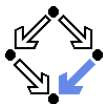


Programs with Undefined Expressions

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>





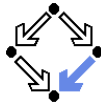
1. Programs with Undefined Expressions

2. Avoiding Undefined Expressions

3. Introducing Expression Checks

4. Well-definedness of Expressions

Undefined Expressions



$\perp := \text{SUCH } v : v \notin \text{Value}$

$\text{Value}_\perp := \text{Value} \cup \{\perp\}$

$\text{StateFunction}_\perp := \text{State} \rightarrow \text{Value}_\perp$

$\text{StatePredicate} := \mathbb{P}(\text{State})$

$\llbracket _ \rrbracket_D : \text{Expression} \rightarrow \text{StatePredicate}$

$\llbracket E \rrbracket_D(s) \Leftrightarrow \dots$

$\llbracket _ \rrbracket_\perp : \text{Expression} \rightarrow \text{StateFunction}_\perp$

$\llbracket E \rrbracket_\perp(s) = \text{IF } \llbracket E \rrbracket_D(s) \text{ THEN } \llbracket E \rrbracket(s) \text{ ELSE } \perp$

An expression may denote an “undefined” value.

Programs with Undefined Expressions



$EXP := \text{SUCH } k : k \in \text{Key}$

$\text{expthrow} : \text{State} \rightarrow \text{State}$

$\text{expthrow}(s) = \text{throw}(s, EXP, \text{read}(s, \text{VAL}))$

$\llbracket _ \rrbracket_{\perp} : \text{Command} \rightarrow \text{StateRelation}$

...

$\llbracket I=E \rrbracket_{\perp}(s, s') \Leftrightarrow$

LET $v = \llbracket E \rrbracket_{\perp}(s)$ IN

IF $v = \perp$

THEN $s' = \text{expthrow}(s)$

ELSE $s' = \text{write}(s, I, v)$

When encountering an undefined value, every command raises an “evaluation exception”.

Loops and Undefined Expressions



$finiteExecution_{\perp} :$

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

$finiteExecution_{\perp}(k, t, u, s, E, C) \Leftrightarrow$

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

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

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

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

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

THEN $t(i+1) = execute(u(i+1))$

ELSE $t(i+1) = u(i+1)$

Also an evaluation exception lets a loop terminate.

Loops and Undefined Expressions


$$\begin{aligned} \llbracket \text{while } (E) C \rrbracket_{\perp}(s, s') &\Leftrightarrow \\ \exists k \in \mathbb{N}, t, u \in \text{State}^{\infty} : & \\ \text{finiteExecution}_{\perp}(k, t, u, s, \llbracket E \rrbracket_{\perp}, \llbracket C \rrbracket_{\perp}) \wedge & \\ (\llbracket E \rrbracket_{\perp}(t(k)) = \perp \vee \llbracket E \rrbracket_{\perp}(t(k)) \neq \text{TRUE} \vee & \\ \neg(\text{executes}(\text{control}(u(k))) \vee \text{continues}(\text{control}(u(k)))))) \wedge & \\ \text{IF } (\text{executes}(\text{control}(u(k))) \vee \text{continues}(\text{control}(u(k)))) & \\ \wedge \llbracket E \rrbracket_{\perp}(t(k)) = \perp & \\ \text{THEN } \text{expthrow}(t(k)) = s' & \\ \text{ELSE } t(k) = s' & \end{aligned}$$

Evaluation exception lets loop terminate in a “throwing” state.

Relationship to Original Semantics



- If the expression value is not undefined, it is the same as in the original semantics:

$\forall E \in \text{Expression} :$

$\forall s \in \text{State} :$

$$\llbracket E \rrbracket_{\perp}(s) \neq \perp \Rightarrow \llbracket E \rrbracket_{\perp}(s) = \llbracket E \rrbracket(s)$$

- A command produces either an original poststate or an “exception throwing” state (provided that such exceptions are not caught).

$\forall C \in \text{Command} :$

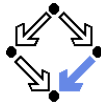
C has no subcommand $\text{try } C_1 \text{ catch}(EXP) C_2 \Rightarrow$

$$\forall s, s' \in \text{State} : \text{executes}(\text{control}(s)) \wedge \llbracket C \rrbracket_{\perp}(s, s') \Rightarrow \\ \llbracket C \rrbracket(s, s') \vee \text{expthrows}(\text{control}(s'))$$

$\text{expthrows} : \mathbb{P}(\text{Control})$

$\text{expthrows}(c) \Leftrightarrow \text{throws}(c) \wedge \text{key}(c) = EXP$

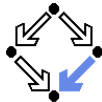
Reasoning about Undefined Expressions



How to reason about programs with undefined expressions?

