A located lambda calculus

Ezra elias kilty Cooper    Philip Wadler
University of Edinburgh

Abstract
Several recent language designs have offered a unified language for programming a distributed system; we call these “location-aware” languages. These languages provide constructs that allow the programmer to control the location (the choice of host, for example) where a piece of code should run, which can be useful for security or performance reasons. On the other hand, a central mantra of web engineering insists that web servers should be “stateless”; that no “session state” should be maintained on behalf of individual clients—that is, no state that pertains to the particular point of interaction at which a client program resides. Thus far, most implementations of unified location-aware languages have ignored this precept, usually keeping a process for each client running on the server, or otherwise storing state information in memory. We show how to implement a location-aware language on top of the stateless-server model.

1. Introduction
Designing a web server requires thinking carefully about user state and how to manage it. Unlike a desktop application, which deals with one user at a time, or a traditional networked system, which may handle multiple simultaneous requests but in a more controlled environment, a modest web system can expect to deal with tens or hundreds of thousands of users in a day, each one can have multiple windows open on the site simultaneously—and these users can disappear at any time without notifying the server. This makes it infeasible for a web server to maintain state regarding a user’s session. The mantra of web programming is: Get the state out!—get it out of the server and into the client. An efficient web server will respond to each request quickly and then forget about it even quicker.

Nonetheless, several recent high-level programming language designs allow the programmer the illusion of a persistent environment encompassing both client and server. This allows the programmer to move control back and forth freely between client and server, using local resources on either side as necessary, but still expressing the program in one language.

Murphy et al. (2004) introduced a core located calculus, Lambda 5, and Murphy (2007) built upon this a full-fledged programming language, ML5, also showing how to compile it to separate code for client and server. Neubauer and Thiemann (2005) also introduced a variant of ML with location annotations and showed how to perform a splitting transformation to produce code for each location. However, each of these works relied on concurrently-running servers. Ours is the first work we’re aware of that shows how to implement a located language on top of a stateless server model.

Our technique involves three essential transformations: defunctionalization à la Reynolds, CPS transformation, and a “trampoline” that allows tunnelling server-to-client requests within server responses.

We have implemented a version of this feature in the Links language (Cooper et al. 2006). In the current version, only calls to top-level functions can pass control between client and server. Here we provide a formalization and show how to relax the top-level function restriction. In particular, the current, limited version requires just a CPS translation and a trampoline; defunctionalization is needed in implementing nested annotations. This paper This paper presents a simple λ-calculus enriched with location annotations, allowing the programmer to indicate the location where a fragment of code should execute. The semantics of this calculus clarifies where each computation step is allowed to take place.

We then give a translation from this calculus to a first-order abstract machine that models an asymmetrical client–server environment, and show that the translation preserves the locative semantics of the source calculus.

As a stepping stone to the full translation, we formally define defunctionalization. By isolating this phase of the larger translation, we hope to clarify the notation and formal techniques that will be used in the full translation. While there are many compact definitions of CPS transformations in the literature, we did not find such a compact and formal definition of defunctionalization. We give a complete formal definition of defunctionalization in eleven lines.

2. Defunctionalization
Source calculus, $\lambda_{ac}$
Figure 1 shows an entirely pedestrian call-by-value λ-calculus, called $\lambda_{ac}$, defined with a big-step reduction relation $M \Downarrow V$, stating that term $M$ reduces to value $V$. We write $N[V/x]$ for the capture-avoiding substitution of a value $V$ for the variable $x$ in the term $N$, and $N[V_1/x_1, \ldots, V_n/x_n]$ for the simultaneous capture-avoiding substitution of each $V_i$ for the corresponding $x_i$. We let $\sigma$ range over such $n$-ary substitutions. We assume terms are equal up to α-equivalence.

First-order machine
The defunctionalized machine, DM, defined in Figure 2, is a first-order machine, in contrast to $\lambda_{ac}$ which allowed expressions in the function position of an application. In DM, terms, ranged over by $L$, $M$, and $N$, are first-order; they are built from constants $c$, variables $x$, constructor applications $F(M)$, function applications
Defunctionalization is a translation from the terms of \( \lambda x.e \) to the terms and definition-sets of DM; it is defined in Figure 3. From a term \( M \) we compute a defunctionalized term \( [M] \) and a corresponding definition set, \( [M]^{\text{top}} \). Let \( \text{arg} \) be a special reserved variable name not appearing in the source program. The coll function is used by \( [M]^{\text{top}} \) to traverse a term: it applies a function \( f \) to each subterm of its argument, collecting the results.

A special feature of our translation is the use of an injective function that maps source terms into the space of constructor names. We write \( \Gamma \) for the name assigned to the term \( M \) by this function. One example of such a function is the one that collects the names assigned to immediate subterms and uses a hash function when the subterms have a few preliminaries to deal with first.

