta, l, x : \llbracket A \rrbracket_l \vdash_l M' : D = \llbracket \Gamma, x : A \vdash M : B \rrbracket_{(\Delta, l, x : \llbracket A \rrbracket_l), l} \\
&\quad \frac{\Delta, l \vdash_{Loc} \lambda^l x. M' : \llbracket A \rrbracket_l \xrightarrow{l} D}{\Delta \vdash_{Loc} \Lambda l. \lambda^l x. M' : \forall l. (\llbracket A \rrbracket_l \xrightarrow{l} D)} \\
&\quad \frac{}{\Delta \vdash_{Loc} (\Lambda l. \lambda^l x. M')[Loc] : (\llbracket A \rrbracket_l \xrightarrow{l} D)\{Loc/l\}} \\
\llbracket \Gamma \vdash M N : B \rrbracket_{\Delta, Loc} &= \text{let } \Delta \vdash_{Loc} M' : C \xrightarrow{Loc'} D = \llbracket \Gamma \vdash M : A \rightarrow B \rrbracket_{\Delta, Loc} \\
&\quad \text{let } \Delta \vdash_{Loc} N'_0 : C_0 = \llbracket \Gamma \vdash N : A \rrbracket_{\Delta, Loc} \\
&\quad \text{let } N'_0 : C_0 \Rightarrow N' : C \\
&\quad \Delta \vdash_{Loc} M' N' : D \\
\llbracket \Gamma \vdash (M_1, M_2) : A_1 \times A_2 \rrbracket_{\Delta, Loc} &= \text{let } \Delta \vdash_{Loc} M'_i : A'_i = \llbracket \Gamma \vdash M_i : A_i \rrbracket_{\Delta, Loc} \text{ for } i = 1, 2 \\
&\quad \Delta \vdash_{Loc} (M'_1, M'_2) : (A'_1 \times A'_2) \\
\llbracket \Gamma \vdash \pi_i M : A_i \rrbracket_{\Delta, Loc} &= \text{let } \Delta \vdash_{Loc} M' : A'_1 \times A'_2 = \llbracket \Gamma \vdash M : A_1 \times A_2 \rrbracket_{\Delta, Loc} \\
&\quad \Delta \vdash_{Loc} \pi_i M' : A'_i \\
\llbracket \Gamma \vdash \Lambda \alpha. V : \forall \alpha. A \rrbracket_{\Delta, Loc} &= \text{let } \Delta, \alpha \vdash_{Loc} V' : A' = \llbracket \Gamma, \alpha \vdash V : A \rrbracket_{(\Delta, \alpha), Loc} \\
&\quad \Delta \vdash_{Loc} \Lambda \alpha. V' : \forall \alpha. A' \\
\llbracket \Gamma \vdash M [B] : A\{B/\alpha\} \rrbracket_{\Delta, Loc} &= \text{let } \Delta \vdash_{Loc} M' : \forall \alpha. C = \llbracket \Gamma \vdash M : \forall \alpha. A \rrbracket_{\Delta, Loc} \\
&\quad \Delta \vdash_{Loc} M' [\llbracket B \rrbracket_{Loc}] : C\{\llbracket B \rrbracket_{Loc}/\alpha\}
\end{aligned}$$

Figure 5: A compilation of λ_{rpc}^{Links} into λ_{rpc}^{\forall}

for free variables in λ_{rpc}^{Links} have been heuristically compiled. The location *Loc* informs where terms being compiled are.

One of the most important compilation rules is for λ -applications $M N$. After M and N are compiled into M' and N'_0 , the compiled argument term is adjusted to the argument type of the compiled functional term.

Unannotated λ -abstractions are compiled into location abstractions that are subsequently instantiated with the location of the λ -abstractions.

Note that compiling $\lambda^a x. M$ of type $A \rightarrow B$ at a locational context *Loc* extends the current type environment with a type binding of x and a heuristically compiled argument type, that is, $\llbracket A \rrbracket_a$ or $\llbracket A \rrbracket_l$ for some fresh location variable.