1. Construct a new calculus for the new semantics.
 - Extra work; rules become considerably more complicated.
2. Rule out that program may encounter such expressions.
 - Typically only erroneous programs encounter such expressions.
 - Introduce calculus for verifying “well-definedness” of programs.
3. Transform programs by adding “checks” for undefined expressions.
 - Transformed programs in original semantics should have same behavior as original programs in new semantics.
 - By (automatically) inserting such checks, we can stick to the verification calculi of the original semantics.

We deal with the last two solutions.



1. Programs with Undefined Expressions

2. Avoiding Undefined Expressions

3. Introducing Expression Checks

4. Well-definedness of Expressions

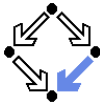


Avoiding Undefined Expressions

Introduce judgement for verifying the “well-definedness” of a program.

$$\begin{aligned} C \checkmark P &\Leftrightarrow \\ C \text{ has no subcommand } &\text{try } C_1 \text{ catch}(EXP I) C_2 \Rightarrow \\ \forall s, s' \in \text{State} : &\text{executes}(\text{control}(s)) \wedge \llbracket P \rrbracket(s) \Rightarrow \\ &(\llbracket C \rrbracket(s, s') \Leftrightarrow \llbracket C \rrbracket_{\perp}(s, s')) \end{aligned}$$

When executed in a state where property P holds, command C does not encounter an undefined value.



Avoiding Undefined Expressions

$$\vdash^D \vdash : \mathbb{P}(\text{Expression} \times \text{Formula})$$

$$E \simeq^D F_D \Leftrightarrow \forall s \in \text{State} : \llbracket E \rrbracket_D(s) \Leftrightarrow \llbracket F_D \rrbracket(s)$$

$$E \simeq^D F_D$$

$$\forall s \in \text{State} : \llbracket (\text{now.executes AND } P) \Rightarrow F_D \rrbracket(s)$$

$$I = E \checkmark P$$

Rules match those of "termination calculus" $C \downarrow F$.



Loops Avoiding Undefined Expressions

$$\begin{array}{l} E \stackrel{D}{\simeq} F_D \\ \forall s \in \text{State} : \llbracket (\text{now.executes AND } P) \Rightarrow F_D \rrbracket (s) \\ E \simeq H \\ C : [F]_{I_1, \dots, I_n}^{F_c, F_b, F_r, \{K_1, \dots, K_m\}} \\ \dots \\ R = H \text{ AND} \\ \quad \text{EXISTS } \$J_1, \dots, \$J_n : \text{EXSTATE } \#I_s : \\ \quad \quad P[\#I_s/\text{now}][\$J_1/I_1, \dots, \$J_n/I_n] \\ C \checkmark R \\ \text{POST}(C, R) = Q \\ \forall s \in \text{State} : \\ \quad \llbracket ((\text{now.executes OR now.continues}) \\ \quad \quad \text{AND } Q) \Rightarrow F_D \rrbracket (s) \\ \hline \text{while}(E) C \checkmark P \end{array}$$



Loops Avoiding Undefined Expressions

...

$Invariant(G, H, F)_{l_1, \dots, l_n}$

$R = H \text{ AND}$

EXISTS $\$J_1, \dots, \J_n : EXSTATE $\#l_s, \#l_t$:

$P[\#l_s/\text{now}][\$J_1/l_1, \dots, \$J_n/l_n]$ AND

$(G[\text{now}/\text{next}][l_1/l_1', \dots, l_n/l_n']$

$[\#l_s/\text{now}][\$J_1/l_1, \dots, \$J_n/l_n] \Rightarrow$

$G[\#l_s/\text{now}][\$J_1/l_1, \dots, \$J_n/l_n]$

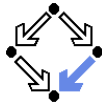
$[\#l_t/\text{next}][l_1/l_1', \dots, l_n/l_n']$ AND

$(!(\#l_t.\text{continues OR } \#l_t.\text{breaks}) \Rightarrow$

$\text{now} == \#l_t))$

...

$\text{while}(E) C \checkmark P$



-
1. Programs with Undefined Expressions
 2. Avoiding Undefined Expressions
 - 3. Introducing Expression Checks**
 4. Well-definedness of Expressions

Checked Programs



Transform program such that it “mimics” new semantics.

$$\begin{aligned} \text{CHECK}(C) = C' &\Leftrightarrow \\ C \text{ has no subcommand } \text{try } C_1 \text{ catch}(EXP I) C_2 &\Rightarrow \\ \forall s, s' \in \text{State} : \text{executes}(\text{control}(s)) &\Rightarrow \\ \llbracket C \rrbracket_{\perp}(s, s') \Leftrightarrow \llbracket C' \rrbracket(s, s') & \end{aligned}$$

Program C' behaves in the old semantics as program C does in the original semantics.