The central claim of this section is that the defunctionalized terms have a few preliminaries to deal with first. This completes the definition of the reverse translation \( \Gamma^{-1} \). The central claim of this section is that the defunctionalized machine simulates the source calculus under this translation. We have a few preliminaries to deal with first.

\[
[f(\vec{y})] \xrightarrow{[M]} = \lambda x.(\Gamma x) \cdot \Gamma \vec{y}
\]
\[
\vec{y} = \text{fv}(\lambda x.e)
\]
\[
[x] = x
\]
\[
[c] = c
\]
\[
[LM] = \text{apply}([L], [M])
\]
During reduction we may lose subterms which would have given rise to defunctionalized definitions; thus the reduction of a term does not have the same definition-set as its ancestor. Still, all the definitions it needs were generated by the original term; we formalize this as follows.

**Definition 2** (Definition containment). A definition set $\mathcal{D}$ contains $\mathcal{D}'$, written $\mathcal{D} \supseteq \mathcal{D}'$, if

$$\text{cases}(\text{apply}, \mathcal{D}) \supseteq \text{cases}(\text{apply}, \mathcal{D}')$$

**Correctness of the translation from $\lambda\text{MC}$ to DM**

Finally, we can state and prove the simulation of $\lambda\text{MC}$ by DM. First a few preliminaries.

The function $[-]^{-1}$ inverts the pair of functions $[\cdot]$ and $[-]^{\text{top}}$, as follows.

**Observation 1.** If $\mathcal{D} \supseteq [M]^{\text{top}}$ then $[[M]]^{-1}_D = M$. Note that the forward translation replaces each variable bound by a $\lambda$-abstraction with the variable $\text{arg}$, and the reverse translation generates fresh names when generating abstractions. This is acceptable since we identify $\alpha$-equivalent terms.

**Lemma 1** (Substitution-Reverse translation). The reverse translation preserves substitutions. Given definition-set $\mathcal{D}$, term $M$ and value $V$, if $\mathcal{D} \supseteq [M]^{\text{top}}$, then we have $[M[V/x]]^{-1}_D = [[M]^V_x]^{-1}_D$.

**Proof.** By induction on the structure of $M$. The proof of each case is a simple matter of pushing the substitutions down through the terms, applying the inductive hypothesis, and pulling them back up the terms. 

**Proposition 1** (Simulation). Given a $\lambda\text{MC}$ term $M$, DM value $V$ and definitions $\mathcal{D}$ with $\mathcal{D} \supseteq [M]^{\text{top}}$, if $[M] \rightarrow V$ then $M \downarrow [V]^{-1}_D$.

**Proof.** This follows directly from the following lemma.

**Lemma 2.** Given a $\lambda\text{MC}$ term $M$ and a DM value $V$, substitution $\sigma$ and definitions $\mathcal{D}$ with $\mathcal{D} \supseteq [M]^{\text{top}}$,

$$\text{if } [M] \sigma \rightarrow V \text{ then } M[\sigma]^{-1}_D \downarrow [V]^{-1}_D$$

**Proof.** By induction on the length of the reduction $[M] \sigma \rightarrow V$.

We take cases on the structure of the term $M$.

- Case $V$. The reduction is of zero steps, $[V] \sigma \rightarrow [V] \sigma$, and $[[V] \sigma]^{-1}_D = [V] \sigma^{-1}_D$. The judgment $V[\sigma]^{-1}_D \downarrow [V]^{-1}_D$ is by $\text{VALUE}$. 

- Case $LM$. By hypothesis, $[LM] \sigma \rightarrow V$. Recall that $[LM] \sigma = \text{apply}([L] \sigma, [M] \sigma)$.

It must be the case that $[L] \sigma$ and $[M] \sigma$ each reduce to some value, for if not the reduction would get stuck. Further, the value to which $[L] \sigma$ reduces must be a constructor application $F(\bar{V})$, for the same reason. Let $W$ be the value to which $[M] \sigma$ reduces.

Let $x$ be a fresh variable and let $N$ be such that $(F(\bar{y}) \rightarrow [N]\{\text{arg}/x\}) \in \text{cases}(\text{apply}, \mathcal{D})$. We know the body of the case for $F$ has the form $[N]\{\text{arg}/x\}$ because it is in the image of the translation $[-]^{\text{top}}$. As such we have $[F(\bar{V})]^{-1}_D = \lambda x.N\{[\bar{V}]^{-1}_D / \bar{y}\}$.

The reduction follows:

$$\begin{align*}
\text{apply}([L] \sigma, [M] \sigma) &\rightarrow \text{apply}(F(\bar{V}), W) \\
&\rightarrow [N]\{\bar{V}/\bar{y}, W/x\} \quad \text{where } \bar{y} = \text{fv}(N) \\
&\rightarrow V
\end{align*}$$

Now we apply the inductive hypothesis three times. Note that all of these reductions are no longer than the present one, and the definitions in $[[L]^{\text{top}}, [M]^{\text{top}}]$ and $[N]^{\text{top}}$ are all contained in $\mathcal{D}$ (we know $[N]^{\text{top}}$ is contained because $[N]$ appeared as the body of a case in $apply$, and so $\lambda x.N$ must have been a subterm of $LM$), so the inductive hypothesis applies.