3.1 Problem: Violating Syntactic Restriction on Polymorphism

We have a problem in the formulation of the compilation of λ_{rpc}^{Links} into λ_{rpc}^{\forall} . Recall that type and location abstractions are defined in the form as $\Lambda \alpha. V$ and $\Lambda l. V$ whose bodies are in the form of values. The compilation of arbitrary ranked λ_{rpc}^{Links} terms could produce λ_{rpc}^{\forall} terms that violate this syntactic restriction. For those λ_{rpc}^{Links} terms of polymorphic types under the rank-1 polymorphism, often called

the Hindley-Milner style polymorphism, the proposed compilation rules will always produce legitimate terms keeping the polymorphic restriction in syntax.

The cause of the problem is that the eta-conversion of type and location abstractions leads to type and location applications in their bodies. Let us explain the cause by example. It is legitimate to have an adjustment as

$$\Lambda\alpha.(\lambda^c x.x, 42) : \forall\alpha.(\alpha \xrightarrow{c} \alpha \times Int) \Rightarrow \Lambda\alpha.(\lambda^s y.(\lambda^c x.x) y, 42) : \forall\alpha.(\alpha \xrightarrow{s} \alpha \times Int)$$

when the body of the type abstraction is known. Note that the body of the type abstraction is in the form of values. Otherwise, when it is difficult to know how the body of the type abstraction is structured, it seems only to have

$$f : \forall\alpha.(\alpha \xrightarrow{c} \alpha \times Int) \Rightarrow \Lambda\alpha.(\lambda^s y.(\pi_1 f[\alpha]) y, \pi_2 f[\alpha]) : \forall\alpha.(\alpha \xrightarrow{s} \alpha \times Int)$$

where the presence of $\pi_2 f[\alpha]$ causes the body of the compiled type abstraction to be beyond the form of values.

4 Compiling λ_{rpc}^\forall back to λ_{rpc}^{Links} extended with GADTs

Now we study how to compile λ_{rpc}^\forall terms back to λ_{rpc}^{Links} terms. For this, the presence of the notion of location in λ_{rpc}^\forall should be encoded in λ_{rpc}^{Links} terms. GADTs can be used to encode locations as was discussed in [2].

For the compilation, a GADT, *Location* α , is introduced to λ_{rpc}^{Links} to have two data constructors, *Client* and *Server*:

- *Client* : *Location ClientType*
- *Server* : *Location ServerType*

where *ClientType* and *ServerType* are some types whose inhabitants are of no interest.

Accordingly, λ_{rpc}^{Links} is assumed to be extended with case analysis terms on values. For brevity, we only consider a special case term for the location GADT as

$$\text{case } L \text{ of } Client \rightarrow M; Server \rightarrow N.$$

For example, one can write $\Lambda\alpha.\lambda x.\text{case } x \text{ of } Client \rightarrow M; Server \rightarrow N$ in the extended λ_{rpc}^{Links} , and the term can be of type $\forall\alpha. Location \alpha \rightarrow A$ for some type A . This term can be applied as $(\dots) [ClientType] Client$ or as $(\dots) [ServerType] Server$.

Using this way of encoding locations, $L[Loc] = A$, the compilation of locations *Loc* in λ_{rpc}^\forall into types A in λ_{rpc}^{Links} , can be defined as Figure 6.

$$L[\mathbf{c}] = ClientType \quad L[\mathbf{s}] = ServerType \quad L[l] = Location \alpha_l$$

Figure 6: A location compilation of λ_{rpc}^\forall into types in λ_{rpc}^{Links}

For compiling back, a reverse of the heuristic type compilation is required. The reverse type compilation would have only to erase locations annotated to function types except for location abstractions. Location abstractions $\forall l.A$, which is not expressible in λ_{rpc}^{Links} , would be replaced by a combination of type abstractions and function types $\forall\alpha_l.Location \alpha_l \rightarrow A$ using the location GADT.

Figure 7 defines a reverse type compilation based on the idea explained.

$$\begin{aligned} T[\alpha] &= \alpha & T[base] &= base & T[A \xrightarrow{Loc} B] &= T[A] \rightarrow T[B] \\ T[\forall\alpha.A] &= \forall\alpha.T[A] & T[\forall l.A] &= \forall\alpha_l.Location \alpha_l \rightarrow T[A] \end{aligned}$$

