Semantics and Verification of Software

Summer Semester 2019

Lecture 2: Operational Semantics of WHILE I (Evaluation of Expressions)

Thomas Noll
Software Modeling and Verification Group
RWTH Aachen University

https://moves.rwth-aachen.de/teaching/ss-19/sv-sw/

Software Modeling and Verification Chair
RWTH AACHEN UNIVERSITY

Recap: Syntax of WHILE

Syntactic Categories

WHILE: simple imperative programming language without procedures or advanced data structures

Syntactic categories:

Category	Domain	Meta variable
Numbers	$\mathbb{Z} = \{0, 1, -1, \ldots\}$	z
Truth values	$\mathbb{B} = \{\text{true}, \text{false}\}$	t
Variables	$\mathit{Var} = \{\mathrm{x}, \mathrm{y}, \ldots\}$	x
Arithmetic expressions	*AExp* (next slide)	a
Boolean expressions	*BExp* (next slide)	b
Commands (statements)	*Cmd* (next slide)	c

Recap: Syntax of WHILE

Syntax of WHILE Programs

Definition (Syntax of WHILE)

The syntax of WHILE Programs is defined by the following context-free grammar:

$$a ::= z \mid x \mid a_1 + a_2 \mid a_1 - a_2 \mid a_1 * a_2 \in \mathit{AExp}$$
$$b ::= t \mid a_1 = a_2 \mid a_1 > a_2 \mid \neg b \mid b_1 \wedge b_2 \mid b_1 \vee b_2 \in \mathit{BExp}$$
$$c ::= \texttt{skip} \mid x := a \mid c_1 \,;\, c_2 \mid \texttt{if } b \texttt{ then } c_1 \texttt{ else } c_2 \texttt{ end} \mid \texttt{while } b \texttt{ do } c \texttt{ end} \in \mathit{Cmd}$$

Remarks: we assume that

- the syntax of numbers, truth values and variables is predefined (i.e., no “lexical analysis”)
- the syntactic interpretation of ambiguous constructs (expressions) is uniquely determined (by brackets or priorities)

Operational Semantics of WHILE

- Idea: define meaning of programs by specifying its behaviour being executed on an (abstract) machine
- Here: evaluation/execution relation for program fragments (expressions, statements)
- Approach based on Structural Operational Semantics (SOS)
  - G.D. Plotkin: *A structural approach to operational semantics*, DAIMI FN-19, Computer Science Department, Aarhus University, 1981
- Employs derivation rules of the form

$$\text{(Name)}\ \frac{\text{Premise(s)}}{\text{Conclusion}}\quad \text{[side conditions]}$$

  - meaning: if every premise [and all side conditions] are fulfilled, then the conclusion can be drawn
  - a rule with no premises is called an axiom
- Derivation rules can be composed to form derivation trees with axioms as leaves (formal definition later)

Evaluation of Arithmetic Expressions

Program States

- Meaning of expression = its value (in the usual sense)
- Depends on the values of the variables in the expression

Definition 2.1 (Program state)

A (program) state is an element of the set

$$\Sigma := \{\sigma \mid \sigma : \mathit{Var} \to \mathbb{Z}\},$$

called the state space.

Thus $\sigma(x)$ denotes the value of $x \in \mathit{Var}$ in state $\sigma \in \Sigma$.

Evaluation of Arithmetic Expressions I

Remember: $a ::= z \mid x \mid a_1 + a_2 \mid a_1 - a_2 \mid a_1 * a_2 \in AExp$

Definition 2.2 (Evaluation relation for arithmetic expressions)

If $a \in AExp$ and $\sigma \in \Sigma$, then $\langle a, \sigma \rangle$ is called a configuration.