From $[L] \sigma \rightarrow F(\bar{V})$ we get $[L][\sigma]^{-1}_D \downarrow \lambda x.N\{[\bar{V}]^{-1}_D / \bar{y}\}$.

From $[M] \sigma \rightarrow W$ we get $M[\sigma]^{-1}_D \downarrow [W]^{-1}_D$.

From $[N]\{\bar{V}/\bar{y}, W/x\} \rightarrow V$ we get $N\{[\bar{V}]^{-1}_D / \bar{y}, [W]^{-1}_D / x\} \downarrow [V]^{-1}$.

By the freshness of $x$, $N\{[\bar{V}]^{-1}_D / \bar{y}, [W]^{-1}_D / x\} \downarrow [V]^{-1}$ follows by $\text{BETA}$. huzzah!
3. The Located \( \lambda \)-calculus

The Located \( \lambda \)-calculus

The located lambda calculus, \( \lambda_{loc} \), is defined in Figure 5. This calculus extends the pedestrian calculus of Section 2 by tagging \( \lambda \)-abstractions with a location; we use the set of locations \( \{ c, s \} \), because we are interested in the client-server setting, but the calculus would be undisturbed by any other choice of location set.

The annotation on a \( \lambda \)-abstraction indicates the location where its body must execute. Thus an abstraction \( \lambda^x.N \) represents a function that, when applied, would evaluate the term \( N \) at the client \( c \), binding the variable \( x \) to the function argument as usual. Constants are assumed to be universal, that is, all locations treat them the same way, and they contain no location annotations.

The semantics, defined in big-step style, uses a judgment of the form \( M \Downarrow_a V \), read “the term \( M \), evaluated at location \( a \), results in value \( V \).” The reader can verify that the body \( N \) of a \( \alpha \)-annotated abstraction \( \lambda^x.N \) is only ever evaluated at location \( a \), and thus the location annotations are honored by the semantics.

Again we write \( N(V/x) \) for the capture-avoiding substitution of a value \( V \) for the variable \( x \) in the term \( N \). We assume terms are equal up to \( \alpha \)-equivalence. The annotation \( a \) on \( \lambda^x.N \) has no effect on the binding behavior of names.

3.1 Client/server machine

Our target abstract machine models a pair of interacting agents, a client and a server, that are each first-order computing machines. Figure 6 defines the machine, called CSM.

Being a first-order machine, the application form \( f(M) \) is \( n \)-ary and allows only a function name, \( f \), in the position function. The machine also introduces constructor applications of the form \( F(M) \), which can be seen as a tagged tuple. Constructor- applications are destructed by the case-analysis form \( \text{case } M \text{ of } A \).

A configuration \( K \) of this machine comes in one of two forms: a client-side configuration \( (M; \cdot) \) consisting of an active client term \( M \) and a quiescent server (represented by the dot), or a server-side configuration \( (E; M) \) consisting of an active server term \( M \) and a suspended client context \( E \), which is waiting for the server’s response. Although the client and server are in some sense independent agents, they interact in a synchronous fashion: upon making a request, the client blocks waiting for the server, and upon completing a request, the server is idle until the next request.

Reduction takes place in the context of a pair of definition sets, one for each agent, thus the reduction judgment takes the form \( K \rightarrow_{c,s} K' \). Each definition \( f(x) = M \) defines the function name \( f \), taking arguments \( x \), to be the term \( M \). The variables \( x \) are thus bound in \( M \). A definition set is only well-formed if it uniquely defines each name. This does not preclude the other definition set, in a pair \( (C, S) \), from also defining the same name.

The reflexive, transitive closure of the relation \( \rightarrow_{c,s} \) is written with a double-headed arrow \( \rightarrow_{c,s} \). This keeps the definition-sets formed throughout the reduction sequence.

Observation 2. CSM reduction steps can be made in any evaluation context: \( (M; \cdot) \rightarrow (N; \cdot) \) implies \( (E[M]; \cdot) \rightarrow (E[N]; \cdot) \) and \( (E; M) \rightarrow (E[N]; \cdot) \) implies \( (E'; E[M]); \cdot) \rightarrow (E'[E]; M) \rightarrow (E'[E]; N) \). This is a direct consequence of the reduction rules, which are already defined in terms of evaluation contexts.