Figure 7: A type compilation of λ_{rpc}^\forall into types in λ_{rpc}^{Links}

Figure 8 shows a term compilation of λ_{rpc}^\forall into λ_{rpc}^{Links} . There are three things to note. Firstly, the definition uses location representation by values. The representation method is defined by the compilation of locations into terms $\llbracket Loc \rrbracket = V$. The client location \mathbf{c} is represented by the data constructor *Client*

$$\begin{aligned}
\llbracket \mathbf{c} \rrbracket &= Client & \llbracket \mathbf{s} \rrbracket &= Server & \llbracket l \rrbracket &= x_l \\
\llbracket x \rrbracket_{\Gamma, Loc, A} &= x \\
\llbracket (M, N) \rrbracket_{\Gamma, Loc, A \times B} &= (\llbracket M \rrbracket_{\Gamma, Loc, A} , \llbracket N \rrbracket_{\Gamma, Loc, B}) \\
\llbracket \pi_i M \rrbracket_{\Gamma, Loc, A_i} &= \pi_i (\llbracket M \rrbracket_{\Gamma, Loc, A_1 \times A_2}) \quad (i=1,2) \\
\llbracket \lambda^{Loc'} x. M \rrbracket_{\Gamma, Loc, A \xrightarrow{Loc'} B} &= \lambda^{Ann \llbracket Loc' \rrbracket} x. \llbracket M \rrbracket_{(\Gamma, x: T \llbracket A \rrbracket), Loc', T \llbracket B \rrbracket} \\
\llbracket M N \rrbracket_{\Gamma, Loc, B} &= \llbracket M \rrbracket_{\Gamma, Loc, A \xrightarrow{Loc} B} \llbracket N \rrbracket_{\Gamma, Loc, A} \quad (\text{local procedure call}) \\
\llbracket M N \rrbracket_{\Gamma, \mathbf{c}, B} &= \llbracket M \rrbracket_{\Gamma, \mathbf{c}, A \xrightarrow{s} B} \llbracket N \rrbracket_{\Gamma, \mathbf{c}, A} \quad (\text{remote procedure call}) \\
\llbracket M N \rrbracket_{\Gamma, \mathbf{s}, B} &= \llbracket M \rrbracket_{\Gamma, \mathbf{s}, A \xrightarrow{c} B} \llbracket N \rrbracket_{\Gamma, \mathbf{s}, A} \quad (\text{remote procedure call}) \\
\llbracket M N \rrbracket_{\Gamma, Loc, B} &= if(\llbracket Loc \rrbracket, if(\llbracket Loc' \rrbracket, f \ arg, f \ arg), if(\llbracket Loc' \rrbracket, f \ arg, f \ arg)) \\
&\quad \text{where } f = \llbracket M \rrbracket_{\Gamma, Loc, A \xrightarrow{Loc'} B} \text{ and } arg = \llbracket N \rrbracket_{\Gamma, Loc, A}. \\
&\quad if(L, M, N) = \text{case } M \text{ of } Client \rightarrow M; Server \rightarrow N. \\
&\quad (\text{local/remote procedure calls depending on conditionals}) \\
\llbracket \Lambda l. V \rrbracket_{\Gamma, Loc, \forall l. B} &= \Lambda \alpha. \lambda \llbracket l \rrbracket. \llbracket V \rrbracket_{\Gamma, Loc, B} \\
\llbracket M \llbracket Loc' \rrbracket \rrbracket_{\Gamma, Loc, B \{ Loc' / l \}} &= \llbracket M \rrbracket_{\Gamma, Loc, B \{ Loc' / l \}} [L \llbracket Loc' \rrbracket] [\llbracket Loc' \rrbracket] \\
\llbracket \Lambda \alpha. V \rrbracket_{\Gamma, Loc, \forall \alpha. B} &= \Lambda \alpha. \llbracket V \rrbracket_{\Gamma, Loc, B} \\
\llbracket M \llbracket A \rrbracket \rrbracket_{\Gamma, Loc, B \{ A / \alpha \}} &= \llbracket M \rrbracket_{\Gamma, Loc, \forall \alpha. B} [T \llbracket A \rrbracket] \\
Ann \llbracket a \rrbracket &= a & Ann \llbracket l \rrbracket &= (unknown)
\end{aligned}$$

Figure 8: A term compilation of λ_{rpc}^{\forall} into terms in λ_{rpc}^{Links}

while \mathbf{s} is so by *Server*. Note that the representation method is reminiscent of one used in the compilation of the polymorphic CS calculus into the untyped CS [3].

