

Modular Reasoning and a Definition of Supertype Abstraction

Gary T. Leavens and David A. Naumann University of Central Florida and Stevens Inst. of Technology Support from US National Science Foundation NII Shonan Workshop on OO Specification and Verification, September 22, 2015

jmlspecs.org

www.cs.ucf.edu/~leavens



- Modular reasoning for OO programs
 - Proving soundness and completeness
 - In general, without restriction to some particular proof system.



Supertype abstraction
T x;
// ...
{pre^T_m[x/self]} x.m(); {post^T_m[x/self]}

Formalize semantically:
 Independent of program logic

Contributions in [LN15]: Semantic treatments of:

- Refinement
- Modular correctness
- Supertype abstraction
- Behavioral Subtyping
- Necessity and sufficiency of behavioral subtyping
- Specification inheritance



Related Work: Liskov 1988 (p. 25)

"If for each object o_1 of type S there is an object o_2 of type T such that for all programs P defined in terms of T, the behavior of P is unchanged when o_1 is substituted for o_2 , then S is a subtype of T."

Problems:

- "Unchanged" behavior is too restrictive,
- What does substitution mean in OO programs?

Related Work: Liskov & Wing 1994 (fig. 4)

For S to be a behavioral subtype of T,

- Subtype's invariant must imply the supertype's:
 ∀ self: S . inv^S(self) ⇒ inv^T(self)
 - Subtype methods preserve the supertype method's behavior."
 - For each method m of type T and self:S $pre_{m}^{T}(self) \Rightarrow pre_{m}^{S}(self)$ $post_{m}^{S}(self) \Rightarrow post_{m}^{T}(self)$

Problems:

- No proofs of soundness
- Postcondition rule is too strong

Why Liskov and Wing's postcondition rule is too strong

```
class TrustingAnimal {
    public model int age;
    int ag; represents age := ag;
```

```
meth setAge(int a)
    requires 0 ≤ a ∧ a ≤ 150;
    ensures age = a;
    { ag := a; }
}
```

public class Animal extends TrustingAnimal {

meth setAge(int a); requires ($0 \le a \land a \le 150$) || a < 0; ensures (old($0 \le a \land a \le 150$) \Rightarrow age = a) && (old(a < 0) \Rightarrow age = old(age)); if ($0 \le a \land a \le 150$) then { ag := a }

An Idealized Java-like OO Language

- interfaces
- classes
- exceptions as objects
- type tests (is) and type casts
- expressions with effects

Omits:

- constructors
- super calls
 - concurrency

Language Semantics overview

- Denotational Semantics
- State transformers
 - Separate state spaces for initial and final states
 - Commands: final state variable exc
 - Expressions: final state variables exc and res
 - Only two kinds of outcomes: \perp or a state
- Two kinds of semantics:
 - Dynamic, models dynamic dispatch
 - Static, models supertype abstraction in reasoning

Language Grammar (Abstract Syntax)

$$T \qquad ::= K \mid I \mid \text{bool} \mid \text{int}$$

$$msig \qquad ::= m(\overline{x}:T): T$$

$$mdec \qquad ::= \text{ meth } msig \{ C \}$$

$$C \qquad ::= x := E \mid x.f := x$$

$$\mid \text{ var } x: T \text{ in } C$$

$$\mid C; C \mid \text{ if } x \text{ then } C \text{ else } C$$

$$\mid \text{ throw } x \mid \text{try } C \text{ catch}(x:T) \mid C$$

$$E \qquad ::= x \mid \text{null} \mid \text{true} \mid 0 \dots$$

$$\mid x.f \mid x = y$$

$$\mid x \text{ is } T \mid (T) \mid x \mid \text{new } K()$$

$$\mid x.m(\overline{x}) \mid \text{ let } x \text{ be } E \text{ in } E$$

 $K, L \in ClassName$ $I \in InterfaceName$ x, y, zfm Names of declared classes Names of declared interfaces Variable names (for parameters and locals) Field names Method names

Basic Domains

$o \in dom \ r \text{ means}$: $o \text{ is allocated and has type } r \ o$

Values: Val(int, r) = ℤ Val(bool, r) = {true, false} Val(K, r) = {null} ∪ {o | o ∈ dom r ∧ r o ≤ K} Val(I, r) = {null} ∪ {o | ∃K· K ≤ I ∧ o ∈ Val(K, r)}

