A Relational Framework for Higher-Order Shape Analysis

Meta-theory

Calculus λ_R

x,y,z, u	\in	variables	
C	::=	Cons Nil	constructors
v	::=	$\lambda(x: au).e\midv: au\midC\overline{v}$	value
e	::=	$x \mid e x \mid$ let $x = e$ in e	
		match x with $C \overline{x} \Rightarrow e$ else e	
		$\mid e: au$	expression
T	::=	unit int intlist	datatypes
au	::=	$\{\nu: T \mid \phi\} \mid x: \tau \to \tau$	dependent types

. . .

Specification Language

$$\begin{array}{lll} R & \in & relation \ names \\ \phi & ::= & r = r \ | \ r \subset r \ | \ \phi \wedge \phi \ | \ \phi \vee \phi \ | \ \top & type \ refinement \\ r & ::= & R(x) \ | \ r \cup r \ | \ r \times r & relational \ expression \\ \Sigma_R & ::= & \langle R, \tau_R, \overline{C \overline{x} \Rightarrow r} \rangle & relation \ definition \\ & \mid \langle R, \tau_R, R^* \rangle & \\ \theta & ::= & T \ | \ T * \theta & tuple \ sort \\ \tau_R & ::= & \operatorname{intlist} : \to \{\theta\} & relation \ sort \\ \end{array}$$

Figure 1: Language

Core Language Metatheory

To simplify meta-theory, we use a calculus with propositional atoms as our intermediate language. We refer to the language as λ_{ϕ} . The language is described in 3.

For a λ_R type environment Γ , we define its λ_{ϕ} projection (denoted Γ^F) as following:

$$\begin{array}{lcl} (\cdot)^F & = & \cdot \\ (\Gamma, \, x : \{\nu : T \,|\, \phi\})^F & = & \Gamma^F, \, x : \mathcal{F}(T) \\ (\Gamma, \, R :: \tau_R)^F & = & \Gamma^F, \, R : \llbracket \tau_R \rrbracket_L \end{array}$$

LEMMA 1.1. If $\Gamma^F \vdash \nu_1^F : \tau^F$ and $\Gamma^F \vdash \nu_2^F : \tau^F$, then $\gamma \sqcup (\nu_1^F, \odot, \nu_2^F)$ is a λ_ϕ value ν^F such that $\Gamma^F \vdash \nu^F : \tau^F$, where $\odot \in \{\lor, \Rightarrow, \Leftrightarrow\}$

Proof By structural induction over τ^F .

• Case bool : By inversion on type derivation of ν_1^F and ν_2^F , we know that ν_1^F , and ν_2^F are λ_ϕ propositions ϕ_1^F and ϕ_2^F respectively. From the definition, $\gamma_\sqcup(\phi_1^F,\odot,\phi_2^F)=\phi^F\odot\phi^F$,

where $\odot \in \{\lor, \Rightarrow, \Leftrightarrow\}$. Proof follows from the type rules of λ_{ϕ} for propositions.

• Case $T^F \to \tau_1^F$: By inversion on type derivations of ν_1^F and ν_2^F , we know that

$$\nu_1^F = \lambda(k:T^F). \, e_1, \, \text{and} \, \nu_2^F = \lambda(k:T^F). \, e_2,$$

for some k, e_1, e_2 , such that

$$\Gamma^F$$
, $k:T^F \vdash e_1:\tau_1^F$ and Γ^F , $k:T^F \vdash e_2:\tau_1^F$.

Now, by structural induction on ν_1^F and ν_2^F , and eliminating inconsistent cases, we have the following inductive hypothesis

$$\gamma \sqcup (e_1,\odot,e_2)$$
 is a value ν_k^F , and $\Gamma^F,\, k:T^F \vdash \nu_k^F:\, au_1^F$

From first conjunct of IH: $\gamma_{11}(\nu_1^F, \odot, \nu_2^F) = \lambda(k:T^F).\nu_k^F$ is a

From the second conjunct of IH: Observing that this is the premise of type rules for functions in λ_{ϕ} , we conclude that:

$$\Gamma^F \vdash \lambda(k:T^F).\gamma_{\sqcup}(e_1,\odot,e_2): T^F \to \tau_1^F$$

Definition (*JoinType*) JoinType of τ_1^F and τ_2^F is defined by structural recursion on the type structure:

- JoinType(bool,bool) = bool
- $\begin{array}{l} \bullet \ \ JoinType(\mathsf{bool}, T^F \to \tau^F) = T^F \to JoinType(\mathsf{bool}, \tau^F) \\ \bullet \ \ JoinType(T^F \to \tau_1^F, \tau_2^F) = T^F \to JoinType(\tau_1^F, \tau_2^F) \end{array}$

Lemma 1.2. If $\Gamma^F \vdash \nu_1^F : \tau_1^F$ and $\Gamma^F \vdash \nu_2^F : \tau_2^F$, then $\gamma_{\bowtie}(\nu_1^F, \wedge, \nu_2^F)$ is a λ_ϕ value ν^F , such that $\Gamma^F \vdash \nu^F : JoinType(\tau_1^F, \tau_2^F)$

Proof By induction on the structure of ν_1^F and ν_2^F , followed by inversion on their typing derivations.

