1 GhostLambda: Lambda-Calculus with Ghost Code

Annotating a program with logical formulae predicting statically it dynamic behaviour is a keystone of the deductive software verification approach.

Typically, these logical formulae are assertions about program assigned variables, loop invariants, recursive function variants, etc.

However, it is often useful to annotate programs with some data that appear in program that appear in assertions but do not affect the program dynamic behaviour in any way.

When a correct-by-construction executable code is extracted from the specified program, this , called *ghost code* disappear together with specification.

In this section we describe a language where one can annotate programs with such *ghost code*. We start by formalizing *ghostλ-calculus*, a tiny language of simply typed λ-calculus enriched with ghost variables and ghost expressions. We then define ghost code *erasure*, which transforms a well-typed ghostLambda term to a term of standard λ-calculus. Finally we state and proof a few basic preservation properties of such translation.

1.1 Syntax and Semantics

The syntax and small-step operational semantics of *ghost-λ* is summarized in Figure 1.

Syntax

$\mathtt{t}$	$::=$		TERMS:
		$()$	*unit*
		$\mathtt{x}_\tau^{\mathfrak{B}}$	*variable*
		$\lambda \mathtt{x}_\tau^{\mathfrak{B}}.\mathtt{t}$	*abstraction*
		$\mathtt{t}\,\mathtt{t}$	*application*
		$\mathtt{ghost\ t}$	*ghost term*
$\mathtt{v}$	$::=$		VALUES:
		$()$	*unit*
		$\lambda \mathtt{x}_\tau^{\mathfrak{B}}.\mathtt{t}$	*abstraction*
$\mathfrak{B}$	$::=$		GHOST STATUS :
		$\top \ (*true*)$	*ghost code*
		$\bot \ (*false*)$	*physical code*
τ	$::=$		TYPES:
		$\mathtt{unit}$	*unit type*
		$\tau^{\mathfrak{B}} \to \tau$	*function type*

Evaluation

$$(\lambda x_\tau^b.\mathtt{t})v \xrightarrow{\epsilon} \mathtt{t}[x_\tau^b \leftarrow v] \qquad \text{(E-AppFun)}$$

$$\mathtt{ghost\ t} \xrightarrow{\epsilon} \mathtt{t} \qquad \text{(E-DeGhost)}$$

$$\frac{\mathtt{t} \to \mathtt{t}'}{\mathtt{E\ t} \to \mathtt{E\ t}'} \qquad \text{(E-Context)}$$

where $\mathtt{E} ::= \square \mid \mathtt{E\ t} \mid \mathtt{v\ E}$

Figure 1: ***ghost-λ* syntax and semantics**

1.2 Typing

The typing relation of *ghost-λ* is summarized in Figure 2.

$$\frac{}{\vdash_{g\lambda} () : (unit, \bot)} \quad \text{(T-Unit)}$$

$$\frac{}{\vdash_{g\lambda} x_{\tau}^{\mathfrak{B}} : (\tau, \mathfrak{B})} \quad \text{(T-Var)}$$

$$\frac{\vdash_{g\lambda} \mathtt{t} : (\tau, \mathfrak{B})}{\vdash_{g\lambda} \mathtt{ghost\ t} : (\tau, \top)} \quad \text{(T-Ghost)}$$

$$\frac{\vdash_{g\lambda} \mathtt{t} : (\tau_2, \mathfrak{B}_2)}{\vdash_{g\lambda} \lambda x_{\tau_1}^{\mathfrak{B}_1}.\mathtt{t} : (\tau_1^{\mathfrak{B}_1} \to \tau_2, \mathfrak{B}_2)} \quad \text{(T-Abs)}$$

$$\frac{\vdash_{g\lambda} \mathtt{t}_1 : (\tau_2^{\mathfrak{B}_2} \to \tau_1, \mathfrak{B}_1) \qquad \vdash_{g\lambda} \mathtt{t}_2 : (\tau_2, \mathfrak{B}'_2) \qquad \mathfrak{B}_2 \Rightarrow \mathfrak{B}'_2}{\vdash_{g\lambda} \mathtt{t}_1\ \mathtt{t}_2 : (\tau_1, \mathfrak{B}_1 \vee (\neg\mathfrak{B}_2 \wedge \mathfrak{B}'_2))} \quad \text{(T-App)}$$

Figure 2: ***ghost-λ*** typing relation

1.3 Ghost Code Erasure

Once we formally defined the simply typed lambda-calculus enriched with ghost expressions, our goal is to show that terms which are not ghost themselves have the same computational behaviour as their translation to lambda-calculus, which preserves the structure of terms except for ghost sub-expressions.