Stores, Heaps, States, and State Transformers

 $s \in Store(\Gamma, r) \Leftrightarrow s \in ((x : dom \Gamma) \rightarrow Val(x, r))$ \land (self \in dom $\Gamma \Rightarrow$ s(self) = null) Obrecord(K, r) = Store(fields K, r) $h \in Heap(r) = (o : dom r) \rightarrow Obrecord(r o, r)$ $\sigma \in State(\Gamma) = (r : RefCtx) \times Heap(r) \times Store(\Gamma, r)$ $\varphi \in STrans(\Gamma, \Gamma') =$ $(\sigma: State(\Gamma)) \rightarrow \{\bot\} \cup \{\sigma' \mid \sigma' \in State(\Gamma')\}$ $\land extState(\sigma, \sigma') \land imuSelf(\sigma, \sigma')$ $extState((r, h, s), (r', h', s')) \Leftrightarrow r \subseteq r'$ $imuSelf((r, h, s), (r', h', s')) \Leftrightarrow$ $(self \in (dom \ s \cap dom \ s') \Rightarrow s(self) = s'(self)).$

Semantics of Expressions, Commands, and Methods $SemExpr(\Gamma, T) = STrans(\Gamma, [res : T, exc : Exc])$ $SemCommand(\Gamma, \Gamma') = STrans(\Gamma, [\Gamma, exc : Exc])$ $SemMeth(T,m) = STrans([self : T, z_1:U_1,...,z_n:U_n],$ [res: U, exc: Exc]where $mtype(T,m) = (z_1:U_1,...,z_n:U_n) \rightarrow U$

Method Environments

Normal method environments:

η ∈ MethEnv = (K : ClassName) × (m: Meths K) → SemMeth(K,m)

Extended method environment: $\eta \in XMethEnv = (T : RefType) \times (m: Meths T)$ $\rightarrow SemMeth(T,m)$

Example Semantics Clauses Common to Dynamic and Static $[[\Gamma \vdash \text{let } x \text{ be } E \text{ in } E1 : U]](\eta)(r, h, s)$ = lets $(r_0, h_0, s_0) = [[\Gamma \vdash E : T]](\eta)(r, h, s)$ in if $s_0 exc \neq null$ then $(r_0, h_0, [res: default U, exc: s_0 exc])$ else let $s_1 = [s, x : s_0 \text{ res}]$ in $[[\Gamma, x : T \vdash E1 : U]](\eta)(r_0, h_0, s_1)$

$$[[\Gamma \vdash x := E]](\eta)(r, h, s)$$

= lets $(r_1, h_1, s_1) = [[\Gamma \vdash E : T]](\eta)(r, h, s)$ in
if $s_1 \exp (r_1 + n_1) = [[\Gamma \vdash E : T]](\eta)(r, h, s)$ in
then $(r_1, h_1, [s_1 + s_1] + s_1) = [[\Gamma \vdash E : T]](\eta)(r, h, s)$ in
else $(r_1, h_1, [s_1 + s_1] + s_1) = [[\Gamma \vdash E : T]](\eta)(r, h, s)$ in
else $(r_1, h_1, [s_1 + s_1] + s_1) = [[\Gamma \vdash E : T]](\eta)(r, h, s)$ in
else $(r_1, h_1, [s_1 + s_1] + s_1) = [[\Gamma \vdash E : T]](\eta)(r, h, s)$ in
else $(r_1, h_1, [s_1 + s_1] + s_1) = [[\Gamma \vdash E : T]](\eta)(r, h, s)$ in

Dynamic and Static Semantics for method calls

$$\mathcal{D}[[\Gamma \vdash x.m(y_1,...,y_n): U]](\eta)(r, h, s)$$
= if $s x = null$ then $except(r, h, U, NullDeref)$
else let $K = r(s x)$ in let $z_1,...,z_n = formals(K,m)$ in
let $s_1 = [self : s x, z_1 : s y_1,...,z_n : s y_n]$ in
 $\eta(K,m)(r, h, s_1)$

$$\begin{split} & \mathcal{S}[[\Gamma \vdash x.m(y_1,...,y_n):U]](\eta')(r, h, s) \\ &= \text{if } s \ x = null \ \text{then } except(r, h, U, \text{NullDeref}) \\ &= \text{let } T = \Gamma \ x \ \text{in } \text{let } z_1,...,z_n = formals(T,m) \ \text{in} \\ &= \text{let } s_1 = [\text{self}: s \ x, \ z_1:s \ y_1,...,z_n:s \ y_n] \ \text{in} \\ &= \eta'(T,m)(r, h, s_1) \end{split}$$

Approximation Orderings

On State Transformers ϕ and ψ in STrans(Γ , Γ): define $\phi \preccurlyeq \psi$ if and only if for all σ in State(Γ), either $\phi \sigma = \psi \sigma$ or $\phi \sigma = \bot$.