Secondly, location abstractions are represented actually by λ -abstractions guarded by type abstractions to introduce a type variable for a parameter of the type *Location* α of the λ -variable. For example,

$$\Lambda l. \lambda^l f. f \ 42$$

would be compiled to

$$\Lambda \alpha. \lambda y_l. \lambda f. f \ 42$$

where the type of y_l is *Location* α . Then an application term in λ_{rpc}^{\forall}

$$(\Lambda l. \lambda^l f. f \ 42) [\mathbf{c}]$$

would be a term in λ_{rpc}^{Links}

$$(\Lambda \alpha. \lambda y_l. \lambda f. f \ 42) [ClientType] \ Client$$

where the type of y_l is *Location* *ClientType*. After applying it to the client location value *Client*, this location can be examined in runtime through a variable y_l bound to the location representation value.

Thirdly, there are four compilation rules for application terms in λ_{rpc}^{\forall} . The first three rules for application terms provides static location contexts. In the static location contexts, there is no need to check the location of functions to invoke by the runtime system for λ_{rpc}^{Links} , though our compilation rules have not expressed this by using different syntactic terms such as $V(W)$ for local procedure calls and $\text{req}(V, W)$ and $\text{call}(V, W)$ for remote procedure calls as in the Client-Server calculus [3]. The last rule for application terms is about dynamic location contexts where location information is provided by two variables (or one constant and one variable) $\llbracket Loc \rrbracket$ and $\llbracket Loc' \rrbracket$ that represent the location of the application term and the location of its function, respectively.

Note that $Ann \llbracket Loc \rrbracket$ determines annotations over locations *Loc*. Location variables disappear from the λ -abstractions after the compilation, but they would remain as term variables as explained previously.

5 Conclusion

This progress report describes an attempt to recover static location contexts in the Links-RPC calculus, which represents the RPC aspect of Links, by compiling to the polymorphic RPC calculus and back to itself.

We have identified a technical problem involving polymorphic constructs in the adjustment of locations to resolve the location discrepancies resulting from our heuristic type compilation. We will continue to study to solve this problem.

References

- [1] Kwanghoon Choi and Byeong-Mo Chang. A theory of RPC calculi for client-server model. *Journal of Functional Programming*, 29:e5, 2019.
- [2] Kwanghoon Choi, James Cheney, Simon Fowler, and Sam Lindley. A polymorphic rpc calculus. *Science of Computer Programming*, 197:102499, 2020.
- [3] Kwanghoon Choi, James Cheney, Sam Lindley, and Bob Reynders. A typed slicing compilation of the polymorphic rpc calculus. submitted, May 2021.
- [4] Ezra K Cooper, Sam Lindley, Philip Wadler, and Jeremy Yallop. Links: Web programming without tiers. In *Proceedings of the 5th International Conference on Formal Methods for Components and Objects*, FMCO’06, pages 266–296, Berlin, Heidelberg, 2007. Springer-Verlag.
- [5] Ezra K. Cooper and Philip Wadler. The rpc calculus. In *Proceedings of the 11th ACM SIGPLAN Conference on Principles and Practice of Declarative Programming*, PPDP ’09, pages 231–242, New York, NY, USA, 2009. ACM.

A The Polymorphic RPC Calculus

This section reminds the reader of the polymorphic RPC calculus [2]. It is a polymorphically typed call-by-value λ -calculus with location annotations on λ -abstractions specifying where to run. The calculus offers the notion of polymorphic location to write polymorphically located functions succinctly, which is convenient for programmers.

A.1 The Syntax and the Semantics

Syntax

Location	a, b	$::=$	$\mathbf{c} \mid \mathbf{s}$
	Loc	$::=$	$a \mid l$
Term	L, M, N	$::=$	$V \mid L M \mid M[A] \mid M[Loc] \mid (L, M) \mid \pi_i(M)$
Value	V, W	$::=$	$x \mid \lambda^{Loc} x.M \mid \Lambda\alpha.V \mid \Lambda l.V \mid (V, W)$

