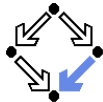


Reasoning about Programs as State Relations

Wolfgang Schreiner
Wolfgang.Schreiner@risc.uni-linz.ac.at

Research Institute for Symbolic Computation (RISC)
Johannes Kepler University, Linz, Austria
<http://www.risc.uni-linz.ac.at>

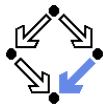




1. Core Idea

2. The Semantics of Programs

3. Outlook



Reasoning about Programs

How to describe the behavior of programs?

- Hoare calculus: triples of statements.

$$\{x = a\} x=x+1 \{x = a + 1\}$$

- Dynamic logic: formulas with modalities.

$$\forall a : x = a \Rightarrow [x=x+1] x = a + 1$$

- My approach: formulas with primed variables.

$$x=x+1: x' = x + 1$$

Core idea: translate programs to state relations described by formulas in classical predicate logic with classical rules of reasoning.



Example

- Program

```
x=x+1; if (x == 0) return 1; else y = x*x
```

- Formula

```
x' = ADD32(x,1) AND
```

```
IF x' = 0
```

```
THEN next.returns AND next.value = 1 AND y' = y
```

```
ELSE next.executes AND y' = MULT32(x,x)
```

Effect of command fully described in a classical logical framework.



Semantics of Commands

Take command C , context c , states s and s' .

- **Transition relation** $\llbracket _ \rrbracket^c$:

$$\llbracket C \rrbracket^c(s, s') \Leftrightarrow \dots$$

- $\llbracket C \rrbracket^c$ defines a relation on states.
- Which state transitions are possible by execution of C in c ?

- **Termination condition** $\llbracket _ \rrbracket_T^c$:

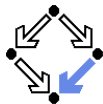
$$\llbracket C \rrbracket_T^c(s) \Leftrightarrow \dots$$

- $\llbracket C \rrbracket_T^c$ defines a condition on states.
- For which prestates must the execution of C yield a poststate?

- **Constraint:**

$$\forall s : \llbracket C \rrbracket_T^c(s) \Rightarrow \exists s' : \llbracket C \rrbracket^c(s, s').$$

The semantics of a command is described by a pair of a state relation and a state condition.



Semantics of Formulas

Take logical formula F , context c , environment e , states s and s'

- **Formula semantics** $\llbracket _ \rrbracket$:

$$\llbracket F \rrbracket^c(e)(s, s') \Leftrightarrow \dots$$

- $\llbracket F \rrbracket^c(e)$ defines a relation on states.

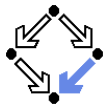
- **Formula semantics** $\llbracket _ \rrbracket$:

$$\llbracket F \rrbracket^c(e)(s) \Leftrightarrow \dots \llbracket F \rrbracket^c(e)(s, s)$$

- $\llbracket F \rrbracket^c(e)$ defines a condition on states.

We use a classical logical language for specifying state relations and state conditions.

Judgements



Various kinds of judgements that describe properties of commands.

$$C : F \Leftrightarrow \\ \forall c, s, s', e : \dots \Rightarrow \\ \llbracket C \rrbracket^c(s, s') \Rightarrow \llbracket F \rrbracket^c(e)(s, s')$$

$$C \downarrow F \Leftrightarrow \\ \forall c, s, e : \dots \Rightarrow \\ \llbracket F \rrbracket^c(s) \Rightarrow \llbracket C \rrbracket_T^c(e)(s)$$

...

We have a calculus for deriving only true judgements.



Specifications of Commands

requires F_C ensures $F_R \{C\}$

- Verification of partial correctness:

1. Derivation of transition relation F of C

$C : F$ (bottom-up application of calculus)

2. Proof that transition relation meets specification:

$$\forall c, es, s' : \dots \llbracket F \rrbracket^c(e)(s, s') \Rightarrow \llbracket F_C \Rightarrow F_R \rrbracket^c(e)(s, s')$$

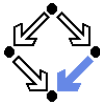
- Verification of total correctness:

- Derivation of termination condition F_C for C .

$C \downarrow F_C$ (top-down application of calculus)

- ...

As usual, the details are a bit more complicated.