Expression a evaluates to $z \in \mathbb{Z}$ in state σ (notation: $\langle a, \sigma \rangle \to z$) if this relationship is derivable by means of the following rules:

Axioms: $\dfrac{}{\langle z, \sigma \rangle \to z} \qquad \dfrac{}{\langle x, \sigma \rangle \to \sigma(x)}$

Rules:

$$\frac{\langle a_1, \sigma \rangle \to z_1 \quad \langle a_2, \sigma \rangle \to z_2}{\langle a_1 + a_2, \sigma \rangle \to z} \text{ where } z := z_1 + z_2$$

$$\frac{\langle a_1, \sigma \rangle \to z_1 \quad \langle a_2, \sigma \rangle \to z_2}{\langle a_1 - a_2, \sigma \rangle \to z} \text{ where } z := z_1 - z_2$$

$$\frac{\langle a_1, \sigma \rangle \to z_1 \quad \langle a_2, \sigma \rangle \to z_2}{\langle a_1 * a_2, \sigma \rangle \to z} \text{ where } z := z_1 \cdot z_2$$

Evaluation of Arithmetic Expressions II

Example 2.3

$a = (\mathrm{x}{+}3){*}(\mathrm{y}{-}2)$, $\sigma(\mathrm{x}) = 3$, $\sigma(\mathrm{y}) = 9$:

$$\dfrac{\dfrac{\dfrac{}{\langle \mathrm{x}, \sigma\rangle \to 3} \quad \dfrac{}{\langle 3, \sigma\rangle \to 3}}{\langle \mathrm{x}{+}3, \sigma\rangle \to 6} \quad \dfrac{\dfrac{}{\langle \mathrm{y}, \sigma\rangle \to 9} \quad \dfrac{}{\langle 2, \sigma\rangle \to 2}}{\langle \mathrm{y}{-}2, \sigma\rangle \to 7}}{\langle (\mathrm{x}{+}3){*}(\mathrm{y}{-}2), \sigma\rangle \to 42}$$

$$\dfrac{\langle a_1, \sigma\rangle \to z_1 \quad \langle a_2, \sigma\rangle \to z_2}{\langle a_1 {*} a_2, \sigma\rangle \to z} \text{ where } z := z_1 \cdot z_2 \qquad \dfrac{\langle a_1, \sigma\rangle \to z_1 \quad \langle a_2, \sigma\rangle \to z_2}{\langle a_1 {+} a_2, \sigma\rangle \to z} \text{ where } z$$

Here: structure of derivation tree = structure of program fragment
(not generally true)

Free Variables I

First formal result: value of an expression only depends on valuation of variables which occur (freely) in the expression

Definition 2.4 (Free variables)

The set of free variables of an expression is given by the function

$$FV : AExp \to 2^{Var}$$

where

$$\begin{aligned} FV(z) &:= \emptyset & FV(a_1 + a_2) &:= FV(a_1) \cup FV(a_2) \\ FV(x) &:= \{x\} & FV(a_1 - a_2) &:= FV(a_1) \cup FV(a_2) \\ & & FV(a_1 * a_2) &:= FV(a_1) \cup FV(a_2) \end{aligned}$$

Result will be shown by structural induction on the expression

Excursus: Proof by Structural Induction I

Proof principle

Given: an inductive set, i.e., a set S whose elements are either

- atomic or
- obtained from atomic elements by (finite) application of certain operations

To show: property $P(s)$ applies to every $s \in S$

Proof: we verify:

Induction base: $P(s)$ holds for every atomic element s

Induction hypothesis: assume that $P(s_1)$, $P(s_2)$ etc.

Induction step: then also $P(f(s_1, \ldots, s_n))$ holds for every operation f of arity n

Remark: structural induction is a special case of well-founded induction

Excursus: Proof by Structural Induction II

Application: natural numbers (“mathematical induction”)

Definition: $\mathbb{N}$ is the least set which

- contains 0 and
- contains $n+1$ whenever $n \in \mathbb{N}$

Induction base: $P(0)$ holds

Induction hypothesis: $P(n)$ holds

Induction step: $P(n+1)$ holds

Generalisation: complete (strong, course-of-values) induction

- induction step: $P(0), P(1), \ldots, P(n) \Rightarrow P(n+1)$
- corresponds to well-founded induction over natural numbers

Excursus: Proof by Structural Induction III

Example 2.5 (Mathematical induction)

We prove that $P(n) : \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ holds for every $n \in \mathbb{N}$.