Checked Programs

$\sqsubseteq \stackrel{D}{\simeq} \sqsubseteq : \mathbb{P}(\text{Expression} \times \text{Expression})$

$E \stackrel{D}{\simeq} E_D : \Leftrightarrow \forall s \in \text{State} : \llbracket E \rrbracket_D(s) \Leftrightarrow \llbracket E_D \rrbracket(s) = \text{TRUE}$

$\text{CHECK } E_D \equiv \text{if } (!E_D) \text{ throw } EXP \text{ VAL}$

$$\frac{E \stackrel{D}{\simeq} E_D}{\text{CHECK}(I = E) = \text{CHECK } E_D; I = E}$$

Sound calculus for all commands (except loops).



Checked Loops (without continue)

$$\frac{E \stackrel{D}{\simeq} E_D \quad [C]_{I_1, \dots, I_n}^{\text{FALSE}, F_b, F_r, \{K_1, \dots, K_m\}} \quad \text{CHECK}(C) = C'}{\text{CHECK}(\text{while } (E) C) = \text{CHECK } E_D; \text{ while } (E) (C'; \text{ CHECK } E_D)}$$

Rule is sound, if loop does not contain continue.



Checked Loops (with continue)

Every execution of `continue` is immediately followed by the evaluation of the expression of the enclosing loop.

- **Idea:** precede every occurrence of `continue` by a check of the corresponding loop expression.

```
// transformation of continue
{check E_D; continue }
```

- Calculus needs to take into account enclosing loop expression.
- **Caveat:** `continue` may occur in a different variable context.

```
while (I > 0) {...{var I; ...continue; ...}...}
```

- Can be prevented by automatic renaming of local variables.

Adapt calculus to perform transformation correspondingly.

Checked Loops



$$E \stackrel{D}{\simeq} F_D$$

$$\forall s \in \text{State} : \llbracket (\text{now.executes AND } P) \Rightarrow F_D \rrbracket (s)$$

$$\text{CHECK}_{E'}(I=E) = P$$

$$\forall s_1, s_2 \in \text{State} : s_1 = s_2 \text{ EXCEPT } I \Rightarrow \llbracket E' \rrbracket_{\perp}(s_1) = \llbracket E' \rrbracket_{\perp}(s_2)$$

$$\text{CHECK}_{E'}(C) = C'$$

$$\text{CHECK}_{E'}(\text{var } I; C) = \text{var } I; C'$$

$$E' \stackrel{D}{\simeq} E_D$$

$$\text{CHECK}_{E'}(\text{continue}) = \text{check } E_D; \text{continue}$$

$$E \stackrel{D}{\simeq} E_D$$

$$\text{CHECK}_E(C) = C'$$

$$\text{CHECK}_{E'}(\text{while } (E) C) =$$

$$\text{CHECK } E_D; \text{while } (E) (C'; \text{CHECK } E_D)$$

What about the soundness of the transformation?



Soundness of Transformation

Transformed command is *not* completely equivalent to original one.

- Original command may (in new semantics) yield a “continuing” state with an undefined loop expression value.
 - Transformed command then generates (in original semantics) a “throwing” state with an “evaluation exception”.

Equivalence of programs is not completely preserved!

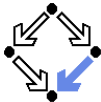


Soundness of Transformation

Assume C has no subcommand `try` C_1 `catch` $(EXP I) C_2$.

$$\begin{aligned} \text{CHECK}_{E'}(C) = C' &\Leftrightarrow \\ \forall s, s' \in \text{State} : \text{executes}(\text{control}(s)) &\Rightarrow \\ (\llbracket C \rrbracket_{\perp}(s, s') \wedge \neg \llbracket C' \rrbracket(s, s') &\Rightarrow \\ \text{continues}(\text{control}(s')) \wedge \llbracket E' \rrbracket_{\perp}(s') = \perp \wedge & \\ \llbracket C' \rrbracket(s, \text{expthrow}(s')) \wedge & \\ (\llbracket C' \rrbracket(s, s') \wedge \neg \llbracket C \rrbracket_{\perp}(s, s') &\Rightarrow \\ \exists s'' \in \text{State} : \llbracket C \rrbracket_{\perp}(s, s'') \wedge s' = \text{expthrow}(s'') \wedge & \\ \text{continues}(\text{control}(s'')) \wedge \llbracket E' \rrbracket_{\perp}(s'') = \perp) \wedge & \\ (\llbracket C' \rrbracket(s, s') \Rightarrow & \\ \neg(\text{continues}(\text{control}(s')) \wedge \llbracket E' \rrbracket_{\perp}(s') = \perp)) & \end{aligned}$$

There is a set of states on which the transformed command behaves different from the original one.