Semantics

(Abs)	$\frac{}{\lambda^b x.M \Downarrow_a \lambda^b x.M}$	(App)	$\frac{L \Downarrow_a \lambda^b x.N \quad M \Downarrow_a W \quad N\{W/x\} \Downarrow_b V}{L M \Downarrow_a V}$
(Tabs)	$\frac{}{\Lambda\alpha.V \Downarrow_a \Lambda\alpha.V}$	(Tapp)	$\frac{M \Downarrow_a \Lambda\alpha.V}{M[B] \Downarrow_a V\{B/\alpha\}}$
(Labs)	$\frac{}{\Lambda l.V \Downarrow_a \Lambda l.V}$	(Lapp)	$\frac{M \Downarrow_a \Lambda l.V}{M[b] \Downarrow_a V\{b/l\}}$
(Pair)	$\frac{L \Downarrow_a V \quad M \Downarrow_a W}{(L, M) \Downarrow_a (V, W)}$	(Proj-i)	$\frac{M \Downarrow_a (V_1, V_2) \quad i \in \{1, 2\}}{\pi_i(M) \Downarrow_a V_i}$

Figure 9: The polymorphic RPC calculus λ_{rpc}^\forall

Figure 9 shows the syntax and semantics of the polymorphic RPC calculus, λ_{rpc}^\forall that allows programmers to use the same syntax of λ -application for both local and remote calls, and allows them to compose differently located functions arbitrarily. An important feature is the notion of location variable l for which a location constant a can be substituted. A syntactic object Loc is either a location constant or a location variable. Assuming the client-server model in the calculus, location constants are either \mathbf{c} denoting client or \mathbf{s} denoting server.

In the syntax, M denotes terms, and V denotes values. Every λ -abstraction $\lambda^{Loc}x.M$ has a location annotation of Loc . By substituting a location b for a location variable annotation, $(\lambda^l x.M)\{b/l\}$ becomes a monomorphic λ -abstraction $\lambda^b x.(M\{b/l\})$. This location variable is abstracted by the location abstraction construct $\Lambda l.V$, and it is instantiated by the location application construct $M[Loc]$. Term applications are denoted by $L M$. The rest of the syntax are straightforward.

The semantics of λ_{rpc}^\forall is defined in the style of a big-step operational semantics whose evaluation judgments, $M \Downarrow_a V$, denote that a term M evaluates to a value V at location a . In the semantics, location annotated λ -abstractions, type abstractions, and location abstractions are all values. So, (Abs), (Tabs), and (Labs) are straightforwardly defined as an identity evaluation relation over them. (App) defines local calls when $a = b$ and remote calls when $a \neq b$ in the same syntax of lambda applications. The evaluation of an application $L M$ at location a performs β -reduction at location b , where a λ -abstraction $\lambda^b x.N$ from L has as an annotation, with a value W from M , and it continues to evaluate the β -reduced term $N\{W/x\}$, which is a substitution of W for x in N , at the same location. The remaining semantics rules are easily understood.

Types

Type	$A, B, C ::= base \mid A \xrightarrow{Loc} B \mid \alpha \mid A \times B \mid \forall \alpha. A \mid \forall l. A$
Type environment	$\Gamma ::= \emptyset \mid \Gamma, x : A \mid \Gamma, \alpha \mid \Gamma, l$

Typing Rules

(T-Var)	$\frac{\Gamma(x) = A}{\Gamma \vdash_{Loc} x : A}$	(T-Abs)	$\frac{\Gamma, x : A \vdash_{Loc} M : B}{\Gamma \vdash_{Loc'} \lambda^{Loc} x. M : A \xrightarrow{Loc} B}$
(T-App)	$\frac{\Gamma \vdash_{Loc} L : A \xrightarrow{Loc'} B \quad \Gamma \vdash_{Loc} M : A}{\Gamma \vdash_{Loc} L M : B}$		
(T-Tabs)	$\frac{\Gamma, \alpha \vdash_{Loc} V : A}{\Gamma \vdash_{Loc} \Lambda \alpha. V : \forall \alpha. A}$	(T-Tapp)	$\frac{\Gamma \vdash_{Loc} M : \forall \alpha. A}{\Gamma \vdash_{Loc} M[B] : A\{B/\alpha\}}$
(T-Labs)	$\frac{\Gamma, l \vdash_{Loc} V : A}{\Gamma \vdash_{Loc} \Lambda l. V : \forall l. A}$	(T-Lapp)	$\frac{\Gamma \vdash_{Loc} M : \forall l. A}{\Gamma \vdash_{Loc} M[Loc'] : A\{Loc'/l\}}$
(T-Pair)	$\frac{\Gamma \vdash_{Loc} L : A \quad \Gamma \vdash_{Loc} M : B}{\Gamma \vdash_{Loc} (L, M) : A \times B}$		
(T-Proj-i)	$\frac{\Gamma \vdash_{Loc} M : A_1 \times A_2 \quad i \in \{1, 2\}}{\Gamma \vdash_{Loc} \pi_i(M) : A_i}$		