Lemma 3 (Substitution-Evaluation). For any substitution \( \sigma \), we have

\[
(M; \cdot) \rightarrow_{c,s} (N; \cdot) \iff (M\sigma; \cdot) \rightarrow_{c,s} (N\sigma; \cdot) \text{ and } (E; M) \rightarrow_{c,s} (E; N) \iff (E; M\sigma) \rightarrow_{c,s} (E; N\sigma).
\]
The result for sequence. and server terms, respectively. The translation respects this and
case for both locations' abstractions; for its own abstractions it gets
is the location of the context where
The bodies of the apply functions will have a case for each abstraction
appearing in the source term. Each location will have a case for both locations' abstractions; for its own abstractions it gets
\[ (\lambda^a x. N) \overset{\gamma}{\mapsto} \lambda x. N^\gamma(y) \]
\[ x \overset{\gamma}{=} x \]
\[ c \overset{\gamma}{=} c \]
\[ (\lambda^x x. N)^* = \lambda^x x. N^\gamma(y) \]
\[ V^* = V^\gamma \]
\[ (LM)^* = \text{apply}(L^*, M^*) \]
\[ V^\top K = \text{cont}(K, V^\gamma) \]
\[ (LM)^\top K = L^\gamma(\lambda^x x. N^\gamma(y), K) \]

**Figure 7.** Term-level translations from \( \lambda_{bsc} \) to CSM.

\[ \text{coll}_{fa} f LM = f_a(LM) \cup \text{coll}_{fa} f L \cup \text{coll}_{fa} f M \]
\[ \text{coll}_{fa} f \lambda^x x. N = f_a(\lambda^x x. N) \cup \text{coll}_{fa} f N \]
\[ \text{coll}_{fa} f x = f_a(x) \]
\[ \text{coll}_{fa} f c = f_a(c) \]

**Figure 8.** Generic traversal function for \( \lambda_{bsc} \) terms.

**Proof.** We show that \( K_\gamma \overset{\gamma}{\mapsto} K' \gamma \) implies \( K\sigma \overset{\gamma}{\mapsto} K'\gamma \sigma \).
The result for \( \overset{\gamma}{\mapsto} \) then follows by induction on the reduction sequence.

Whether \( K \) is of the form \( (M; \cdots) \) or \( (E; M) \), decompose \( M \)
into the form \( E'[M'] \). The proposition then follows by induction on \( M' \).
The proof of each case is a simple matter of pushing the substitutions down through the terms, applying the inductive hypothesis, and pulling them back up the terms.

**Translation from \( \lambda_{bsc} \) to CSM**

Figures 7–10 give a translation from the client-server \( \lambda_{bsc} \) to CSM. Figure 7 gives term-level translations \( (\lambda^x x. N)^\gamma \), \( (\lambda^x x. N)^\gamma \) and \( (\lambda^x x. N)^{\gamma'} \), which generate values, client terms, and server terms, respectively. The functions \( (\lambda^x x. N)^\gamma \) and \( (\lambda^x x. N)^{\gamma'} \) are only defined for \( \lambda_{bsc} \) client
and server terms, respectively. The translation respects this and applies these functions only to terms of the appropriate sort. The translation assumes that the \( \lambda_{bsc} \) program being translated is a client term. This is in keeping with the client-server paradigm, where the client initiates the interaction.

The resulting terms make use of function definitions for apply, tramp and cont. The necessary definition sets are produced by the top-level translations, \( \Lambda^{\text{e,top}} \) and \( \Lambda^{\text{b,top}} \), defined in Figures 9–
10, making use of the traversal function coll of Figure 8. Intuitively, the apply functions handle all function applications, the cont function handles continuation application, and tramp is the trampoline, which tunnels server-to-client requests through responses. For this translation, let \( a \) and \( k \) be special reserved variable names not appearing in the source program.

The function coll traverses a term in a location-sensitive way: it computes the union of \( f_a(N) \) for each subterm \( N \) of \( M \), where \( a \) is the location of the context where \( N \) appears.

A full definition but for the other's abstractions the case will be a mere stub. This stub dispatches a request to the other location, to apply the function to the given arguments. The cont function is defined only on the server, because it arises from the CPS transformation; it has a case for evaluating the r.h.s. of each server-located application, a case App for applying a function to an argument, and a case for Top which is the continuation for returning a value to the client. These correspond to the defunctionalizations of the abstractions that appear in the classic CPS transformation. Recall the translation for applications.

\[ (LM)^{\text{e,top}} K = L^\gamma(\lambda^x x. N^\gamma(y), \text{coll}\, (\Lambda^{\text{e,top}} A) \)
\]
and recall that we always need a “top-level” continuation \( \lambda x. x \) to extract a value from a CPS term; this corresponds to \( \lambda x. x \).
Note that the traversals used in \([-M]^{\text{top}}\) and \([-M]^{\text{top}}\) pass to coll an outermost current location of \(c\), this reflects the requirement that the source program is a client term.