Relationship to Original Program

Assume $\text{CHECK}_{E'}(C) = C'$ can be derived.

$$\begin{aligned} \forall s, s' \in \text{State} : \text{executes}(\text{control}(s)) \wedge \llbracket C \rrbracket_{\perp}(s, s') \Rightarrow \\ \text{IF } \text{continues}(\text{control}(s')) \wedge \llbracket E' \rrbracket_{\perp}(s'') = \perp \\ \text{THEN } \llbracket C' \rrbracket(s, \text{exptrow}(s')) \\ \text{ELSE } \llbracket C' \rrbracket(s, s') \end{aligned}$$

(Only) if it can be ruled out that the execution of C yields a continuing state with undefined loop expression value, then a valid specification $C' : F$ (in the original semantics) also yields a valid specification $C : F$ (in the new semantics).

Consequences for Reasoning



$C = \text{while } (E) C_0$

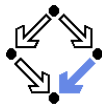
$\text{CHECK}_E(C_0) = C'_0$

$C'_0 : [F]_{I_1, \dots, I_n}^{F_c, F_b, F_r, \{K_1, \dots, K_m\}}$

- Case $F_c = \text{F}$
 - Execution of loop body C_0 does not result in a continuing state.
 - F is valid specification of C_0 in new semantics.
- Case $C \checkmark^L P$
 - Variant of “well-definedness” calculus to show that loop expressions are only evaluated in states where their values are defined.
 - F is valid specification of C_0 in new semantics.

(Only) in these cases, equivalence to new semantics holds.

Consequences for Reasoning


$$C = \text{while } (E) C_0$$
$$\text{CHECK}_E(C_0) = C'_0$$
$$C'_0 : [F]_{I_1, \dots, I_n}^{F_c, F_b, F_r, \{K_1, \dots, K_m\}}$$

■ General:

- None of above restrictions applies.
- Only $F' = (\text{NOT next.throws } EXP) \Rightarrow F$ is valid specification of C_0 in new semantics.

■ Consequences for invariant reasoning:

- Either use weaker specification F' of loop body C_0 .
- Or write invariant for C'_0 (in original semantics) rather than for C_0 (in new semantics).

To preserve full reasoning power, the transparency of the transformation has to be sacrificed.

Reasoning about Undefined Expressions



Develop verifier where user may choose between two modes.

- **“Well-Definedness”**: check for well-definedness of programs, report violations of well-definedness as errors.
 - Default mode (only mode in e.g. ESC/Java2).
 - Evaluation exceptions are unintended in most programs.
- **“Expression Checks”**: transform programs by introducing expression checks and verify transformed programs.
 - Undefined expression values give rise to evaluation exceptions.
 - Users may mostly think of programs as in the “new semantics”.
 - In loop invariants users must replace “continuing” states with undefined loop expressions by states with “evaluation exceptions”.
 - Transformation not completely transparent.

All in all, more complex than originally (naively) expected.



-
1. Programs with Undefined Expressions
 2. Avoiding Undefined Expressions
 3. Introducing Expression Checks
 - 4. Well-definedness of Expressions**

Expressions



For introducing expression checks, expression language must be rich enough to describe well-definedness of any expression.

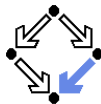
$E \in \text{Expression.}$

$E ::=$

$I \mid$
 $0 \mid 1 \mid -E \mid E_1 + E_2 \mid E_1 / E_2 \mid$
 $\text{true} \mid \text{false} \mid !E \mid E_1 \&\&E_2 \mid E_1 \mid \mid E_2 \mid$
 $E_1 == E_2 \mid E_1 < E_2 \mid E_1 <= E_2 \mid$
 $\text{null} \mid \text{new } E \mid E_1.\text{length} \mid E_1[E_2] \mid E_1[E_2] \rightarrow E_3]$

We omit semantics of the expressions.

Definedness Expressions



$\llbracket _ \rrbracket_{DE} : \text{Expression} \rightarrow \text{Expression}$

$\llbracket / \rrbracket_{DE} = \text{true}$

...

$\llbracket -E \rrbracket_{DE} = \llbracket E \rrbracket_{DE}$

$\llbracket E_1 + E_2 \rrbracket_{DE} = \llbracket E_1 \rrbracket_{DE} \ \&\& \ \llbracket E_2 \rrbracket_{DE}$

$\llbracket E_1 / E_2 \rrbracket_{DE} = \llbracket E_1 \rrbracket_{DE} \ \&\& \ \llbracket E_2 \rrbracket_{DE} \ \&\& \ !(E_2 == 0)$

...