Figure 10: A type system for the polymorphic RPC calculus

A.2 The Type System

Figure 10 shows a type system for the polymorphic RPC calculus [2] that can identify remote procedure calls at the type level, supporting location polymorphism. The type language allows function types $A \xrightarrow{Loc} B$. Then every λ -abstraction at unknown location gets assigned $A \xrightarrow{l} B$ using some location variable l . A universal quantifier over a location variable, $\forall l. A$, is also introduced to allow to abstract such occurrences of location variables.

Typing judgments are in the form of $\Gamma \vdash_{Loc} M : A$, saying a term M at location a has type A under a type environment Γ . The location annotation, Loc , is either a location variable or constant. Typing environments Γ have location variables, type variables, and types of variables, as $\{l_1, \dots, l_n, \alpha_1, \dots, \alpha_k, x_1 : A_1, \dots, x_m : A_m\}$. They are used to keep track of a set of free location, type, and value variables in the context of a given term.

The typing rules for the polymorphic RPC calculus are defined as follows. (T-App) is a refinement of the conventional typing rule for λ -applications with respect to the combinations of location Loc (where to evaluate the application) and location Loc' (where to evaluate the function). When $Loc = Loc'$, one can statically decide that it is a local procedure call. Otherwise, Loc is different from Loc' . When both locations are constants as $Loc = a$ and $Loc' = b$, $L M$ is statically found to be a remote procedure call: if $a = \mathbf{c}$ and $b = \mathbf{s}$, it is to invoke a server function from the client, and if $a = \mathbf{s}$ and $b = \mathbf{c}$, it is to invoke a client function from the server. When at least one of them is a location variable, we cannot make a decision statically. (T-Labs) and (T-Lapp) are similar to the typing rules for type abstraction and type application. (T-Labs) checks if its bound location variable does not appear in the type environment and in the contextual location. (T-Lapp) substitutes Loc' for all occurrences of a location variable l on λ -abstractions in M .

The type soundness of the type system for the polymorphic RPC calculus, which was formulated as Theorem 1 and was proved by [2], guarantees that every remote procedure call thus identified statically will never change to a local procedure call under evaluation. This enables compilers to generate call instructions for local calls and network communication for remote calls safely even though both are in the same syntax of lambda applications.

Theorem 1 (Type soundness for λ_{rpc}^\forall [2]). *For a closed term M , if $\emptyset \vdash_a M : A$ and $M \Downarrow_a V$, then $\emptyset \vdash_a V : A$.*

B A Polymorphic RPC Calculus with No Value Restriction

Recall that the polymorphic RPC calculus restricts the body of location and type abstractions to values as $\Lambda l.V$ and $\Lambda \alpha.V$. This is a form of value restriction on polymorphism. The value restriction is well-known in the ML programming language family for the purpose of integrating polymorphism with *effectful* features, such as mutable references, without breaking type safety.

In the polymorphic RPC calculus, value restriction also provides a simple mean to integrate polymorphism with a somewhat unusual effectful feature of the statically located context of application terms.

B.1 Why is naively relaxing value restriction on polymorphism dangerous?

In a naive adoption of the existing typing rules (T-Tabs) and (T-Labs), the location of a polymorphic term Loc becomes that of the body of the polymorphic term.

$$(T\text{-}Tabs) \frac{\Gamma \vdash_{Loc} M : A}{\Gamma \vdash_{Loc} \Lambda \alpha.M : \forall \alpha.A} \quad (T\text{-}Labs) \frac{\Gamma, l \vdash_{Loc} M : A}{\Gamma \vdash_{Loc} \Lambda l.M : \forall l.A}$$

Let us see how this naive typing rules can break the type safety of statically located contexts of applications. Suppose the following piece of a polymorphic term g of type $\forall \alpha. \alpha \xrightarrow{s} \alpha$ is located in the client.

$$\lambda^c z. \dots \text{let } g = \Lambda \alpha. ((\lambda^s x : Int. \lambda^s y : \alpha. y) 123) \text{ in } \dots$$

By the statically located context of application terms, it means that the application of a server function to an integer 123

$$(\lambda^s x : Int. \dots) 123$$

is located at the client. This application is viewed as a remote procedure call since a server function is invoked at the client.