Pre/Postcondition Reasoning

$$C : [F]_{I_1, \dots, I_n}$$

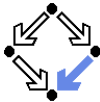
...

$$\text{PRE}(C, Q) =$$
$$\text{FORALL } \$J_1, \dots, \$J_n:$$
$$F[\$J_1/I_1', \dots, \$J_n/I_n'] \Rightarrow Q[\$J_1/I_1, \dots, \$J_n/I_n]$$
$$C : [F]_{I_1, \dots, I_n}$$

...

$$\text{POST}(C, P) =$$
$$\text{EXISTS } \$J_1, \dots, \$J_n: P[\$J_1/I_1, \dots, \$J_n/I_n] \text{ AND}$$
$$F[\$J_1/I_1, \dots, \$J_n/I_n, I_1/I_1', \dots, I_n/I_n']$$

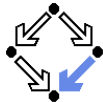
Pre/post-conditions can be computed from transition relations.



Advantages of Framework

- Translation of programs to classical logical formulas.
 - Well-understood semantics and rules of reasoning.
 - Only logic most people writing programs have ever learned.
 - May be investigated by programmer to yield insight.
 - Should be instructive (after appropriate simplification).
 - Application of general-purpose tools possible.
 - Simplification and verification.
- Relational semantical framework.
 - Simpler and more abstract semantics than functional framework.
 - No need for unique poststates, no need to be constructive.
 - Programs and specifications on same level.
 - Not functions versus relations.
 - Non-determinism covered.
 - Prepared for integration into concurrency framework.

Framework fully covers control flow interruptions.



1. Core Idea

2. The Semantics of Programs

3. Outlook



What are Programs?

$P \in$ Program, $Ms \in$ Methods, $M \in$ Method, $S \in$ Specification, $C \in$ Command, $E \in$ Expression, $I, J, K \in$ Identifier, $F \in$ Formula.

$P ::= Ms S \{C\}$.

$Ms ::= _ \mid Ms M$.

$M ::= \text{method } I_m(J_1, \dots, J_p) S \{C\}$.

$S ::= \text{writesonly } I_1, \dots, I_n \text{ throwonly } K_1, \dots, K_m$
 $\quad \text{requires } F_C \text{ ensures } F_R$

$C ::=$

$I = E \mid \text{var } I; C \mid \text{var } I = E; C \mid C_1; C_2$

$\text{if } (E) C \mid \text{if } (E) C_1 \text{ else } C_2$

$\text{while } (E) C$

$\text{continue} \mid \text{break} \mid \text{return } E \mid \text{throw } I E$

$\text{try } C_1 \text{ catch}(I_k I_v) C_2 \mid$

$I_r = I_m(E_1, \dots, E_p)$.



What is a State?

- Approximation 1: a state is a store.

$Store := Variable \rightarrow Value$

$State := Store$

- Sufficient for modeling basic programs.

- Approximation 2: a store also holds control data.

$Flag := \{E, C, B, R, T\}$

$Key := Value$

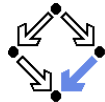
$Control := Flag \times Key \times Value$

$State := Store \times Control$

- Needed for modeling programs with control flow interruptions.
continue, break, return, throw.

First model can be in some sense “embedded” into second one.

What is the Meaning Program Identifiers?

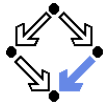


How are they related to variables in the store?

- Approximation 0: identifiers = variables.
 - Used in basic presentations of program semantics and reasoning.
- Approximation 1: identifiers denote variables.
 - Different identifiers denote different variables.
 - Not all variables need to be denoted by identifiers.
 $\llbracket I \rrbracket \in \text{Variable}, \llbracket I \rrbracket = \dots$
 $\text{DifferentVars} \Leftrightarrow \forall I, J \in \text{Identifier} : I \neq J \Rightarrow \llbracket I \rrbracket \neq \llbracket J \rrbracket$
- Approximation 2: given a context, identifiers denote variables.
 - Contexts may change.
 - In each context, different identifiers denote different variables.
 $\llbracket I \rrbracket : \text{Context} \rightarrow \text{Variable}, \llbracket I \rrbracket^c = \text{view}(c)(I)$
 $\text{DifferentVars}(c) \Leftrightarrow \forall I, J \in \text{Identifier} : I \neq J \Rightarrow \llbracket I \rrbracket^c \neq \llbracket J \rrbracket^c$

For the moment, we stick to a semantics without contexts.

