

Formal Definition of SWIN language

Chenglong Wang Jun Li Yingfei Xiong Zhenjiang Hu

September 12, 2014

1 Featherweight Java

1.1 Syntax

Class Declaration

$$CL ::= \text{class } C \text{ extends } C \{ \bar{C} \bar{f}; K \bar{M} \}$$

Constructor Declaration

$$K ::= C(\bar{C} \bar{f}) \{ \text{super}(\bar{f}); \text{this}.\bar{f} = \bar{f} \}$$

Method Declaration

$$M ::= C m(\bar{C} \bar{x}) \{ \text{return } t; \}$$

Term

$$t ::= x \mid t.f \mid t.m(\bar{t}) \mid \text{new } C(\bar{t}) \mid (C) t$$

1.2 Type System

1.2.1 Subtyping

$$\frac{}{C <: C} \text{ (S-SELF)} \qquad \frac{C <: D \quad D <: E}{C <: E} \text{ (S-TRANS)}$$

$$\frac{CT(C) = \text{class } C \text{ extends } D \{ \dots \}}{C <: D} \text{ (S-DEF)}$$

1.2.2 Auxiliary Functions

$$\frac{}{\text{fields}(\text{Object}) = \{ \}} \text{ FIELD-OBJECT}$$

$$\frac{\begin{array}{l} CT(C) = \text{class } C \text{ extends } D \{ \bar{C} \bar{f}; K \bar{M} \} \\ \text{fields}(D) = \bar{D} \bar{g} \end{array}}{\text{fields}(C) = \bar{D} \bar{g}, \bar{C} \bar{f}} \text{ (FIELD-LOOKUP)}$$

$$\frac{\begin{array}{l} CT(C) = \text{class } C \text{ extends } D \{ \bar{C} \bar{f}; K \bar{M} \} \\ B m(\bar{B} \bar{x}) \{ \text{return } t; \} \in \bar{M} \end{array}}{\text{mtype}(m, C) = \bar{B} \rightarrow B} \text{ (METHOD-LOOKUP1)}$$

$$\frac{\begin{array}{l} CT(C) = \text{class } C \text{ extends } D \{ \bar{C} \bar{f}; K \bar{M} \} \\ m \text{ is not defined in } \bar{M} \end{array}}{\text{mtype}(m, C) = \text{mtype}(m, D)} \text{ (METHOD-LOOPUP2)}$$

$$\frac{\text{mtype}(m, D) = \bar{D} \rightarrow D_0 \text{ implies } \bar{C} = \bar{D} \text{ and } C_0 = D_0}{\text{override}(m, D, \bar{C} \rightarrow C_0)} \text{ (OVERRIDE)}$$

1.2.3 Typing

$$\frac{x : C \in \Gamma}{\Gamma \vdash x : C} \text{ (FJ-VAR)}$$

$$\frac{\Gamma \vdash t_0 : C_0 \quad \text{fields}(C_0) = \bar{C} \bar{f}}{\Gamma \vdash t_0.f_i : C_i} \text{ (FJ-FIELD)}$$

$$\frac{\Gamma \vdash t_0 : C_0 \quad \text{mtype}(m, C_0) = \bar{D} \rightarrow C \quad \Gamma \vdash \bar{t} : \bar{C} \quad \bar{C} <: \bar{D}}{\Gamma \vdash t_0.m(\bar{t}) : C} \text{ (FJ-INVK)}$$

$$\frac{\text{fields}(C) = \bar{D} \bar{f} \quad \Gamma \vdash \bar{t} : \bar{C} \quad \bar{C} <: \bar{D}}{\Gamma \vdash \text{new } C(\bar{t}) : C} \text{ (FJ-NEW)}$$

$$\frac{\Gamma \vdash t_0 : D \quad D <: C}{\Gamma \vdash (C)t_0 : C} \text{ (FJ-UCAST)}$$

$$\frac{\Gamma \vdash t_0 : D \quad C <: D \quad C \neq D}{\Gamma \vdash (C)t_0 : C} \text{ (FJ-DCAST)}$$