However, this view can be broken when this polymorphic term is used in differently located contexts, that is, at the server that requires this application to be a local procedure call. Consider

$$\dots (\lambda^c f. \textcolor{green}{f} [\textcolor{green}{Int}] 456) g \dots (\lambda^s f. \textcolor{red}{f} [\textcolor{red}{String}] \textcolor{red}{“456”}) g \dots$$

In the green colored part, the polymorphic term happens to be used at the same location as the declared one. So, a view of the application of a server function to an integer 123 as a remote procedure call is preserved. In the red colored part, however, the used location is the server, which is different from the declared location. So, the view as a remote procedure call is broken.

With value restriction on polymorphism, we will never face this problem because both $f [\textcolor{green}{Int}]$ and $f [\textcolor{red}{String}]$ are enforced to be syntactic values that are relocatable everywhere. This explains why relaxing value restriction in the polymorphic RPC calculus can be dangerous to statically located contexts of applications.

This section will discuss an extended polymorphic RPC calculus supporting $\lambda l.M$ and $\Lambda l.M$ with no value restriction on polymorphism. A simple idea is to enforce dynamic location checking whenever such a conflict arises. This enforcement can be implemented by having a fresh location l_0 for typing the body of type and location abstractions as follows.

$$\text{(T-Tabs)} \quad \frac{\Gamma, l_0 \vdash_{l_0} M : A \quad \text{fresh } l_0}{\Gamma \vdash_{Loc} \Lambda \alpha. M : \forall \alpha. A} \quad \text{(T-Labs)} \quad \frac{\Gamma, l, l_0 \vdash_{l_0} M : A \quad \text{fresh } l_0}{\Gamma \vdash_{Loc} \Lambda l. M : \forall l. A}$$

Note that the fresh location variable replaces the lexically inherited location Loc in the naive typing rules. Using this idea, we are able to design a new RPC calculus that relaxes the value restriction on polymorphism.

Users in need of relaxed polymorphism in the RPC calculus have only to adopt the two typing rules (T-Tabs) and (T-Labs) as this. But to prove whether this idea does make the relaxed calculus be type safe in terms of statically located contexts of applications involves more than this. The next section will discuss how to prove the type safety of the relaxed polymorphic calculus.

B.2 Why does relaxing value restriction on polymorphism interest us?

There are two motivations on studying value restriction on polymorphism in the RPC calculus and its relaxation. A motivation comes from a theoretical interest. The existing studies on the typed RPC calculus have not clearly explained what value restriction on polymorphism means to the calculus itself. They explained it indirectly or implicitly by compiling it into client-server calculi where application terms are compiled according to statically determined location contexts. The correctness of the compilation can explain the notion of value restriction in the RPC calculus.

The other motivation comes from a practical purpose. A goal of location inference is to infer as many static location as possible from RPC programs with minimal location annotations like the Links programs. In Links, location annotations are so minimal that programmers might not feel so hard in writing location annotations. We are designing a location inference algorithm for Links-like programs to discover as many static location contexts as possible. One problem with this algorithm is that it requires a program transformation for location adjustment. This can break the value restriction on polymorphism in the RPC calculus. So there is a need to take a look at whether the value restriction can be relaxed enough to permit the transformation required by the location inference algorithm.

B.3 A Relaxed Polymorphic RPC Calculus

Figure 11 presents terms and types in a new relaxed polymorphic RPC calculus together with a dynamic semantics.

Figure 12 defines a set of typing rules.