The Algebra of States



$store : State \rightarrow Store, store(s, c) = s$

$read : State \times Identifier \rightarrow Value$

$read(s, l) = store(s)([l])$

$write : State \times Identifier \times Value \rightarrow State$

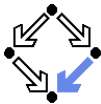
$write(s, l, v) = (store(s)([l] \mapsto v), control(s))$

$writes(s, l_1, v_1, \dots, l_n, v_n) \equiv$

$(store(s)([l_1] \mapsto v_1) \dots ([l_n] \mapsto v_n), control(s))$

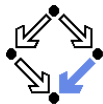
$s_0 = s_1 \text{ EXCEPT } l_1, \dots, l_n \equiv$

$\forall l \in Identifier : l \neq l_1 \wedge \dots \wedge l \neq l_n \Rightarrow read(s_0, l) = read(s_1, l)$



Semantics of Commands

$$\begin{aligned} \llbracket I = E \rrbracket(s, s') &\Leftrightarrow \\ & s' = \text{write}(s, I, \llbracket E \rrbracket(s)) \\ \llbracket \text{var } I; C \rrbracket(s, s') &\Leftrightarrow \\ & \exists s_0, s_1 \in \text{State} : \\ & \quad s_0 = s \text{ EXCEPT } I \wedge \llbracket C \rrbracket(s_0, s_1) \wedge \\ & \quad s' = \text{write}(s_1, I, \text{read}(s, I)) \\ \llbracket \text{var } I=E; C \rrbracket(s, s') &\Leftrightarrow \\ & \exists s_0, s_1 \in \text{State} : \\ & \quad s_0 = \text{write}(s, I, \llbracket E \rrbracket(s)) \wedge \llbracket C \rrbracket(s_0, s_1) \wedge \\ & \quad s' = \text{write}(s_1, I, \text{read}(s, I)) \\ \llbracket C_1; C_2 \rrbracket(s, s') &\Leftrightarrow \\ & \exists s_0 \in \text{State} : \llbracket C_1 \rrbracket(s, s_0) \wedge \llbracket C_2 \rrbracket(s_0, s') \\ \llbracket \text{if } (E) C \rrbracket(s, s') &\Leftrightarrow \\ & \text{IF } \llbracket E \rrbracket(s) = \text{TRUE} \text{ THEN } \llbracket C \rrbracket(s, s') \text{ ELSE } s' = s \\ \llbracket \text{if } (E) C_1 \text{ else } C_2 \rrbracket(s, s') &\Leftrightarrow \\ & \text{IF } \llbracket E \rrbracket(s) = \text{TRUE} \text{ THEN } \llbracket C_1 \rrbracket(s, s') \text{ ELSE } \llbracket C_2 \rrbracket(s, s') \end{aligned}$$



Semantics of Loops

For the moment, we stick to a semantics without interruptions.

$StateFunction := State \rightarrow Value$

$finiteExecution :$

$\mathbb{P}(\mathbb{N} \times State^\infty \times State \times StateFunction \times StateRelation)$

$finiteExecution(k, t, s, E, C) \Leftrightarrow$

$t(0) = s \wedge \forall i \in \mathbb{N}_k : E(t(i)) = \text{TRUE} \wedge C(t(i), t(i+1))$

$\llbracket \text{while } (E) C \rrbracket(s, s') \Leftrightarrow$

$\exists k \in \mathbb{N}, t \in State^\infty :$

$finiteExecution(k, t, s, \llbracket E \rrbracket, \llbracket C \rrbracket) \wedge$

$\llbracket E \rrbracket(t(k)) \neq \text{TRUE} \wedge t(k) = s'$

The semantics of a loop can be described by a finite sequence of states.



The Algebra of Control Data

$control : State \rightarrow Control, control(s, c) = c$

$flag : Control \rightarrow Flag, flag(f, k, v) = f$

$key : Control \rightarrow Key, key(f, k, v) = k$

$value : Control \rightarrow Value, value(f, k, v) = v$

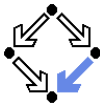
$executes : \mathbb{P}(Control), executes(c) \Leftrightarrow flag(c) = E$