$P(0)$ holds: $\sum_{i=1}^{0} i = 0 = \frac{0(0+1)}{2}$ ✓

Assume $P(n)$: $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$

Show $P(n+1)$:

$$\begin{aligned}\sum_{i=1}^{n+1} i &= \sum_{i=1}^{n} i + (n+1)\\ &= \frac{n(n+1)}{2} + (n+1)\\ &= \frac{n(n+1)}{2} + \frac{2(n+1)}{2}\\ &= \frac{(n+2)(n+1)}{2}\\ &= \frac{(n+1)((n+1)+1)}{2} \quad ✓\end{aligned}$$

Excursus: Proof by Structural Induction IV

Application: arithmetic expressions (Def. 1.2)

Definition: $AExp$ is the least set which

- contains all integers $z \in \mathbb{Z}$ and all variables $x \in Var$ and
- contains $a_1 + a_2$, $a_1 - a_2$ and $a_1 * a_2$ whenever $a_1, a_2 \in AExp$

Induction base: $P(z)$ and $P(x)$ holds (for every $z \in \mathbb{Z}$ and $x \in Var$)

Induction hypothesis: $P(a_1)$ and $P(a_2)$ holds

Induction step: $P(a_1 + a_2)$, $P(a_1 - a_2)$ and $P(a_1 * a_2)$ holds

Free Variables II

Lemma 2.6

Let $a \in AExp$ and $\sigma, \sigma' \in \Sigma$ such that $\sigma(x) = \sigma'(x)$ for every $x \in FV(a)$. Then, for every $z \in \mathbb{Z}$,

$$\langle a, \sigma \rangle \to z \iff \langle a, \sigma' \rangle \to z.$$

Proof.

by structural induction on a (on the board) □

Evaluation of Boolean Expressions I

Definition 2.7 ((Strict) evaluation relation for Boolean expressions)

For $b \in BExp$, $\sigma \in \Sigma$, and $t \in \mathbb{B}$, the evaluation relation $\langle b, \sigma \rangle \to t$ is defined by:

$$\frac{}{\langle t, \sigma \rangle \to t}$$

$$\frac{\langle a_1, \sigma \rangle \to z \quad \langle a_2, \sigma \rangle \to z}{\langle a_1{=}a_2, \sigma \rangle \to \text{true}} \qquad \frac{\langle a_1, \sigma \rangle \to z_1 \quad \langle a_2, \sigma \rangle \to z_2}{\langle a_1{=}a_2, \sigma \rangle \to \text{false}} \text{ if } z_1 \neq z_2$$

$$\frac{\langle a_1, \sigma \rangle \to z_1 \quad \langle a_2, \sigma \rangle \to z_2}{\langle a_1{>}a_2, \sigma \rangle \to \text{true}} \text{ if } z_1 > z_2 \qquad \frac{\langle a_1, \sigma \rangle \to z_1 \quad \langle a_2, \sigma \rangle \to z_2}{\langle a_1{>}a_2, \sigma \rangle \to \text{false}} \text{ if } z_1 \leq z_2$$

$$\frac{\langle b, \sigma \rangle \to \text{false}}{\langle \neg b, \sigma \rangle \to \text{true}} \qquad \frac{\langle b, \sigma \rangle \to \text{true}}{\langle \neg b, \sigma \rangle \to \text{false}}$$

$$\frac{\langle b_1, \sigma \rangle \to \text{true} \quad \langle b_2, \sigma \rangle \to \text{true}}{\langle b_1 \wedge b_2, \sigma \rangle \to \text{true}} \qquad \frac{\langle b_1, \sigma \rangle \to \text{true} \quad \langle b_2, \sigma \rangle \to \text{false}}{\langle b_1 \wedge b_2, \sigma \rangle \to \text{false}}$$