Syntax

Location	a, b	$::=$	$\mathbf{c} \mid \mathbf{s}$
	Loc	$::=$	$a \mid l$
Term	L, M, N	$::=$	$V \mid L \ M \ @ \ Loc \mid M[A] \mid M[Loc] \mid (L, M) \mid \pi_i(M)$
Value	V, W	$::=$	$x \mid \lambda^{Loc} x. M \mid \Lambda \alpha. \mathbf{ctx} M \mid \Lambda l. \mathbf{ctx} M \mid (V, W)$
Context	\mathbf{ctx}	$::=$	l

Semantics

(Abs)	$\frac{}{\lambda^b x. M \Downarrow_a \lambda^b x. M}$	(App)	$\frac{L \Downarrow_a \lambda^b x. N \quad M \Downarrow_a W \quad N\{W/x\} \Downarrow_b V}{L \ M \ @ \ a \Downarrow_a V}$
(Tabs)	$\frac{}{\Lambda \alpha. \mathbf{ctx} M \Downarrow_a \Lambda \alpha. \mathbf{ctx} M}$	(Tapp)	$\frac{M \Downarrow_a \Lambda \alpha. \mathbf{ctx} N \quad N\{B/\alpha\}\{a/\mathbf{ctx}\} \Downarrow_a V}{M[B] \Downarrow_a V}$
(Labs)	$\frac{}{\Lambda l. \mathbf{ctx} M \Downarrow_a \Lambda l. \mathbf{ctx} M}$	(Lapp)	$\frac{M \Downarrow_a \Lambda l. \mathbf{ctx} N \quad N\{b/l\}\{a/\mathbf{ctx}\} \Downarrow_a V}{M[b] \Downarrow_a V}$
(Pair)	$\frac{L \Downarrow_a V \quad M \Downarrow_a W}{(L, M) \Downarrow_a (V, W)}$	(Proj-i)	$\frac{M \Downarrow_a (V_1, V_2) \quad i \in \{1, 2\}}{\pi_i(M) \Downarrow_a V_i}$

Figure 11: The polymorphic RPC calculus with no value restriction on polymorphism: $\lambda_{rpc}^{\forall, M}$

Types

Type	A, B, C	$::=$	$base \mid A \xrightarrow{Loc} B \mid \alpha \mid A \times B \mid \forall \alpha. \mathbf{ctx} A \mid \forall l. \mathbf{ctx} A$
Type environment	Γ	$::=$	$\emptyset \mid \Gamma, x : A \mid \Gamma, \alpha \mid \Gamma, l$

Typing Rules

(T-Var)	$\frac{\Gamma(x) = A}{\Gamma \vdash_{Loc} x : A}$	(T-Abs)	$\frac{\Gamma, x : A \vdash_{Loc} M : B}{\Gamma \vdash_{Loc'} \lambda^{Loc} x. M : A \xrightarrow{Loc} B}$
(T-App)	$\frac{\Gamma \vdash_{Loc} L : A \xrightarrow{Loc'} B \quad \Gamma \vdash_{Loc} M : A}{\Gamma \vdash_{Loc} L \ M \ @ \ Loc : B}$		
(T-Tabs)	$\frac{\Gamma, \alpha, \mathbf{ctx} \vdash_{\mathbf{ctx}} M : A}{\Gamma \vdash_{Loc} \Lambda \alpha. \mathbf{ctx} M : \forall \alpha. \mathbf{ctx} A}$	(T-Tapp)	$\frac{\Gamma \vdash_{Loc} M : \forall \alpha. \mathbf{ctx} A}{\Gamma \vdash_{Loc} M[B] : A\{B/\alpha\}\{Loc/\mathbf{ctx}\}}$
(T-Labs)	$\frac{\Gamma, l, \mathbf{ctx} \vdash_{\mathbf{ctx}} M : A}{\Gamma \vdash_{Loc} \Lambda l. \mathbf{ctx} M : \forall l. \mathbf{ctx} A}$	(T-Lapp)	$\frac{\Gamma \vdash_{Loc} M : \forall l. \mathbf{ctx} A}{\Gamma \vdash_{Loc} M[Loc'] : A\{Loc'/l\}\{Loc/\mathbf{ctx}\}}$
(T-Pair)	$\frac{\Gamma \vdash_{Loc} L : A \quad \Gamma \vdash_{Loc} M : B}{\Gamma \vdash_{Loc} (L, M) : A \times B}$		
(T-Proj-i)	$\frac{\Gamma \vdash_{Loc} M : A_1 \times A_2 \quad i \in \{1, 2\}}{\Gamma \vdash_{Loc} \pi_i(M) : A_i}$		

Figure 12: A type system for $\lambda_{rpc}^{\forall, M}$