The \text{tramp} function implements the trampoline. Its protocol is as follows: when the client first needs to make a server call, it makes a request wrapped in \text{tramp}. The server will either complete this call itself, without any client calls, or it will have to make a client call along the way. If it needs to make a client call, it returns a specification of that call as a value \text{Call}(\text{fun}, \text{arg}, k), where \text{fun} and \text{arg} specify the call and \(k\) is the current continuation. The \text{tramp} function recognizes this as the result of the original server call, so it simply returns \(x\). As an invariant, the client always wraps its server-requests in a call to \text{tramp}. This way it can always handle \text{Call} responses.

3.2 Correctness of the translation from \(\lambda_{\text{loc}}\) to CSM

As with the defunctionalization formalization, we again need to compare definition-sets. This time we define the containment relation more generally; it holds just when the names defined in the r.h.s. are all defined in the l.h.s. and upon inspecting corresponding function definitions, either the bodies are identical or they are both case analyses where the l.h.s. contains all the alternatives of the r.h.s.

\textbf{Definition 3} (Definition containment). We say a definition set \(D\) contains \(D'\), written \(D \supseteq D'\), if for each definition \(f(x) = M'\) in \(D'\) there is a definition \(f(x) = M\) in \(D\) and either \(M \equiv M'\) or cases\((f, D) \supseteq \text{cases}(f, D')\).

Figure 11 defines a reverse translation for the \(\lambda_{\text{loc}} \rightarrow\) CSM translation. This translation allows us to retract results from CSM into \(\lambda_{\text{loc}}\). All of the functions used in this translation are parameterized on the definition sets \(C\) and \(S\). To extract the alternatives of case-expressions from function bodies, we use a function cases, defined as follows:

\textbf{Definition 4}. Define cases\((f, D) = A\) when
\[(f(x, y) = \text{case } x \text{ of } A) \in D.\]

Note we are relying on the fact that each of our special functions dispatches on the first of its arguments, whether that be the argument \(\text{fun}\) for \text{apply}, or \(k\) for \text{cont}; the dispatching argument is conveniently always the first.

\textbf{Lemma 4} (Retraction). \(\lambda_{\text{loc}}\) is \(\lambda_{\text{loc}}\) and \(S \supseteq [M]^{\star}\), for each \(M\)
(i) \((M^\star)^{\underline{\lambda_{\text{loc}}}}_C, S = M\),
(ii) \((M^\star)^{\underline{\lambda_{\text{loc}}}}_C, S = M\quad \text{and} \quad \text{for each } K \text{ in CSM: } (M^\top K)^{\underline{\lambda_{\text{loc}}}}_C, S = K^\top M^\star_C, S\)

for those terms on which the relevant operation (respectively \(-\)\(\ast\), \(-(\cdot)\), \((-\cdot)\)) is defined.

\textbf{Proof}. By induction on \(M\).

We omit the \((C, S)\) argument to each of the reverse-translation functions, since it never changes.

- Case \(x\) for (i). Trivial.
- Case \(c\) for (i). Trivial.
- Case \(\lambda x. N\) for (i).

Let \(x\) be a fresh variable.

\[(\lambda x. N)^{**} = (\gamma x. N^\gamma(y))\]
\[\tilde{y} = \text{fv}(\lambda x. N)\]
\[= \lambda x.((N^\gamma k)\{x/y\}^\gamma\{x/y\})\]
\[= \lambda x.((N^\gamma k)(x/x))^\gamma\]
\[= \lambda x.(N^\gamma k)^\gamma\]
\[= \lambda x.N\quad \text{ (ind. hyp.)}\]
\[\text{huzzah!}\]
Case \( \alpha = c \):

\[
(\lambda^\alpha x.N)^* = (\lambda^\alpha x.N^*(\vec{y}))^*
\]

\( \vec{y} = \text{fv}(\lambda x.N) \)

\[
= \lambda x^\alpha.(N^*(\arg/x)/x^\alpha)^*
\]

\[
= \lambda x.N^*
\]

(\text{ind. hyp.})

\[
\lambda x.N
\]

\textit{huzzah!}

• Cases \( V \) for \((-)^* \).

\[
V^{**} = V^{\circ *}
\]

\[
= V^{\circ *}
\]

(\text{ind. hyp.})

\[
= V
\]

\textit{huzzah!}

• Cases \( V \) for \((ii)\).

\[
(V^1 K)^\dagger = (\text{cont}(K, V^{**}))^\dagger
\]

\[
= K^\dagger(V^{**})
\]

(\text{ind. hyp.})

\[
= K^\dagger V
\]

\textit{huzzah!}

• Case \( LM \) for \((ii)\).

\[
((LM)^*)^* = (\text{apply}(L^*, M^*))^*
\]

\[
= L^{**} M^{**}
\]

(\text{ind. hyp.})

\[
LM
\]

\textit{huzzah!}

• Case \( LM \) for \((iii)\).

\[
((LM)^1 K)^\dagger = (L^1(\text{M}^\dagger(\vec{y}, K))^\dagger
\]

\( \vec{y} = \text{fv}(M) \)