$$\frac{\langle b_1, \sigma \rangle \to \text{false} \quad \langle b_2, \sigma \rangle \to \text{true}}{\langle b_1 \wedge b_2, \sigma \rangle \to \text{false}} \qquad \frac{\langle b_1, \sigma \rangle \to \text{false} \quad \langle b_2, \sigma \rangle \to \text{false}}{\langle b_1 \wedge b_2, \sigma \rangle \to \text{false}}$$

($\vee$ analogously)

Evaluation of Boolean Expressions II

Remarks:

- Binary Boolean operators $\wedge$ and $\vee$ are interpreted as strict, i.e., always evaluate both arguments.

  Important in situations like

  ```
  while p <> nil and p^.key < val do ...!
  ```

  (see following slides for alternatives)
- $FV : BExp \to 2^{Var}$ can be defined in analogy to Def. 2.4.
- Lemma 2.6 holds analogously for Boolean expressions, i.e., the value of $b \in BExp$ does not depend on variables in $Var \setminus FV(b)$.

Evaluation of Boolean Expressions III

Definition 2.8 (Sequential evaluation of Boolean expressions)

For $b \in BExp$, $\sigma \in \Sigma$, and $t \in \mathbb{B}$, the sequential evaluation relation $\langle b, \sigma \rangle \to t$ is defined by the following rules (truth values/relational expressions/negation as before):

$$\vdots$$

$$\frac{\langle b_1, \sigma \rangle \to \text{false}}{\langle b_1 \wedge b_2, \sigma \rangle \to \text{false}} \qquad \frac{\langle b_1, \sigma \rangle \to \text{true} \quad \langle b_2, \sigma \rangle \to t}{\langle b_1 \wedge b_2, \sigma \rangle \to t}$$

$$\frac{\langle b_1, \sigma \rangle \to \text{true}}{\langle b_1 \vee b_2, \sigma \rangle \to \text{true}} \qquad \frac{\langle b_1, \sigma \rangle \to \text{false} \quad \langle b_2, \sigma \rangle \to t}{\langle b_1 \vee b_2, \sigma \rangle \to t}$$

Remarks: yields same result as strict evaluation for our simple language

- (Boolean) expressions have no side effects (assignments, exceptions, ...)
- evaluation always terminates

Evaluation of Boolean Expressions IV

Definition 2.9 (Parallel evaluation of Boolean expressions)

For $b \in BExp$, $\sigma \in \Sigma$, and $t \in \mathbb{B}$, the parallel evaluation relation $\langle b, \sigma \rangle \to t$ is defined by the following rules (truth values/relational expressions/negation as before):

$$\vdots$$

$$\frac{\langle b_1, \sigma \rangle \to \text{false}}{\langle b_1 \wedge b_2, \sigma \rangle \to \text{false}} \qquad \frac{\langle b_2, \sigma \rangle \to \text{false}}{\langle b_1 \wedge b_2, \sigma \rangle \to \text{false}}$$

$$\frac{\langle b_1, \sigma \rangle \to \text{true} \quad \langle b_2, \sigma \rangle \to \text{true}}{\langle b_1 \wedge b_2, \sigma \rangle \to \text{true}}$$

$$\frac{\langle b_1, \sigma \rangle \to \text{true}}{\langle b_1 \vee b_2, \sigma \rangle \to \text{true}} \qquad \frac{\langle b_2, \sigma \rangle \to \text{true}}{\langle b_1 \vee b_2, \sigma \rangle \to \text{true}}$$

$$\frac{\langle b_1, \sigma \rangle \to \text{false} \quad \langle b_2, \sigma \rangle \to \text{false}}{\langle b_1 \vee b_2, \sigma \rangle \to \text{false}}$$