$continues : \mathbb{P}(Control), continues(c) \Leftrightarrow flag(c) = C$

$breaks : \mathbb{P}(Control), breaks(c) \Leftrightarrow flag(c) = B$

$returns : \mathbb{P}(Control), returns(c) \Leftrightarrow flag(c) = R$

$throws : \mathbb{P}(Control), throws(c) \Leftrightarrow flag(c) = T$



The Algebra of Control Data

$execute : State \rightarrow State$

$execute(s) =$

LET $c = control(s)$ IN $(store(s), (E, key(c), value(c)))$

$continue : State \rightarrow State$

$continue(s) =$

LET $c = control(s)$ IN $(store(s), (C, key(c), value(c)))$

$break : State \rightarrow State$

$break(s) =$

LET $c = control(s)$ IN $(store(s), (B, key(c), value(c)))$

$return : State \times Value \rightarrow State$

$return(s, v) =$

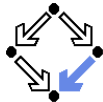
LET $c = control(s)$ IN $(store(s), (R, key(c), v))$

$throw : State \times Key \times Value \rightarrow State$

$throw(s, k, v) =$

LET $c = control(s)$ IN $(store(s), (B, k, v))$

Semantics of Commands with Interruptions

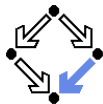


The semantics of some commands has to be redefined.

$$\begin{aligned} \llbracket \text{var } l; C \rrbracket(s, s') &\Leftrightarrow \\ &\exists s_0, s_1 \in \text{State} : \\ & s_0 = s \text{ EXCEPT } l \wedge \text{control}(s_0) = \text{control}(s) \wedge \\ & \llbracket C \rrbracket(s_0, s_1) \wedge s' = \text{write}(s_1, l, \text{read}(s, l)) \end{aligned}$$

$$\begin{aligned} \llbracket C_1; C_2 \rrbracket(s, s') &\Leftrightarrow \\ &\exists s_0 \in \text{State} : \\ & \llbracket C_1 \rrbracket(s, s_0) \wedge \\ & \text{IF } \text{executes}(\text{control}(s_0)) \text{ THEN } \llbracket C_2 \rrbracket(s_0, s') \text{ ELSE } s' = s_0 \end{aligned}$$

Semantics of Commands with Interruptions



$\llbracket \text{continue} \rrbracket (s, s') \Leftrightarrow s' = \text{continue}(s)$

$\llbracket \text{break} \rrbracket (s, s') \Leftrightarrow s' = \text{break}(s)$

$\llbracket \text{return } E \rrbracket (s, s') \Leftrightarrow s' = \text{return}(s, \llbracket E \rrbracket (s))$

$\llbracket \text{throw } l \ E \rrbracket (s, s') \Leftrightarrow s' = \text{throw}(s, l, \llbracket E \rrbracket (s))$

$\llbracket \text{try } C_1 \text{ catch}(l_k \ l_v) \ C_2 \rrbracket (s, s') \Leftrightarrow$

$\exists s_0, s_1, s_2 \in \text{State} :$

$\llbracket C_1 \rrbracket (s, s_0) \wedge$

IF $\text{throws}(\text{control}(s_0)) \wedge \text{key}(\text{control}(s_0)) = l_k$ THEN

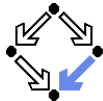
$s_1 = \text{write}(\text{execute}(s_0), l_v, \text{value}(\text{control}(s_0))) \wedge$

$\llbracket C_2 \rrbracket (s_1, s_2) \wedge$

$s' = \text{write}(s_2, l_v, \text{read}(s_0, l_v))$

ELSE $s' = s_0$

Semantics of Loops with Interruptions



finiteExecution :

$\mathbb{P}(\mathbb{N} \times \text{State}^\infty \times \text{State}^\infty \times \text{State} \times \text{StateFunction} \times \text{StateRelation})$

$\text{finiteExecution}(k, t, u, s, E, C) \Leftrightarrow$

$t(0) = s \wedge u(0) = s \wedge$

$\forall i \in \mathbb{N}_k :$

$\neg \text{breaks}(\text{control}(u(i))) \wedge \text{executes}(\text{control}(t(i))) \wedge$

$E(t(i)) = \text{TRUE} \wedge C(t(i), u(i+1)) \wedge$

$\text{IF } \text{continues}(\text{control}(u(i+1))) \vee \text{breaks}(\text{control}(u(i+1)))$

$\text{THEN } t(i+1) = \text{execute}(u(i+1))$

$\text{ELSE } t(i+1) = u(i+1)$

$\llbracket \text{while } (E) C \rrbracket (s, s') \Leftrightarrow$

$\exists k \in \mathbb{N}, t, u \in \text{State}^\infty :$

$\text{finiteExecution}(k, t, u, s, \llbracket E \rrbracket, \llbracket C \rrbracket) \wedge$

$(\llbracket E \rrbracket(t(k)) \neq \text{TRUE} \vee$

$\neg(\text{executes}(\text{control}(u(k))) \vee$

$\text{continues}(\text{control}(u(k)))) \wedge$

$t(k) = s'$

Expressions and Interruptions

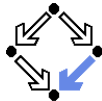


What happens on the evaluation of $1/0$?