Therefore we need to define at first a type-erasure and a term-erasure translations from gλ-calculus to standard simply typed λ-calculus:

Definition 1.1 (Type-Erasure). We define type-erasure by induction on the structure of types :

$\mathcal{E}_{\top}(\tau) = \mathtt{unit}$
$\mathcal{E}_{\bot}(\mathtt{unit}) = \mathtt{unit}$
$\mathcal{E}_{\bot}(\tau_2^{\mathfrak{B}_2} \to \tau_1) = \mathcal{E}_{\mathfrak{B}_2}(\tau_2) \to \mathcal{E}_{\bot}(\tau_1)$.

Definition 1.2 (Term-Erasure). Let $\mathtt{t}$ be a term such that $\vdash_{g\lambda} \mathtt{t} : (\tau, \mathfrak{B})$ holds. We define term-erasure function by induction on the structure of $\mathtt{t}$

$\mathcal{E}_{\top}(\mathtt{t}) = \mathtt{()}$
$\mathcal{E}_{\bot}(\mathtt{()}) = \mathtt{()}$
$\mathcal{E}_{\bot}(x_{\tau}^{\bot}) = x_{\mathcal{E}_{\bot}(\tau)}$
$\mathcal{E}_{\bot}(\lambda x_{\tau_2}^{\mathfrak{B}_2}.t) = \lambda x_{\mathcal{E}_{\mathfrak{B}_2}(\tau_2)}.\mathcal{E}_{\bot}(t)$
$\mathcal{E}_{\bot}(\mathtt{t}_1\ \mathtt{t}_2) = \mathcal{E}_{\bot}(\mathtt{t}_1)\ \mathcal{E}_{\bot}(\mathtt{t}_2)$.

As it can be seen, the erasure function is a morphism that preserve the structure of operational (not ghost) terms and their types ($\sim \mathcal{E}_{\bot}(\star)$), and sends ghost expressions and types to $\mathtt{()}$ and $\mathtt{unit}$ respectively ($\sim \mathcal{E}_{\top}(\star)$).

1.4 Properties of Ghost Code Erasure

Now that we defined the erasure-translation of $g\lambda$-calculus to λ-calculus, our concern is to prove that evaluation of well-typed operational terms is preserved under erasure. However we need to state a few basic lemmas before we can prove the main theorem.

Lemma 1.1 [SUBSTITUTION UNDER ERASURE]. *If* $\vdash_{g\lambda} t_1 : (\tau_1, \mathfrak{B}_1)$ *and* $\vdash_{g\lambda} v_2 : (\tau_2, \mathfrak{B}_2)$ *hold, then* $\mathcal{E}_{\mathfrak{B}_1}(t_1[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail v_2]) = \mathcal{E}_{\mathfrak{B}_1}(t_1)[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_{\mathfrak{B}_1}(v_2)]$

Proof. First by case analysis on ghost indicator variable $\mathfrak{B}_1$ and then by induction on the structure of $\mathtt{t}_1$

Case $\mathfrak{B}_1 = \top$:

$$\mathcal{E}_\top(\mathtt{t}_1[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail \mathtt{v}_2]) = () = \mathcal{E}_\top(\mathtt{t}_1)[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_\top(\mathtt{v}_2)].$$

Case $\mathfrak{B}_1 = \bot, \quad \mathtt{t}_1 = x^{\mathfrak{B}_2}_{\tau_2}$:

$$\begin{aligned}\mathcal{E}_\bot(x^{\mathfrak{B}_2}_{\tau_2}[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail \mathtt{v}_2]) = \mathcal{E}_\bot(\mathtt{v}_2) &= x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}}[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_\bot(\mathtt{v}_2)] \\ &= \mathcal{E}_\bot(x^{\mathfrak{B}_2}_{\tau_2})[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_\bot(\mathtt{v}_2)]\end{aligned}$$