(\text{ind. hyp.})

\[
= (\text{M}^\dagger(\text{App}(\arg, k))^\dagger)\{L/\arg\}
\]

(\text{ind. hyp.})

\[
= (\text{M}^\dagger(\text{App}(\arg, k))^\dagger(\text{M}))\{L/\arg\}
\]

\[
= (\text{M}^\dagger(\arg M))\{L/\arg\}
\]

\[
= K^\dagger(LM)
\]

\textit{huzzah!}

\textit{Lemma 6 (Simulation).} For any term \( M \) and substitution \( \sigma \) in \( \lambda \text{sec} \), together with definition sets \( C \) and \( S \) such that \( C \supseteq [M]^{\text{c,top}} \) and \( S \supseteq [M]^{\text{s,top}} \), we have the following implications:

\[
\begin{align*}
\text{(i)} & \quad \text{if } M^* \sigma \rightarrow^c S V \\
\text{then } M^*_{\sigma,S} & \psi_v V_{\sigma,S} \text{ and } \\
\text{(ii)} & \quad \text{if } \text{tramp}(\{\}; (M^1 k)^\sigma) \\
\text{then } M^*_{\sigma,S} & \psi_v V_{\sigma,S}.
\end{align*}
\]

\textit{Proof.} By induction on the length of the CSM reduction sequence. Throughout this proof, we make free use of Observation 2 and Lemma 3, showing (respectively) that we can place a reduction inside an evaluation context to get another reduction, and that we can apply a substitution to a reduction to get another reduction.

In this proof we omit the definitions \( C, S \) on reductions, because they are unchanged through each reduction sequence and on the reverse-translation functions because they are unchanged throughout the recursive calls thereof.

Take cases on the structure of \( M \) and split the conclusion into cases for (i) and (ii); in the case of an application \( LM \), take cases on whether \( L \) reduces to a client abstraction \( \lambda^\alpha x.N \) or to a server abstraction \( \lambda^\alpha x.N \):

• Case \( LM \) for (i) where \( L \) reduces to a client abstraction.

By hypothesis, we have \( (LM)^1 \sigma \rightarrow^c V \). Recall \( (LM)^1 \sigma = \text{apply}(L^*, M^*) \). It must be that \( M^* \sigma \) reduces to a value; call it \( W \). It must be that \( L^* \sigma \) reduces to a value of the form \( F^\alpha(\vec{y}) \).

Let \( N \) be such that \( F^\alpha(\vec{y}) \rightarrow N^*\{\arg/x\} \in \text{cases}(apply, C) \).

(We know it has this form because the cases of \( \text{apply} \) are in the image of the \([-]^\text{c,low} \) translation.) Let \( x \) be a fresh variable.

The reduction follows:

\[
(LM)^* \sigma \rightarrow \text{apply}(L^* \sigma, M^* \sigma) \\
\rightarrow \text{apply}(F^\alpha(\vec{y}), M^* \sigma) \\
\rightarrow \text{apply}(F(\vec{y}), W) \\
\rightarrow N^*\{\vec{y}/\vec{g}, W/x\} \\
\rightarrow V
\]

We have \( N^*\{\vec{y}/\vec{g}, W/x\} = N^*\{\vec{y}/\vec{g}\}\{W/x\} \) due to the freshness of \( x \).

Applying the inductive hypothesis three times, we get

\[
L^*\sigma \psi_v F^\alpha(\vec{y}) = \lambda x^\alpha.N^*(\vec{y}/\vec{g}),
\]

\[
M^* \sigma \psi_v W^* \quad \text{and}
\]

\[
N^*\{\vec{y}/\vec{g}\}\{W^*/x\} \psi_v V^*.
\]

The judgment \( (LM)^* \sigma \psi_v V^* \) follows by \text{Beta}.

• Case \( LM \) for (i) where \( L \) reduces to a server abstraction.

By hypothesis, we have \( (LM)^* \sigma \rightarrow^c V \). Recall \( (LM)^* \sigma = \text{apply}(L^*, M^*) \). It must be that \( M^* \sigma \) reduces to a value; call
it W. It must be that $L^\*\sigma$ reduces to a value of the form $F(\bar{y})$.

Let $N$ be such that

$$(F(\bar{y}) \rightarrow (N^k(\{arg/x\})x)) \in \text{cases}(\text{apply}, S)$$

and

$$(F(\bar{y}) \rightarrow \text{tramp}(\text{req apply} (F(\bar{y}), \text{arg}, \text{Top}))) \in \text{cases}(\text{apply}, C).$$

Let $x$ be a fresh variable. The reduction follows:

$$(LM)^*\sigma = \text{apply}(L^*\sigma, M^*\sigma)$$

$\quad \text{apply}(F(\bar{V}), W)$$

$\quad \text{tramp}(\text{req apply}; (F(\bar{V}), W, \text{Top})))$$

$\quad \text{tramp}(\text{Top})); \text{cont}(\text{Top}(V), V)$$

$\quad \text{V}.$

Applying the inductive hypothesis three times, we get

$$L^*\sigma \downarrow x. (N^k(\bar{V}/\bar{y})(\bar{V}/\bar{y})^\text{eta}.$$

$$M^*\sigma \downarrow ^x. N^*\{\bar{V}^*/\bar{y}/\bar{z}/k\}.$$}

Also $N\{\bar{V}^*/\bar{y}, W^*/x\} = N\{\bar{V}^*/\bar{y}\}$ because of the freshness of $x$.