On Method Environments η and η' : define $\eta \leq \eta'$ if and only if $\eta(K,m) \leq \eta'(K,m)$, for all K,m.

Dynamic semantics of class tables: $\mathcal{D}[[CT]]$ is lub of chains of method environments

Specification Semantics Basics

- Semantics, not syntax
- One-state predicate on Γ-states = SO(State(Γ))

General specifications of methods: def: A general specification of type $\Gamma \dashrightarrow \Gamma'$ is a triple (J, pre, post) consisting of: a nonempty set J and J-indexed families of predicates: $pre \in J \rightarrow \mathscr{O}(State(\Gamma))$ and $post \in J \rightarrow \mathscr{O}(State(\Gamma))$.

Relation to Two-State Specifications

Consider a method specification of the form: requires 0 ≤ age ∧ age < 150; ensures age = old(age+1);

Can encode this as the general specification of type [age: int] \dashrightarrow [exc: Exc] with index set [age: int]-States: (\wp (State([age:int]), pre_{σ}, post_{σ}) where pre_{σ} = { $\tau \mid \sigma = \tau \land \sigma = (r,h,s)$ $\land 0 \le s(age) \land s(age) < 150$ } and post_{σ} = { $\tau \mid \sigma = (r,h,s) \land \tau = (r',h',s')$ $\land s'(age) = s(age) + 1$ }

Satisfaction (total correctness) for General Specifications

def: $\phi \models (J, pre, post)$ if and only if for all $i \in J$, $\forall \sigma \cdot \sigma \in pre_i \Rightarrow \phi(\sigma) \in post_i$.

Correctness for Method Specifications



Intrinsic Refinement of General Specifications

Idea: Subtype's (stronger) specifications have implementations that can be used in place of those of supertype's (weaker) specifications.

Problem:

Subtype's specification knows that self has its subtype (or lower).

Thus type of **self** changes covariantly!

So types of the corresponding state transformers are not related by subtyping!

Dealing with type of self

Two flavors:

Exact: self has exactly the subtype
 Downward: self has the subtype or lower

Define:

selftype(r, h, s) = r(s(self)) $\sigma \in pre \mid T \Leftrightarrow selftype(\sigma) = T \land \sigma \in pre$ $\sigma \in pre \mid^*T \Leftrightarrow selftype(\sigma) \leq T \land \sigma \in pre$

Refinement (standard)

Let $spec_0 : \Gamma \dashrightarrow \Gamma'$ and $spec_1 : \Delta \dashrightarrow \Delta'$, where $\Delta \dashrightarrow \Delta' \leq \Gamma \dashrightarrow \Gamma'$ (i.e., $\Gamma \leq \Delta$ and $\Delta' \leq \Gamma'$). Then $spec_1$ refines $spec_0$, written $spec_1 \supseteq spec_0$, if and only if for all $\phi \in STrans(\Delta, \Delta')$, $\phi \models spec_1 \Rightarrow \phi \models spec_0$

Refinement at a Subtype

Let $spec_0 : \Gamma \dashrightarrow \Gamma'$ and $spec_2 : [\Delta | self : S] \dashrightarrow \Delta'$. where $\Delta \longrightarrow \Delta' \leq \Gamma \longrightarrow \Gamma'$ (i.e., $\Gamma \leq \Delta$ and $\Delta' \leq \Gamma'$) and $S \leq \Gamma$ self. Refinement at exact subtype S, spec₂ \exists Spec₀, is defined by $spec_2 \supseteq Spec_0 \Leftrightarrow spec_2 \supseteq spec_0 \downarrow S$. Refinement at a downward subtype S, spec₂ \exists^{*S} spec₀, is defined by $spec_2 \supseteq^{*S} spec_0 \Leftrightarrow spec_2 \supseteq spec_0 l^*S.$

Refinement at type S



Refinement at type S



Refinement at type S



Characterization of Refinement at a subtype

Suppose that (I, pre, post): $\Gamma \rightarrow \Gamma'$ and (J, pre', post'): $spec_2$: $[\Delta \mid self : S] \rightsquigarrow \Delta'$, where $S \leq \Gamma$ self and $\Delta \longrightarrow \Delta' \leq \Gamma \longrightarrow \Gamma'$. If (J, pre', post') is satisfiable, then the following are equivalent: (a) $(J, pre', post') \supseteq^{S} (I, pre, post)$ (b) $\forall i \in I, \sigma \in State(\Gamma) \cdot \sigma \in pre_i \downarrow S$ $\Rightarrow (\exists j \in J \cdot \sigma \in pre'_i)$ $\land (\forall \tau \in State(\Delta))$ $\cdot (\forall k \in J \cdot \sigma \in pre'_k \Rightarrow \tau \in post'_k)$ $\Rightarrow \tau \in post_i$).