Case $\mathfrak{B}_1 = \bot, \quad \mathtt{t}_1 = y^\bot_{\tau'_2} \neq x^{\mathfrak{B}_2}_{\tau_2}$:

$$\begin{aligned}\mathcal{E}_\bot(y^\bot_{\tau'_2}[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail \mathtt{v}_2]) = \mathcal{E}_\bot(y^\bot_{\tau'_2}) &= y_{\mathcal{E}_\bot(\tau'_2)} \\ = y_{\mathcal{E}_\bot(\tau'_2)}[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_\bot(\mathtt{v}_2)] &= \mathcal{E}_\bot(y^\bot_{\tau'_2})[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_\bot(\mathtt{v}_2)]\end{aligned}$$

Case $\mathfrak{B}_1 = \bot, \quad \mathtt{t}_1 = \lambda y^{\mathfrak{B}'_2}_{\tau'_2}.t_{11}$ with $y^{\mathfrak{B}'_2}_{\tau'_2}$fresh :

$$\begin{aligned}\mathcal{E}_\bot((\lambda y^{\mathfrak{B}'_2}_{\tau'_2}.t_{11})[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail \mathtt{v}_2]) &= \mathcal{E}_\bot[\lambda y^{\mathfrak{B}'_2}_{\tau'_2}.(t_{11}[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail \mathtt{v}_2])] \\ &= \lambda y_{\mathcal{E}_{\mathfrak{B}'_2}(\tau'_2)}.\mathcal{E}_\bot(\mathtt{t}_{11}[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail \mathtt{v}_2])] \\ &\overset{Ind.Hyp.}{=} \lambda y_{\mathcal{E}_{\mathfrak{B}'_2}(\tau'_2)}.(\mathcal{E}_\bot(\mathtt{t}_{11}))[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_\bot(\mathtt{v}_2)] \\ &= (\lambda y_{\mathcal{E}_{\mathfrak{B}'_2}(\tau'_2)}.\mathcal{E}_\bot(\mathtt{t}_{11}))[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_\bot(\mathtt{v}_2)] \\ &= \mathcal{E}_\bot(\lambda y^{\mathfrak{B}'_2}_{\tau'_2}.t_{11})[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_\bot(\mathtt{v}_2)]\end{aligned}$$

Case $\mathfrak{B}_1 = \bot, \quad \mathtt{t}_1 = \mathtt{t}_{11}\mathtt{t}_{12}$

$$\begin{aligned}\mathcal{E}_\bot(\mathtt{t}_{11}\mathtt{t}_{12}[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail \mathtt{v}_2]) &= \mathcal{E}_\bot(\mathtt{t}_{11}[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail \mathtt{v}_2]\ \mathtt{t}_{12}[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail \mathtt{v}_2]) \\ &= \mathcal{E}_\bot(\mathtt{t}_{11}[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail \mathtt{v}_2])\ \mathcal{E}_\bot(\mathtt{t}_{12}[x^{\mathfrak{B}_2}_{\tau_2} \leftarrowtail \mathtt{v}_2]) \\ &\overset{Ind.Hyp.}{=} (\mathcal{E}_\bot(\mathtt{t}_{11})[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_\bot(\mathtt{v}_2)])(\mathcal{E}_\bot(\mathtt{t}_{12})[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_\bot(\mathtt{v}_2)]) \\ &= \mathcal{E}_\bot(\mathtt{t}_{11}\mathtt{t}_{12})[x_{\mathcal{E}_{\mathfrak{B}_2(\tau_2)}} \leftarrowtail \mathcal{E}_\bot(\mathtt{v}_2)])\end{aligned}$$

□

1.4.1 Typing Erasure

Lemma 1.2 [Translation of Typing Relation].
If $\vdash_{g\lambda}\ t : (\tau, \bot)$ *then* $\vdash_{\lambda}\ \mathcal{E}_{\bot}(t)\ :\ \mathcal{E}_{\bot}(\tau)$.

Proof. By induction on a derivation of the statement $\vdash_{g\lambda}\ \mathcal{E}_{\bot}(\mathtt{t})\ :\ \mathcal{E}_{\bot}(\tau)$. For a given derivation, we proceed by case analysis on the final typing rule used in the proof.