The judgment $(LM)^*\sigma \downarrow x. V^*$ follows by Beta. huzzah!

• Case $LM$ for (ii) where $L$ reduces to a server abstraction.

By hypothesis, we have

$$\text{tramp}(\text{Top})); \text{cont}(\text{Top}(V), V)$$

$\quad \text{V}.$

Recall that $(LM)^k k = L^1(C^{-\sigma}(\bar{z}, k))$, letting $\bar{z} = \text{inv}(M)$. It must be that $(M^1 k)^\sigma$ reduces to a term of the form $\text{cont}(k, F(\bar{y}))$. Let $N$ be such that $(F(\bar{y}) \rightarrow (N^k(\{arg/x\})x)) \in \text{cases}(\text{apply}, S)$. Let $x$ be a fresh variable. The reduction follows:

$$((LM)^k k)^\sigma = \text{tramp}(\text{Top})); \text{cont}(\text{Top}(V), V)$$

$\quad \text{V}.$

Because of the freshness of $x$, we have $N\{\bar{V}^*/\bar{y}, W^*/x\} = N\{\bar{V}^*/\bar{y}\}$.

Applying the inductive hypothesis three times, we get

$$L^*\sigma \downarrow x. (N^k(\bar{V}/\bar{y}))^\text{eta}.$$

$$M^*\sigma \downarrow ^x. N^*\{\bar{V}^*/\bar{y}/\bar{z}/k\}.$$}

$\quad \text{V}.$

The judgment $(LM)^*\sigma \downarrow x. V^*$ follows by Beta. huzzah!

• Case $V$. Using the substitution lemma (Lemma 5) and the inferences of $(\odot)^1$ to $(\odot)^n$, we get $(V^\sigma)^\sigma = V^\sigma$. In both cases, for server and client, the reduction is of zero steps:

For client, $V^\sigma \rightarrow V^\sigma$. The judgment $V^\sigma \downarrow \text{Cont}(V^\sigma)$ follows by Value.

For server,

$$\text{tramp}(\text{Top})); \text{cont}(\text{Top}(V), V^\sigma).$$

The judgment $V^\sigma \downarrow \text{Cont}(\text{Cont}(V^\sigma))$ follows by Value.

\begin{itemize}
  \item \textbf{4. A richer calculus}:
  
  As a hypothetical language feature, we consider detecting location annotations from abstractions, permitting any term to be wrapped in location brackets.
\end{itemize}
Syntax

- constants  
- variables  
- locations  
- terms  

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$L, M, N$</td>
<td>$\ ::= \ (M)^a \mid \lambda x.N \mid LM \mid V$</td>
</tr>
<tr>
<td>$V, W$</td>
<td>$\ ::= \ \lambda^a x.N \mid x \mid c$</td>
</tr>
</tbody>
</table>

Semantics (big-step)

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ellbracket{V} \Downarrow \ellbracket{a}$ $V$</td>
</tr>
<tr>
<td>$\ellbracket{\lambda x.N} \Downarrow \ellbracket{a} \lambda^a x.N$</td>
</tr>
<tr>
<td>$\ellbracket{L \ellbracket{a} M \ellbracket{b} W} \Downarrow \ellbracket{a} N{W/x} \ellbracket{b} V$</td>
</tr>
<tr>
<td>$\ellbracket{LM} \Downarrow \ellbracket{a} V$</td>
</tr>
<tr>
<td>$\ellbracket{\lambda^a V} \Downarrow \ellbracket{a} V$</td>
</tr>
</tbody>
</table>

Figure 12. The bracket-located lambda calculus, $\lambda_{\ellbracket{a}}$.

\[
\begin{align*}
\llbracket (M)^a \rrbracket_a &= (\lambda^a x.\llbracket M \rrbracket^b)() \quad x \text{ fresh} \\
\llbracket \lambda x.N \rrbracket_a &= \lambda^a x.\llbracket N \rrbracket^b \\
\llbracket \lambda^a x.N \rrbracket_a &= \lambda^a x.\llbracket N \rrbracket^b \\
\llbracket x \rrbracket_a &= x \\
\llbracket c \rrbracket_a &= c
\end{align*}
\]

Figure 13. Translation from $\lambda_{\ellbracket{a}}$ to $\lambda_{\ellbracket{loc}}$.