LEMMA 1.3. if $\Gamma \vdash R$:: intlist : $\rightarrow \{\theta\}$ then, $\Gamma^F \vdash \llbracket R \rrbracket_L$: $A_1 \to \llbracket \theta \rrbracket_L \to bool$

 $\mathbf{Proof} \ [\![\mathsf{intlist}: \to \{\theta\}]\!]_L \stackrel{def}{=} [\![\mathsf{intlist}]\!]_L \to [\![\theta]\!]_L \to bool.$ Since $\mathcal{F}(\text{intlist}) = A_1$, we have that: $[\![R]\!]_L = [\![R :: intlist : \rightarrow \{\theta\}]\!]_L$ $=\eta_{wrap}(R, [[intlist: \rightarrow \{\hat{\theta}\}]]_L)$ $= \eta_{wrap}(R, \{\{\Pi^{child}: \gamma \in \{\emptyset\}\}\}\}_{L}) \to bool).$ But, when $\Gamma^{F} \vdash \phi^{F} : \tau^{F}$, then $\Gamma^{F} \vdash \eta_{wrap}(\phi^{F}, \tau^{F}) : \tau^{F}$. Therefore, $\Gamma^{F} \vdash [\![R]\!]_{L} : A_{1} \to [\![\theta]\!]_{L} \to bool$

LEMMA 1.4. If $\Gamma \vdash r : \theta$, then $\Gamma^F \vdash \llbracket r \rrbracket_L : \llbracket \theta \rrbracket_L \to bool$

Proof By induction on the sort derivation. Cases:

• r = R(x), where $\Gamma \vdash R :: \mathsf{intlist} : \to \{\theta\} \qquad \Gamma \vdash x : \mathsf{intlist}.$

From Lemma 1.3, we know that

1

[Copyright notice will appear here once 'preprint' option is removed.]

Sort Checking Specification Language $\Gamma \vdash r :: \{\theta\}$

Well-Formedness
$$\Gamma \vdash r, \quad \Gamma \vdash \phi, \quad \Gamma \vdash \tau$$

WF-RPRED WF-REF WF-BASE WF-FUN

$$\begin{array}{c|c} \odot \in \{=,\subset\} \\ \hline \Gamma \vdash r_1 :: \theta & \Gamma \vdash r_2 :: \theta \\ \hline \Gamma \vdash r_1 \odot r_2 \end{array} \qquad \begin{array}{c|c} \odot \in \{\land,\lor,\Rightarrow\} \\ \hline \Gamma \vdash \phi_1 & \Gamma \vdash \phi_2 \\ \hline \Gamma \vdash \phi_1 \odot \phi_2 \end{array} \qquad \begin{array}{c|c} \hline \Gamma, \nu : T \vdash \phi \\ \hline \Gamma \vdash \{\nu : T \mid \phi\} \end{array} \qquad \begin{array}{c|c} \Gamma \vdash \tau_1 & \Gamma, x : \tau_1 \vdash \tau_2 \\ \hline \Gamma \vdash x : \tau_1 \rightarrow \tau_2 \end{array}$$

Figure 2: Static semantics of λ_R specification language

Typed Calculus with Propositions (λ_{ϕ})

$$\begin{array}{lll} \boldsymbol{\nu}^F & ::= & \lambda(k:T^F).\,\boldsymbol{\nu}^F \mid \phi^F\\ e & ::= & \boldsymbol{\nu}^F \mid e \; e\\ T^F & ::= & A \mid bool\\ \boldsymbol{\tau}^F & ::= & bool \mid T^F \rightarrow \boldsymbol{\tau}^F \end{array}$$

MSFOL

$$\begin{array}{lll} \phi^L & ::= & \forall (k:T^F).\, \phi^L \ | \ \phi^F \ | \ \phi^L \wedge \phi^L & proposition \\ & | & \phi^L \vee \phi^L \ | \ \phi^L \Rightarrow \phi^L \end{array}$$

variable Auxiliary Functions

Semantics of Relational Expressions $\llbracket r \rrbracket_L$

Semantics of Type Refinements $\llbracket \phi \rrbracket_L$

Figure 3: Semantics of Specification Language

$$\Gamma^F \vdash \llbracket R \rrbracket_L : A_1 \to \llbracket \theta \rrbracket_L \to bool \quad \Gamma^F \vdash x : A_1$$