$$\frac{\Gamma \vdash t_0 : D \quad C \not<: D \quad D \not<: C \quad \text{stupid warning}}{\Gamma \vdash (C)t_0 : C} \text{ (FJ-SCAST)}$$

$$\frac{\bar{x} : \bar{C}, \text{this} : C \vdash t_0 : E_0 \quad E_0 <: C_0 \quad \text{CT}(C) = \text{class } C \text{ extends } D \{ \dots \} \quad \text{override}(m, D, \bar{C} \rightarrow C_0)}{C_0 \text{ m } (\bar{C} \bar{x}) \{ \text{return } t_0; \} \text{ OK in } C} \text{ (FJ-M-OK)}$$

$$\frac{K = C(\bar{C} \bar{f}) \{ \text{super}(\bar{f}); \text{this}.\bar{f} = \bar{f} \} \quad \text{fields}(D) = \bar{D} \bar{g} \quad \bar{M} \text{ OK in } C}{\text{class } C \text{ extends } D \{ \bar{C} \bar{f}; K \bar{M} \} \text{ OK}} \text{ (FJ-C-OK)}$$

2 SWIN

2.1 Syntax

Π	$::= \{ \bar{\pi} \}$	SWIN program
π	$::= (\bar{d}) [l : C_1 \rightarrow r : C_r]$	rule
d	$::= x : C_1 \hookrightarrow C_2$	variable declaration
l	$::= x.f \mid \text{new } C(\bar{x}) \mid x.m(\bar{x})$	code pattern
r	$::= t$	FJ term

2.2 Environment

$$\begin{array}{ll} \text{API} ::= \{ \overline{\text{CL}} \} & \text{API definition} \\ \text{E} ::= \{ \overline{x : C_1 \hookrightarrow C_2} \} & \text{SWIN typing context} \end{array}$$

2.3 Evaluation Rules

$$\begin{array}{c} \frac{\text{CL} = \text{class } C_1 \text{ extends } C_2 \{ \overline{C_1} \ \overline{f_1}; \ K \ \overline{M} \}}{\Pi(\text{CL}) = \text{class } \Pi(C_1) \text{ extends } \Pi(C_2) \{ \Pi(\overline{C_1}) \ \overline{f_1}; \ \Pi(K) \ \overline{\Pi(M)} \}} \quad (\text{E-DECLARATION}) \\ \\ \frac{K = C_1 (\overline{C_2} \ \overline{f_2}) \{ \text{super}(\overline{f_3}); \ \text{this}.\overline{f_1} = \overline{f_j} \}}{\Pi(K) = \Pi(C_1) (\Pi(\overline{C_2}) \ \overline{f_2}) \{ \text{super}(\overline{f_3}); \ \text{this}.\overline{f_1} = \overline{f_j} \}} \quad (\text{E-CONSTRUCTOR}) \\ \\ \frac{M = C_1 \ m(\overline{C_m} \ \overline{x}) \{ \text{return } t; \}}{\Pi(M) = \Pi(C_1) \ m(\Pi(\overline{C_m}) \ \overline{x}) \{ \text{return } \Pi(t); \}} \quad (\text{E-METHOD}) \\ \\ \frac{C_0 \hookrightarrow C_1 \in \mathbf{TypeMapping}(\Pi)}{\Pi(C_0) = C_1} \quad (\text{E-CLASS}) \\ \\ \frac{\forall C. C_0 \hookrightarrow C \notin \mathbf{TypeMapping}(\Pi)}{\Pi(C_0) = C_0} \quad (\text{E-ALTER-CLASS}) \\ \\ \frac{}{\Pi(x) = x} \quad (\text{E-T-VALUE}) \\ \\ \frac{(\overline{x : C_1 \hookrightarrow C_2})[\ \overline{x.f : C_1} \rightarrow \ \overline{r : C_r}] \in \Pi \quad \text{Type}(t) < C_1}{\Pi(t.f) = [\ \overline{x} \rightarrow \Pi(t) \]r} \quad (\text{E-T-FIELD}) \\ \\ \frac{}{\Pi((C) \ t) = (\Pi(C)) \ \Pi(t)} \quad (\text{E-T-CAST}) \\ \\ \frac{(\overline{d})[\ \text{new } C_0(\ \overline{x} \) : C_1 \rightarrow \ \overline{r : C_r}] \in \Pi \quad \{ \ \overline{x : C_1 \hookrightarrow C_2} \} \subseteq \overline{d} \quad \text{Type}(\overline{t_u}) <: \overline{C_1}}{\Pi(\text{new } C_0(\overline{t_u})) = [\ \overline{x} \rightarrow \Pi(\overline{t_u}) \](r)} \quad (\text{E-T-NEW}) \\ \\ \frac{(\overline{d})[\ \overline{x_0.m_0}(\ \overline{y} \) : C_1 \rightarrow \ \overline{r : C_r}] \in \Pi \quad \{ \overline{y : C_1 \hookrightarrow C_2}, \ \overline{x_0 : C_3 \hookrightarrow C_4} \} \subseteq \overline{d} \quad \text{Type}(t_0) <: C_3 \quad \text{Type}(\overline{t_u}) <: \overline{C_1}}{\Pi(t_0.m_0(\overline{t_u})) = [\ \overline{x_0} \rightarrow \Pi(t_0), \ \overline{y} \rightarrow \Pi(\overline{t_u}) \](r)} \quad (\text{E-T-INVOKE}) \\ \\ \frac{\text{no other inference rule can be applied}}{\Pi(\text{new } C_0(\overline{t_u})) = \text{new } C_0(\ \Pi(\overline{t_u}) \)} \quad (\text{E-ALTER-NEW}) \\ \\ \frac{\text{no other inference rule can be applied}}{\Pi(t_0.m_0(\overline{t_u})) = \Pi(t_0).m(\ \Pi(\overline{t_u}) \)} \quad (\text{E-ALTER-INVOKE}) \\ \\ \frac{\text{no other inference rule can be applied}}{\Pi(t.f) = \Pi(t).f} \quad (\text{E-ALTER-FIELD}) \end{array}$$