Modular Correctness

Modular verifiers and proof systems:

- Focus on one method at a time
- Assume specification of all other methods

Domains for Modular Correctness

CT ∈ ClassTable = (K:ClassName) × (m:MethodName) → SemMeth(K,m)

ST ∈ SpecTable = (T:RefType) × (m:MethodName) → ([self:T, formals(T,m)] ~** [res: resType(T,m), exc:Exc])

Satisfaction for Spec Tables

An extended method environment η satisfies ST, written $\eta' \models ST$, if and only if for all ref types T and $m \in Meths T$, $\eta'(T,m) \models ST(T,m)$.

An normal method environment η satisfies ST, written $\eta \mid = ST$,

if and only if for all classes K and $m \in Meths K$, $\eta(K,m) \models ST(K,m)$.

Modular Correctness

For command $\Gamma \vdash C$ and Γ -specification spec, *C* modularly satisfies spec with respect to *ST*, written *ST*, $(\Gamma \vdash C) \models \mathcal{D}$ spec if and only if $\forall \eta \in MethEnv \cdot \eta \models ST \Rightarrow \mathcal{D}[[\Gamma \vdash C]](\eta) \models spec.$

For command $\Gamma \vdash C$ and Γ -specification spec, *C* modularly satisfies spec with respect to *ST* under static dispatch, written *ST*, $(\Gamma \vdash C) \models^{S}$ spec if and only if $\forall \eta \in XMethEnv \cdot$ $\eta \models ST \Rightarrow S[[\Gamma \vdash C]](\eta) \models spec.$

Supertype Abstraction (1)

A specification table ST allows supertype abstraction when $ST, (\Gamma \vdash C) \models Spec$ implies $ST, (\Gamma \vdash C) \models Spec$ and similarly for expressions.

However, we don't want to reason about all method environments as in the definitions of satisfaction!

Predicate Transformers to the Rescue

- A proof system would use axiomatic semantics
 Method m in type T would be dealt with as: assert pre^Tm; assume post^Tm;
 - which acts as a predicate transformer.
- Notation:
 - [spec]} is the predicate transformer for spec.
 - [ST] is the extended method environment composed of such transformers
 - = least refined environment that satisfies ST.
 - $S{[\Gamma \vdash C]}({[ST]})$ is the predicate transformer denoted by C in ${[ST]}$.

Modular Verification (2)

For command $\Gamma \vdash C$ and Γ -specification spec, C is modularly verified for spec with respect to ST, if and only if $\mathbb{S}\{[\Gamma \vdash C]\}(\{[ST]\}) \supseteq \{[spec]\}\}$.

Supertype Abstraction

Modular verification implies modular correctness when: $S{[\Gamma \vdash C]}{[ST]} \supseteq {[spec]}$ implies $ST, (\Gamma \vdash C) \models \mathcal{D}$ spec and similarly for expressions.



Main Results

The following are equivalent:

(a) *ST* has behavioral subtyping.

(b) Modular correctness under static dispatch implies modular correctness.

(c) Modular verification implies modular correctness.



Related Work

- Work with Naumann [LN06][LN15], basis for this talk. Proved exact conditions on behavioral subtyping for validity of supertype abstraction
 Liskov and Wing [LW94] "subtype requirement" like supertype abstraction. Abstraction functions implicit in JML.
- Several program logics for Java, [Mül02] [Par05] [Pie06] [PHM99], use supertype abstraction.
- America [Ame87] [Ame91] first proved soundness with behavioral subtyping.

Conclusions

- Supertype abstraction defined semantically, based on modular reasoning.
- Supertype abstraction is valid if:
 - invariant methodology enforced, and
 - subtypes are behavioral subtypes.

Plus: a story about specification inheritance.