From function application type rule of λ_ϕ , we have:

$$\Gamma^F \vdash Eval(\llbracket R \rrbracket_L x) : \llbracket \theta \rrbracket_L \to bool$$

• $r = r_1 \cup r_2$, where

$$\Gamma \vdash r_1 :: \theta$$
, and $\Gamma \vdash r_1 :: \theta$

Inductive hypothesis (IH):

$$\Gamma^F \vdash \llbracket r_1 \rrbracket_L : \tau^F$$
, and $\Gamma^F \vdash \llbracket r_2 \rrbracket_L : \tau^F$.

Where, $\tau^F = [\![\theta]\!]_L \to bool$.

We know that $[r_1 \cup r_2]_L \stackrel{def}{=} \gamma_{\sqcup}([r_1]_L, \vee, [r_2]_L)$. From lemma 1.1, proof for this case follows.

• $r = r_1 \times r_2$, where

$$\Gamma \vdash r_1 :: \theta_1, \quad \Gamma \vdash r_1 :: \theta_2, \text{ and } \theta = \theta_1 * \theta_2$$

Inductive hypothesis (IH):

$$\Gamma^F \vdash \llbracket r_1 \rrbracket_L : \tau_1^F$$
, and $\Gamma^F \vdash \llbracket r_2 \rrbracket_L : \tau_2^F$.

Where,
$$\tau_1^F = \llbracket \theta_1 \rrbracket_L \to bool$$
, and $\tau_2^F = \llbracket \theta_2 \rrbracket_L \to bool$.

We know that $[r_1 \times r_2]_L \stackrel{def}{=} \gamma_{\sqcup}([r_1]_L, \wedge, [r_2]_L)$. From Lemma 1.2, we have:

$$\Gamma^{F} \vdash \llbracket r \rrbracket_{L} : JoinType(\tau_{1}^{F}, \tau_{2}^{F}) \Rightarrow \Gamma^{F} \vdash \llbracket r \rrbracket_{L} : \llbracket \theta_{1} \rrbracket_{L} \rightarrow \llbracket \theta_{2} \rrbracket_{L} \rightarrow bool \Rightarrow$$

It remains to prove that $\llbracket \theta_1 * \theta_2 \rrbracket_L = \llbracket \theta_1 \rrbracket_L \to \llbracket \theta_2 \rrbracket_L$, which is trivial to establish by structural induction on θ_1 .

LEMMA 1.5. If $\Gamma \vdash \phi$, then $\llbracket \phi \rrbracket_L$ is an λ_{ϕ} value.

2 2014/3/1

Calculus $\lambda_{\forall R}$

```
\in tuple-sort\ variables
                                                                           x, y, k \in variables
            \in type variables
         ::= C \overline{T} \overline{\theta} \overline{\mathcal{R}} \overline{\nu} \mid \Lambda a. e
                                                                                                          value
          \mathring{\Lambda}t.\Lambda(R::a:\to t).e \mid ...
         ::= eT \mid e\theta \mathcal{R} \mid \mathring{\Lambda}t.\Lambda(R :: a : \rightarrow t).e
                                                                                                expression
                   match x with C \overline{T} \overline{\theta} \overline{\mathcal{R}} \overline{x} \Rightarrow e
                            \verb"else" e \mid \dots
                   T \mid t
Tt
        ::=
T
         ::=
                   'a | 'a list | ...
                                                                                                  datatypes
                   \{\nu:T\,|\,\Phi\}\ |\ \dots
                                                                                        dependent type
                   \forall t. \forall (R :: \ '\mathtt{a} : \rightarrow t). \ \delta \ | \ \tau
                                                                                      parametric type
         ::=
                   \forall'a.\sigma \mid \delta
                                                                                             type scheme
```

Specification Language

```
::=\quad \rho=\rho\ |\ \rho\subset\rho\ |\ \Phi\wedge\Phi
                                                                                      type\ refinement
                      \Phi \lor \Phi \mid \top
            \mathcal{R}(x) \mid \rho \cup \rho \mid \rho \times \rho
                                                                                         rel.\ expression
           ::=
ρ
                     \mathcal{R}T \mid \mathcal{R}\theta\mathcal{R} \mid R
\mathcal{R}
           ::=
                                                                                            instantiation
                    t \mid t * \theta \mid \dots
                                                                                                   tuple \ sort
           ::= \forall t. ('a : \rightarrow t) : \rightarrow ('a list : \rightarrow \theta)
                                                                                             relation\ sort
            \forall a. \ \tau_R \mid \tau_R
                                                                                               sort scheme
\sigma_R
          ::=
\Sigma_R ::= \langle R, R_p, \sigma_R, \overline{C\overline{x} \Rightarrow r} \rangle
                                                                                          rel.\ definition
                     \langle R, R_p, \sigma_R, \mathcal{R}^* \rangle \mid \dots
```