2.4 Auxiliary Functions

$$\text{TypeMapping}([(\bar{x} : \overline{C_1 \hookrightarrow C_2}) [1 : C_1 \rightarrow r : C_r]]) = \{C_1 \hookrightarrow C_r\} \cup \{ \overline{C_1 \hookrightarrow C_2} \}$$

$$\text{TypeMapping}(\{\bar{\pi}\}) = \bigcup_{\pi} (\text{TypeMapping}(\pi)) \quad (\text{Extract type migration information})$$

$$\text{Decl}(\text{class } C \text{ extends } D \{ \dots \}) = C \quad (\text{Extract the declared class name})$$

2.5 Type Checking Rules

$$\frac{\begin{array}{c} x : C_1 \hookrightarrow C'_1, \bar{y} : \overline{C_2 \hookrightarrow C'_2} \in E \quad \text{class } C_1 \text{ extends } D \{ \bar{C} \bar{f}; K \bar{M} \} \in \text{API}_s \\ C_d \text{ m}(\bar{C}_s \bar{u}) \{ \dots \} \in \bar{M} \quad \bar{C}_2 <: \bar{C}_s \end{array}}{E \vdash_1 x.m(\bar{y}) : C_d} \quad (\text{T-L1})$$

$$\frac{\text{class } C \{ \bar{C} \bar{f}; C(\bar{C}_s \bar{u}) \{ \dots \} \bar{M} \} \in \text{API}_s \quad \bar{x} : \overline{C_1 \hookrightarrow C'_1} \in E \quad \bar{C}_1 <: \bar{C}_s}{E \vdash_1 \text{new } C(\bar{x}) : C} \quad (\text{T-L2})$$

$$\frac{E = \{ \bar{x} : \overline{C \hookrightarrow D} \} \quad \{ \bar{x} : \bar{D} \} \vdash_{\text{FJ}}^{\text{API}_d} t : C_d}{E \vdash_r t : C_d} \quad (\text{T-R})$$