- Mathematics: “undefined”
 - A value about which nothing is known.
- Programs: “abortion”.
 - Nowadays: “division by zero exception”.

For the moment, we stick to the mathematical view.

The Algebra of Contexts



$View = Identifier \rightarrow Variable$

$Space = \mathbb{P}(Variable)^\infty$

$Context = View \times Space$

$context : View \times Space \rightarrow Context, context(v, s) = \langle v, s \rangle$

$view : Context \rightarrow View, view(v, s) = v$

$space : Context \rightarrow Space, space(v, s) = s$

$range : Context \rightarrow Space \rightarrow \mathbb{P}(Variable)$

$range(c) = range(view(c)) \cup space(c)$

$take : Space \rightarrow (Variable \times Space)$

$take(s) = LET\ x = SUCH\ x : x \in s\ IN\ \langle x, s \setminus \{x\} \rangle$

$push : Context \times Identifier \rightarrow Context$

$push(c, l) = LET\ \langle v, s \rangle = c, \langle x, s' \rangle = take(s)\ IN\ \langle v[l \mapsto x], s' \rangle$

$push(c, l_1, \dots, l_n) \equiv push(\dots push(c, l_1) \dots, l_n)$

A context consists of a view (mapping of identifiers to variables) and a space (a set of unassigned variables).

Behavior of Contexts



$\forall c \in \text{Context}, I \in \text{Identifier} :$

$$\text{DifferentVariables}(c) \Rightarrow \text{DifferentVariables}(\text{push}(c, I))$$

$\forall c \in \text{Context}, I \in \text{Identifier} :$

$$\text{range}(\text{push}(c, I)) = \text{range}(c) \setminus \{\llbracket I \rrbracket^c\}$$

$\forall c \in \text{Context}, I, J \in \text{Identifier} :$

$$\text{DifferentVariables}(c) \Rightarrow \llbracket I \rrbracket^{\text{push}(c, I)} \neq \llbracket J \rrbracket^c$$

$\forall c \in \text{Context}, I, J \in \text{Identifier} :$

$$\text{DifferentVariables}(c) \wedge I \neq J \Rightarrow \llbracket J \rrbracket^{\text{push}(c, I)} = \llbracket J \rrbracket^c$$

Some useful properties for reasoning about contexts.

Contexts and States



$read : State \times Identifier \times Context \rightarrow Value$

$$read(s, l)^c = store(s)(\llbracket l \rrbracket^c)$$

$write : State \times Identifier \times Value \times Context \rightarrow State$

$$write(s, l, v)^c = \langle store(s)[\llbracket l \rrbracket^c \mapsto v], control(s) \rangle$$

$$writes(s, l_1, v_1, \dots, l_n, v_n)^c \equiv write(\dots write(s, l_1, v_1)^c \dots, l_n, v_n)^c$$

$$s_0 = s_1 \text{ EXCEPT}^c l_1, \dots, l_n \equiv$$

$$\forall l \in Identifier : l \neq l_1 \wedge \dots \wedge l \neq l_n \Rightarrow$$

$$read(s_0, l)^c = read(s_1, l)^c$$

Reading and writing stores now depends on the context.