Figure 4: $\lambda_{\forall R}$ - Complete calculus with parametric relations

```
R(x) \mid r \times r
               \lambda(\overline{x:T}).r
                                                       transformer
F_R
       ::=
       ::=
               \mathtt{bind}\left(R(x),F_{R}\right)
                                                    bind\ expression
e_b
               \lambda(x:T). bind (R(x),F_R)
                                                   bind abstraction
E_b
       ::=
               \hat{R} = E_b
                                                       bind equation
       ::=
                                                    bind\ definition
       ::=
               \lambda R. E_b
```

Figure 6: Bind Syntax

Proof By induction on the derivation of well-formedness judgement. Cases can be proved by straightforward application of Lemmas 1.4, 1.1 and 1.2.

LEMMA 1.6. If $\Gamma \vdash \phi$, then $\gamma_{\forall}(\llbracket \phi \rrbracket_L)$ is an MSFOL formula.

Proof From lemma 1.5, we know that $[\![\phi]\!]_L$ is an λ_ϕ value ν^F . Next, by structural induction on ν^F , we prove that $\gamma_\forall(\nu^F)$ is an MSFOL formula.

LEMMA 1.7. (Completeness of semantics) For every type refinement ϕ , if $\Gamma \vdash \phi$, then compile (Γ, ϕ) terminates and produces an MSFOL formula.

Proof From lemma 1.6

THEOREM 1.8. (**Decidability**) Type checking in λ_R is decidable.

Proof Follows from Lemma 1.7 and decidability proof of EPR logic. ■

2. Parametric Language Meta-theory

2.1 Decidability of Type Checking

Type refinements (Φ) in $\lambda_{\forall R}$ can be elaborated to a conjunction of bind equations representing semantics of instantiated relations, and

a λ_R type refinement (ϕ) . The relevant rewrite rules are shown in figure 5.

THEOREM 2.1. (**Decidability**) *Type checking in* $\lambda_{\forall R}$ *is decidable.*

Proof Follows from 1. completeness of our rewrite rules for well-formed refinements, 2. Completeness of encoding bind equations in MSFOL, 3. Decidability proof of EPR logic, to which bind equations are compiled to, and decidability result (Theorem 1.8) for λ_R .

3 2014/3/1

Kind Rules
$$\Gamma \vdash r :: \delta_R$$

K-PARAM-REL K-BIND K-INST

Rewrite Rules for Type Refinements $\Phi \hookrightarrow \psi \land \phi$

$$\psi ::= true \mid (R = E_b) \wedge \psi \quad bind \ equations$$

RW-REF RW-REF-ATOM RW-PARAM-REL

Semantics of Bind Equations $\llbracket \psi
rbracket_L$

$$\begin{split} \llbracket R_2 &= \lambda(x:T_1). \ \operatorname{bind} \left(R_1(x), \lambda(\overline{k:T_2}).r\right) \rrbracket_L \quad \overset{def}{=} \quad \forall (x:\llbracket T_1 \rrbracket_L). \ \gamma_{\Rightarrow} (\llbracket R_1(x) \rrbracket_L, \ \forall (\overline{(k:\llbracket T_2 \rrbracket_L)} \llbracket r \rrbracket_L, \ \llbracket R_2(x) \rrbracket_L) \\ & \wedge \quad \forall (x:\llbracket T_1 \rrbracket_L). \ \gamma_{\Leftarrow} (\llbracket R_1(x) \rrbracket_L, \ \forall (\overline{k:\llbracket T_2 \rrbracket_L}) \llbracket r \rrbracket_L, \ \llbracket R_2(x) \rrbracket_L) \\ & \gamma_{\Rightarrow} (\forall (\overline{k:T_1^F}).\phi_1^F, \ \forall (\overline{k:T_1^F}).\forall (\overline{j:T_2^F}).\phi_2^F, \ \nu^F) \quad \hookrightarrow \quad \forall (\overline{k:T_1^F}). \ \forall (\overline{j:T_2^F}). \ \forall (\overline{j:T_2$$

Figure 5: Semantics of bind equations for parametric relations in $\lambda_{\forall R}$

4 2014/3/1