Case T-Unit: $\vdash_{g\lambda}\ () : (\mathtt{unit}, \bot)$

As $\mathcal{E}_{\bot}(()) = ()$ and $\mathcal{E}_{\bot}(\mathtt{unit}) = \mathtt{unit}$ we have immediately $\vdash_{\lambda}\ () : \mathtt{unit}$.

Case T-Var: $\vdash_{g\lambda}\ x^{\bot}_{\tau} : (\tau, \bot)$

As $\mathcal{E}_{\bot}(x^{\bot}_{\tau}) = x_{\mathcal{E}_{\bot}(\tau)}$, we have immediately $\vdash_{\lambda}\ x_{\mathcal{E}_{\bot}(\tau)} : \mathcal{E}(\tau)$.

Case T-Abs: $\vdash_{g\lambda}\ \lambda x^{\mathfrak{B}_1}_{\tau_1}.\mathtt{t} : (\tau^1_1 \to \tau_2, \bot)$ with $\vdash_{g\lambda}\ \mathtt{t} : (\tau_2, \bot)$

By induction hypothesis $\vdash_{\lambda}\ \mathcal{E}_{\bot}(\mathtt{t})\ :\ \mathcal{E}_{\bot}(\tau_2)$. There are two cases to consider, depending on whether the parameter of the abstraction is ghost or not. If $\mathfrak{B}_1 = \top$ then $\mathcal{E}_{\top}(\lambda x^{\top}_{\tau_1}.\mathtt{t}) = \lambda().\mathcal{E}(\mathtt{t})$ and therefore

$$\frac{\vdash_{\lambda}\ \mathcal{E}_{\bot}(\mathtt{t}) : \mathcal{E}_{\bot}(\tau_2)}{\vdash_{\lambda}\ \lambda().\mathcal{E}_{\bot}(\mathtt{t}) : \mathtt{unit} \to \mathcal{E}_{\bot}(\tau_2)} \quad \text{(T-Abs)}$$

Otherwise $\mathfrak{B}_1 = \bot$ and again by the rule T-Abs we obtain :

$$\frac{\vdash_{\lambda}\ \mathcal{E}_{\bot}(\mathtt{t}) : \mathcal{E}_{\bot}(\tau_2)}{\vdash_{\lambda}\ \lambda x_{\mathcal{E}_{\bot}(\tau_1)}.\mathcal{E}_{\bot}(\mathtt{t}) : \mathcal{E}_{\bot}(\tau_1) \to \mathcal{E}_{\bot}(\tau_2)} \quad \text{(T-Abs)}$$

Case T-App: $\vdash_{g\lambda}\ \mathtt{t}_1\ \mathtt{t}_2 : (\tau_1, \mathfrak{B}_1)$ with sub-derivations:
$\vdash_{g\lambda}\ \mathtt{t}_1 : (\tau_2^{\mathfrak{B}_2} \to \tau_1, \mathfrak{B}_1)$
$\vdash_{g\lambda}\ \mathtt{t}_2 : (\tau_2, \mathfrak{B}'_2)$, and constraints:
$\mathfrak{B}_1 \vee (\neg\mathfrak{B}_2 \wedge \mathfrak{B}'_2) = \bot$, $\mathfrak{B}_2 \Rightarrow \mathfrak{B}'_2 = \top$

By lemma's statement, $\mathtt{t}_1\ \mathtt{t}_2$ should not be a ghost term. Therefore $\mathfrak{B}_1 \vee (\neg\mathfrak{B}_2 \wedge \mathfrak{B}'_2) = \bot$. From that and from the rule's premise condition $\mathfrak{B}_2 \Rightarrow \mathfrak{B}'_2 = \top$ we deduce that $\mathfrak{B}_1 = \bot$ and that $\mathfrak{B}_2 \Leftrightarrow \mathfrak{B}'_2 = \top$, so we have to cases two consider.

If $\mathfrak{B}_2 = \mathfrak{B}'_2 = \bot$ then by induction hypotheses: $\vdash_{\lambda}\ \mathcal{E}(\mathtt{t}_1) : \mathcal{E}(\tau_2) \to \mathcal{E}(\tau_1)$ and $\vdash_{\lambda}\ \mathcal{E}(\mathtt{t}_2) : \mathcal{E}(\tau_2)$. Applying T-App rule gives us $\vdash_{\lambda}\ \mathcal{E}(\mathtt{t}_1\ \mathtt{t}_2) : \mathcal{E}(\tau_1)$ (where $\mathcal{E}(\mathtt{t}_1)\mathcal{E}(\mathtt{t}_2) = \mathcal{E}(\mathtt{t}_1\ \mathtt{t}_2)$ as $\mathfrak{B}_1 = \bot$).