$$\frac{\{ \bar{x} : \overline{C \hookrightarrow D} \} \vdash_1 1 : C_1, \{ \bar{x} : \overline{C \hookrightarrow D} \} \vdash_r r : C_2}{[\{ \bar{x} : \overline{C \hookrightarrow D} \} 1 : C_1 \rightarrow r : C_2] \text{ ok}} \quad (\text{T-}\pi)$$

$$\text{RuleOK}(\Pi) = \forall \pi. (\pi \in \Pi \Rightarrow \pi \text{ ok})$$

$$\text{ConstrCover}(\Pi, \text{API}_s, \text{API}_d) =$$

$$\begin{array}{l} \forall C_1, \bar{C}. (\text{class } C_1 \text{ extends } _ \{ C_1(\bar{C} _) \dots \} \in (\text{API}_s - \text{API}_d) \\ \Rightarrow \exists C_2, \bar{C}', \bar{x}, r. ((\bar{x} : \overline{C \hookrightarrow C'}) [\text{new } C_1(\bar{x}) : C_1 \rightarrow r : C_2] \in \Pi)) \end{array}$$

$$\text{MethCover}(\Pi, \text{API}_s, \text{API}_d) =$$

$$\begin{array}{l} \forall C_1, C_2, m, \bar{C}. (\text{class } C_1 \text{ extends } _ \{ C_2 \text{ m}(\bar{C} _) \{ \dots \} \dots \} \in (\text{API}_s - \text{API}_d) \\ \Rightarrow \exists x, \bar{y}, C'_1, C'_2, \bar{C}', r. ((x : C_1 \hookrightarrow C'_1, \bar{y} : \overline{C \hookrightarrow C'}) [x.m(\bar{y}) : C_2 \rightarrow r : C'_2] \in \Pi)) \end{array}$$

$$\text{FieldCover}(\Pi, \text{API}_s, \text{API}_d) =$$

$$\begin{array}{l} \forall C_1, C_2, f. (\text{class } C_1 \text{ extends } _ \{ C f; \dots \} \in (\text{API}_s - \text{API}_d) \\ \Rightarrow \exists x, C'_1, C'_2. ((x : C_1 \hookrightarrow C'_1) [x.f : C_2 \rightarrow r : C'_2] \in \Pi)) \end{array}$$

$$\text{MapChecking}(\Pi, \text{API}_s, \text{API}_d) =$$

$$\begin{array}{l} \forall C, D. (C \hookrightarrow D \in \text{TypeMapping}(\Pi) \\ \Rightarrow (\exists CL \in \text{API}_s \cap \text{API}_d. (\text{Decl}(CL) = C \wedge D = C)) \\ \vee (\exists CL \in \text{API}_s - \text{API}_d. (\text{Decl}(CL) = C))) \end{array}$$

$$\text{Subtyping}(\Pi, \text{API}_s, \text{API}_d) =$$

$$\forall C_i, D_i, C_j, D_j. (C_i \hookrightarrow D_i, C_j \hookrightarrow D_j \in \text{TypeMapping}(\Pi) \Rightarrow (C_i <: C_j \Rightarrow D_i <: D_j))$$

$$\text{TypeSafe}(\Pi, \text{API}_s, \text{API}_d) =$$

$$\begin{array}{l} \text{RuleOK}(\Pi) \wedge \text{ConstrCover}(\Pi, \text{API}_s, \text{API}_d) \wedge \text{MethCover}(\Pi, \text{API}_s, \text{API}_d) \\ \wedge \text{FieldCover}(\Pi, \text{API}_s, \text{API}_d) \wedge \text{MapChecking}(\Pi, \text{API}_s, \text{API}_d) \wedge \text{Subtyping}(\Pi, \text{API}_s, \text{API}_d) \end{array}$$

3 Metatheory

Lemma 1 (Typing Context). *Given a SWIN program Π acting on API_s to API_d , suppose the typing context for a term t is $\Gamma_s = \bar{x} : \bar{C}$, then the typing context for $\Pi(t)$ is $\Gamma_d = \bar{x} : \overline{\Pi(C)}$.*

Proof. According to the FJ typing rules, the typing context will not change once it is created in the rule FJ-M-OK. For the typing context Γ , except the variable `this`, all other variables in the typing context are bounded in the definition of a method M .