References

- [Ame87] Pierre America. Inheritance and subtyping in a parallel object-oriented language. In Jean Bezivin et al., editors, ECOOP '87, European Conference on Object-Oriented Programming, Paris, France, pages 234–242, New York, NY, June 1987. Springer-Verlag. Lecture Notes in Computer Science, volume 276.
- [Ame91] Pierre America. Designing an object-oriented programming language with behavioural subtyping. In J. W. de Bakker, W. P. de Roever, and G. Rozenberg, editors, Foundations of Object-Oriented Languages, REX School/Workshop, Noordwijkerhout, The Netherlands, May/June 1990, volume 489 of Lecture Notes in Computer Science, pages 60-90. Springer-Verlag, New York, NY, 1991.
- [BCC+05] Lilian Burdy, Yoonsik Cheon, David R. Cok, Michael D. Ernst, Joeseph R. Kiniry, Gary T. Leavens, K. Rustan M. Leino, and Erik Poll. An overview of JML tools and applications. International Journal on Software Tools for Technology Transfer, 7(3):212–232, June 2005.
- [DL96] Krishna Kishore Dhara and Gary T. Leavens. Forcing behavioral subtyping through specification inheritance. In Proceedings of the 18th International Conference on Software Engineering, Berlin, Germany, pages 258–267. IEEE Computer Society Press, March 1996. A corrected version is ISU CS TR #95-20c, rlhttp://tinyurl.com/s2krg.
- [FF01] Robert Bruce Findler and Matthias Felleisen. Contract soundness for object-oriented languages. In OOPSLA '01 Conference Proceedings, Object-Oriented Programming, Systems, Languages, and Applications, October 14-18, 2001, Tampa Bay, Florida, USA, pages 1–15, October 2001.
- [Hoa69] C. A. R. Hoare. An axiomatic basis for computer programming. Communications of the ACM, 12(10):576-580,583, October 1969.
- [Hoa72] C. A. R. Hoare. Proof of correctness of data representations. Acta Informatica, 1(4):271-281, 1972.
- [LD00] Gary T. Leavens and Krishna Kishore Dhara. Concepts of behavioral subtyping and a sketch of their extension to component-based systems. In Gary T. Leavens and Murali Sitaraman, editors, Foundations of Component-Based Systems, chapter 6, pages 113–135. Cambridge University Press, 2000.
- [Lei98] K. Rustan M. Leino. Data groups: Specifying the modification of extended state. In OOPSLA '98 Conference Proceedings, volume 33(10) of ACM SIGPLAN Notices, pages 144-153. ACM, October 1998.
- [LN06] Gary T. Leavens and David A. Naumann. Behavioral subtyping, specification inheritance, and modular reasoning. Technical Report 06-20b, Department of Computer Science, Iowa State University, Ames, Iowa, 50011, September 2006.
- [LN15] Gary T. Leavens and David A. Naumann. Behavioral Subtyping, Specification Inheritance, and Modular Reasoning. ACM TOPLAS, 37(4):13:1-13:88, Aug. 2015. http://doi.acm.org/10.1145/2766446.
- [LW94] Barbara H. Liskov and Jeannette M. Wing. A behavioral notion of subtyping. ACM Transactions on Programming Languages and Systems, 16(6):1811– 1841, November 1994.
- [Mey97] Bertrand Meyer. Object-oriented Software Construction. Prentice Hall, New York, NY, second edition, 1997.
- [MPHL06] Peter Müller, Arnd Poetzsch-Heffter, and Gary T. Leavens. Modular invariants for layered object structures. Science of Computer Programming, 62(3):253–286, October 2006.
- [Mül02] Peter Müller. Modular Specification and Verification of Object-Oriented Programs, volume 2262 of Lecture Notes in Computer Science. Springer-Verlag, 2002.
- [ParO5] Matthew J. Parkinson. Local reasoning for Java. Technical Report 654, University of Cambridge Computer Laboratory, November 2005. The author's Ph.D. dissertation.
- [PHM99] A. Poetzsch-Heffter and P. Müller. A programming logic for sequential Java. In S. D. Swierstra, editor, European Symposium on Programming (ESOP '99), volume 1576 of Lecture Notes in Computer Science, pages 162–176. Springer-Verlag, 1999.
- [Pie06] Cees Pierik. Validation Techniques for Object-Oriented Proof Outlines. PhD thesis, Universiteit Utrecht, 2006.
- [SBC92] Susan Stepney, Rosalind Barden, and David Cooper, editors. Object Orientation in Z. Workshops in Computing. Springer-Verlag, Cambridge CB2 1LQ, UK, 1992.
- [Wil92] Alan Wills. Specification in Fresco. In Stepney et al. [SBC92], chapter 11, pages 127-135.
- [Win83] Jeannette Marie Wing. A two-tiered approach to specifying programs. Technical Report TR-299, Massachusetts Institute of Technology, Laboratory for Computer Science, 1983.