If $\mathfrak{B}_2 = \mathfrak{B}'_2 = \top$ then by induction hypothesis, $\vdash_{\lambda}\ \mathcal{E}(\mathtt{t}_1) : \mathtt{unit} \to \mathcal{E}(\tau_1)$. Also by definition of $\mathcal{E}$ we have $\mathcal{E}(\mathtt{t}_2) = ()$, and $\mathcal{E}(\tau_2, \mathfrak{B}'_2) = \mathtt{unit}$. Applying T-App rule gives us

$$\dfrac{\dfrac{}{\vdash_{\lambda}\ \mathcal{E}(\mathtt{t}_1\ \mathtt{t}_2) : \mathcal{E}(\tau_1)} \quad \dfrac{}{\vdash_{\lambda}\ () : \mathtt{unit}}\ \text{(T-Unit)}}{\vdash_{\lambda}\ \mathcal{E}(\mathtt{t}_1\ \mathtt{t}_2) : \mathcal{E}(\tau_1)}\ \text{(T-App)}$$

The case (T-Ghost) as well as any other valid derivation where a typed term is marked as ghost do not satisfy lemma's requirement, so these cases are trivially verified. □

1.4.2 Evaluation Preservation

Theorem 1.3 [Evaluation under erasure]. *For any closed $g\lambda$-term $\mathtt{t}$ such that $\vdash_{g\lambda} \mathtt{t} : (\tau, \bot)$ holds, if $\mathtt{t} \to_{g\lambda} \mathtt{t'}$ for some term $\mathtt{t'}$, then $\mathcal{E}_\bot(\mathtt{t}) \to_\lambda \mathcal{E}_\bot(\mathtt{t}')$.*

Proof. By induction on the evaluation relation of $\mathtt{t} \to_{g\lambda} \mathtt{t'}$.

Case E-AppAbs: $\vdash_{g\lambda} (\lambda x^{\mathfrak{B}_2}_{\tau_2}.\mathtt{t})\mathtt{v} : (\tau_1, \mathfrak{B}_1)$, $\vdash_{g\lambda} \mathtt{v} : (\tau_2, \mathfrak{B}'_2)$, $\mathfrak{B}_1 = \bot$, $\vDash \mathfrak{B}_2 \Leftrightarrow \mathfrak{B}'_2$, $(\lambda x^{\mathfrak{B}_2}_{\tau_2}.\mathtt{t})\mathtt{v} \to_{g\lambda} \mathtt{t}[x^{\mathfrak{B}_2}_{\tau_2} \leftarrow \mathtt{v}]$

$$
\begin{array}{ll}
\mathcal{E}_\bot[(\lambda x^{\mathfrak{B}_2}_{\tau_2}.\mathtt{t})\mathtt{v}] & \\
= \lambda x_{\mathcal{E}_{\mathfrak{B}_2}(\tau_2)}.\mathcal{E}_\bot(\mathtt{t}))\mathcal{E}_\bot(\mathtt{v}) & (\text{as } \mathfrak{B}_1 = \bot) \\
\xrightarrow{\epsilon}_\lambda \mathcal{E}_\bot(\mathtt{t})[x_{\mathcal{E}_{\mathfrak{B}_2}(\tau_2)} \leftarrow \mathcal{E}_\bot(\mathtt{v})] & (\text{head red.}) \\
= \mathcal{E}(\mathtt{t}[x^{\mathfrak{B}_2}_{\tau_2} \leftarrow \mathtt{v}]) & (\text{by lemma 1.1})
\end{array}
$$

Case E-DeGhost: Impossible
Case E-Context:

TODO

□

Lemma 1.3 [Evaluation preservation under erasure]. *For any closed $g\lambda$-term $\mathtt{t}$ such that $\vdash_{g\lambda} \mathtt{t} : (\tau, \bot)$ holds, if $\mathtt{t} \to^* v$ for some value v, then $\mathcal{E}(\mathtt{t}) \to^* \mathcal{E}(v)$.*

Proof. TODO □