Commands with Contexts



$$\begin{aligned} \llbracket l = E \rrbracket^c(s, s') &\Leftrightarrow \\ & s' = \text{write}(s, l, \llbracket E \rrbracket^c(s))^c \\ \llbracket \text{var } l; C \rrbracket^c(s, s') &\Leftrightarrow \\ & \text{LET } c' = \text{push}(c, l) \text{ IN } \llbracket C \rrbracket^{c'}(s, s') \\ \llbracket \text{var } l=E; C \rrbracket^c(s, s') &\Leftrightarrow \\ & \text{LET } c' = \text{push}(c, l), s_0 = \text{write}(s, l, \llbracket E \rrbracket^c(s))^c \text{ IN } \llbracket C \rrbracket^{c'}(s_0, s') \\ \llbracket \text{try } C_1 \text{ catch } (l_k \ l_v) \ C_2 \rrbracket^c(s, s') &\Leftrightarrow \\ & \exists s_0, s_1 \in \text{State} : \\ & \quad \llbracket C_1 \rrbracket^c(s, s_0) \wedge \\ & \quad \text{IF } \text{throws}(\text{control}(s_0)) \wedge \text{key}(\text{control}(s_0)) = l_k \text{ THEN} \\ & \quad \quad \text{LET } c' = \text{push}(c, l_v) \text{ IN} \\ & \quad \quad s_1 = \text{write}(\text{execute}(s_0), l_v, \text{value}(\text{control}(s_0)))^{c'} \wedge \\ & \quad \quad \llbracket C_2 \rrbracket^{c'}(s_1, s') \\ & \quad \text{ELSE } s' = s_0 \end{aligned}$$

Semantics of local declarations can be simplified.



Semantics of Method Calls

A command is executed within a *method environments*.

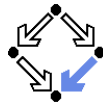
$StateCondition := \mathbb{P}(State)$

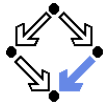
$Behavior := Value^* \rightarrow (StateRelation \times StateCondition)$

$MethodEnv := Identifier \rightarrow (View \times Behavior)$

$$\begin{aligned} \llbracket l_r = l_m(E_1, \dots, E_p) \rrbracket^{c, me}(s, s') &\Leftrightarrow \\ \text{LET } \langle v, b \rangle = me(l_m) \text{ IN} & \\ \text{LET } \langle r, t \rangle = b^{context(v, space(c))}(\llbracket E_1 \rrbracket^c(s), \dots, \llbracket E_p \rrbracket^c(s)) \text{ IN} & \\ \exists s_0 \in State : r(s, s_0) \wedge & \\ \text{IF } throws(control(s_0)) & \\ \text{THEN } s' = s_0 & \\ \text{ELSE } s' = write(s_0, l_r, value(control(s_0)))^c & \end{aligned}$$

Semantics of Method Declarations


$$\begin{aligned} & \llbracket \text{method } l_m(J_1, \dots, J_p) \ S \ \{C\} \rrbracket^V(me) = \\ & \quad \text{LET} \\ & \quad \quad b \in \text{Behavior} \\ & \quad \quad b^c(v_1, \dots, v_p) = \\ & \quad \quad \quad \text{LET} \\ & \quad \quad \quad \quad c' = \text{push}(c, J_1, \dots, J_p) \\ & \quad \quad \quad \quad r \in \text{StateRelation} \\ & \quad \quad \quad \quad r(s, s') \Leftrightarrow \\ & \quad \quad \quad \quad \quad \exists s_0 : \text{State} : \\ & \quad \quad \quad \quad \quad \quad \llbracket C \rrbracket^{c', me}(\text{writes}(s, J_1, v_1, \dots, J_p, v_p)^{c'}, s_0) \wedge \\ & \quad \quad \quad \quad \quad \quad s' = \text{IF } \text{throws}(\text{control}(s_0)) \\ & \quad \quad \quad \quad \quad \quad \text{THEN } s_0 \text{ ELSE } \text{executes}(s_0) \\ & \quad \quad \quad \quad t \in \text{StateCondition} \\ & \quad \quad \quad \quad t(s) \Leftrightarrow \llbracket C \rrbracket_T^{c', me}(\text{writes}(s, J_1, v_1, \dots, J_p, v_p)^{c'}) \\ & \quad \quad \quad \text{IN } \langle r, t \rangle \\ & \quad \text{IN } me[l \mapsto \langle v, b \rangle] \end{aligned}$$



1. Core Idea

2. The Semantics of Programs

3. Outlook

Outlook



- A calculus for $C : F$.
- A calculus for $C \downarrow F$.
- Auxiliary calculi.
 - Preconditions, postconditions, assertions.
- Dealing with “undefined expressions”.
 - A calculus for $C \checkmark F$.
- The overall verification framework.
- Dealing with recursive methods.

Stay tuned for more to come in this seminar.