The calculus $\lambda_{\ellbracket{a}}$ in Figure 12 adds location brackets $\langle \rangle^a$ to $\lambda_{\ellbracket{loc}}$ and allows unannotated $\lambda$-abstractions. The interpretation of a bracketed expression $\langle M \rangle^a$ in a location-$b$ context is a computation that evaluates the term $M$ at location $a$ and returns the value to the location $b$. Note that unannotated $\lambda$-abstractions are not treated as values: all values must be mobile, and yet the interpretation of an unannotated abstraction is that its body will be executed at the location corresponding to the nearest containing annotation. Thus the abstraction itself must get annotated with this location, and the ABSTR rule attaches this annotation when it is not already provided.

Figure 13 gives a translation from $\lambda_{\ellbracket{a}}$ to $\lambda_{\ellbracket{loc}}$. Bracketed terms $\langle M \rangle^a$ are simply treated as applications of located thunks; and as expected, unannotated abstractions $\lambda x.N$ inherit their annotation from the nearest containing brackets.

To argue that this translation is correct in the same way as the previous translation, we would need a reverse translation, but there is a problem: the translation is not injective. We could resort to some hack to distinguish the terms, or we could try to prove a looser relationship, perhaps using a simulation relation. We don’t provide a correctness proof here, but hope that the given interpretation is enough to persuade the reader that the translation is reasonable.

Location brackets such as these may be an interesting language feature, allowing programmers to designate the location of computation of arbitrary terms.

5. Related Work

Location-aware languages  Lambda 5 (Murphy et al. 2004; Murphy 2007) is a small calculus with constructs for controlling the location and movement of terms and values. Lambda 5 offers fine control over the runtime movements of code and data, whereas our calculus uses the simple scope discipline of $\lambda$-binding and is profi-igate with data movements. Like ours, the translation of Lambda 5 to an operational model also involves a CPS translation; and where we have used defunctionalization, it uses closure conversion.

Neubauer and Thiemann give an algorithm for splitting a location-annotated sequential program into separate concurrent programs that communicate over channels (Neubauer and Thiemann 2005). They note that “Our framework is applicable to [the special case of a web application] given a suitable mediator that implements channels on top of HTTP.” The trampoline technique we have given provides that mediator. Neubauer and Thiemann use session types to show that the various processes’ use of channels are type-correct over the course of the interaction. Such a type analysis could be profitably applied to $\lambda_{\ellbracket{loc}}$.

For security purposes, Zdancewic, Grossman and Morrissett developed a calculus with brackets (Zdancewic et al. 1999), which served as the model for our $\lambda_{\ellbracket{a}}$. Their results show how a type discipline, with type translations taking place at the brackets, can be used to prove that certain principals (analogous to locations) cannot inspect certain values passed across an interface. Such a discipline could be applied to our calculus, to address information flow security between client and server.

Defunctionalization  After first being introduced in a lucid but informal account by John Reynolds (Reynolds 1972), defunctionalization has been formalized and verified in a typed setting in several papers (Bell and Hook 1994; Bell et al. 1997; Pottier and Gauthier 1997; Nielsen 2000). These formalizations are complicated by the types; we trade type safety for economy of presentation. Danvy and Nielsen (2001) explore a number of uses and interesting relationships provided by defunctionalization.

Defunctionalization is very similar to lambda-lifting (Johnsson 1985), the essential difference being that lambda-lifting doesn’t rely a closure as a denotable value. Thus it would not be applicable here, where we need to serialize the function to send across the wire.

As noted in the introduction, Murphy (2007) uses closure-conversion in place of our defunctionalization; the distinction here is that the converted closures still contain code pointers, rather than using a stable name to identify each abstraction. These code pointers are only valid as long as the server is actively running, and thus it may be difficult to achieve statelessness with such a system.

Continuation-Passing  The continuation-passing transformation has a long and storied history, going back to the 1970s (Fischer 1972; Plotkin 1975). Our treatment owes much to the presentation and results of “A Reflection on Call-By-Value” (Sabry and Wedler 1997).

6. Conclusions and Future Work

We’ve shown how to compile a located $\lambda$-calculus to an asymmetrical, stateless client-server machine by using a CPS-transformation and “trampoline” effectively to represent the server’s call stack as a value on the client. In the future, we hope to extend the source calculus by adding features such as exceptions and generalizing by allowing each abstraction to consist of a set of permissible locations (rather than a single one). We also hope to implement the “richer calculus” with location brackets in the Links language.

The present work begins with a source calculus with location annotations, but the activity of annotation may burden the program-
mer. Considering that some resources are available only at some locations, it should be possible to automatically assign location annotations so as to reduce communication costs, rather than requiring the programmer to carefully annotate the program. Because the dynamic location behavior of a program may be hard to predict, and because there are a variety of possible communication and computation cost models, and perhaps other issues to consider, such as security, the problem is multifaceted and would be interesting to explore.

References