Induction on Γ . Suppose $\Gamma = \bar{y} : \bar{D}, x : C = \Gamma_1, x : C$, then $\Pi(\Gamma_1) = \bar{y} : \overline{\Pi(D)}$.

There are two cases for $x : C$

- $x = \text{this}$ in Γ . The type C is a client defined class type, so $C \notin \text{TypeMapping}(\Pi)$. According to the rule E-ALTER-CLASS, $\Pi(C) = C$, then we have $\Pi(\Gamma) = \Pi(\Gamma_0), x : \Pi(C) = \bar{y} : \overline{\Pi(D)}, \text{this} : C$.
- x is an argument in method declaration. According to the rule E-METHOD, after transformation, the type of x in the definition is $\Pi(C)$, thus $\Gamma = \Pi(\Gamma_0), x : \Pi(C) = \bar{y} : \overline{\Pi(D)}, x : \Pi(C)$.

With these two cases proved, the lemma is proved. \square

Lemma 2 (Subtyping). *Suppose Π passes SWIN type checking rules, and it transforms an FJ program with API_s to a new program with API_d , then:*

$C_1 <: C_2$ in old program $\implies \Pi(C_1) <: \Pi(C_2)$ in the transformed program.

Proof. First, we suppose $C_1 <: C_2$, in which $C_1 \neq C_2$ and $\nexists C$, s.t. $C_1 <: C' <: C_2$ and $C' \neq C_1, C' \neq C_2$. Consider the two possibilities for C_1 :

- case-1: class C_1 is defined in client code.
In this case, the declaration of C_1 should be $\text{CL} = \text{class } C_1 \text{ extends } C_2 \{ \dots \}$. According to the rule E – DECLARATION, we have $\Pi(\text{CL}) = \text{class } \Pi(C_1) \text{ extends } \Pi(C_2) \{ \dots \}$. Thus we have $\Pi(C_1) <: \Pi(C_2)$.
- case-2: class C_1 is defined in API.
In this case we have C_2 is also a API defined class according to the definition of API in FJ. According to the checking rule **ConstrCover**, there exists $C_1 \hookrightarrow D_1, C_2 \hookrightarrow D_2 \in \text{TypeMapping}(\Pi)$. By the checking rule **Subtyping** and the fact that $C_1 <: C_2$, we have $D_1 = \Pi(C_1) <: D_2 = \Pi(C_2)$.

With this case proved, for any $C_1 <: C_2$, it can be split into $C_1 <: C' <: \dots <: C_2$. Applying the proof on each step by induction, the lemma is proved. \square

Lemma 3 (Variable Substitution). *Suppose that an FJ term t is well typed under context $\Gamma = \Gamma_1, \{ \bar{x} : \bar{C}_x \}$, i.e. $\Gamma \vdash_{\text{FJ}} t : C_t$, then after substituting terms \bar{t}_v for variables \bar{x} , with the property that $\Gamma_1 \vdash_{\text{FJ}} \bar{t}_v : \bar{C}_v$ and $\bar{C}_v <: \bar{C}_x$, t can be typed to C_t or a sub-class of C_t . Namely,*

$$\Gamma_1, \{ \bar{x} : \bar{C}_x \} \vdash_{\text{FJ}} t : C_t \implies \Gamma_1 \vdash_{\text{FJ}} [\bar{x} \rightarrow \bar{t}_v] t : C'_t, C'_t <: C_t$$

Proof. By induction on the derivation on an FJ term t , there are five cases to discuss:

- case-1 $t = x, \Gamma_{\text{FJ}} t : C_t, x : C_t$.

In this case, we substitute an FJ term t_u for x , where $\Gamma_1 \vdash_{\text{FJ}} t_u : C_u$ and $C_u <: C_t$

\square