$\llbracket \text{new } E \rrbracket_{DE} = \llbracket E \rrbracket_{DE} \ \&\& \ 0 \leq E$

$\llbracket E.\text{length} \rrbracket_{DE} = \llbracket E \rrbracket_{DE} \ \&\& \ !(E == \text{null})$

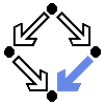
$\llbracket E_1[E_2] \rrbracket_{DE} = \llbracket E_1 \rrbracket_{DE} \ \&\& \ \llbracket E_2 \rrbracket_{DE} \ \&\&$

$!(E_1 == \text{null}) \ \&\& \ 0 \leq E_2 \ \&\& \ E_2 < E_1.\text{length}$

$\llbracket E_1[E_2|->E_3] \rrbracket_{DE} = \llbracket E_1 \rrbracket_{DE} \ \&\& \ \llbracket E_2 \rrbracket_{DE} \ \&\& \ \llbracket E_3 \rrbracket_{DE} \ \&\&$

$!(E_1 == \text{null}) \ \&\& \ 0 \leq E_2 \ \&\& \ E_2 < E_1.\text{length}$

Formulas



For reasoning about well-definedness, formula/term language must be rich enough to describe well-definedness of any expression.

$\llbracket _ \rrbracket : \mathbf{Predicate} \rightarrow \textit{Predicate}$

$$\llbracket \text{eq} \rrbracket(v_1, v_2) \Leftrightarrow v_1 = v_2$$

$$\llbracket \text{lt} \rrbracket(v_1, v_2) \Leftrightarrow v_1 < v_2$$

$$\llbracket \text{le} \rrbracket(v_1, v_2) \Leftrightarrow v_1 \leq v_2$$

$\llbracket _ \rrbracket : \mathbf{Function} \rightarrow \textit{Function}$

$$\llbracket \text{error} \rrbracket = 0$$

$$\llbracket \text{zero} \rrbracket = 0$$

$$\llbracket \text{one} \rrbracket = 1$$

$$\llbracket \text{neg} \rrbracket(v) = \ominus v$$

$$\llbracket \text{add} \rrbracket(v_1, v_2) = v_1 \oplus v_2$$

$$\llbracket \text{div} \rrbracket(v_1, v_2) = v_1 \oslash v_2$$

$$\llbracket \text{null} \rrbracket = \textit{null}$$

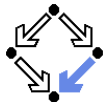
$$\llbracket \text{new} \rrbracket(v) = \textit{new}(v)$$

$$\llbracket \text{length} \rrbracket(v) = \textit{length}(v)$$

$$\llbracket \text{put} \rrbracket(v_1, v_2, v_3) = \textit{put}(v_1, v_2, v_3)$$

$$\llbracket \text{get} \rrbracket(v_1, v_2) = \textit{get}(v_1, v_2)$$

Definedness Formulas



$\llbracket _ \rrbracket_{DF} : \text{Expression} \rightarrow \text{Formula}$

$\llbracket / \rrbracket_{DF} = \text{TRUE}$

...

$\llbracket -E \rrbracket_{DF} = \llbracket E \rrbracket_{DF}$

$\llbracket E_1 + E_2 \rrbracket_{DF} = \llbracket E_1 \rrbracket_{DF} \text{ AND } \llbracket E_2 \rrbracket_{DF}$

$\llbracket E_1 / E_2 \rrbracket_{DF} = \llbracket E_1 \rrbracket_{DF} \text{ AND } \llbracket E_2 \rrbracket_{DF} \ \&\& \ !\text{eq}(\llbracket E_2 \rrbracket_{\text{TERM}}, \text{zero})$

...

$\llbracket \text{new } E \rrbracket_{DF} = \llbracket E \rrbracket_{DF} \text{ AND } \text{le}(\text{zero}, \llbracket E \rrbracket_{\text{TERM}})$

$\llbracket E.\text{length} \rrbracket_{DF} = \llbracket E \rrbracket_{DF} \text{ AND } \! \text{eq}(\llbracket E \rrbracket_{\text{TERM}}, \text{null})$

$\llbracket E_1[E_2] \rrbracket_{DF} = \llbracket E_1 \rrbracket_{DF} \text{ AND } \llbracket E_2 \rrbracket_{DF} \text{ AND}$

$\! \text{eq}(\llbracket E_1 \rrbracket_{\text{TERM}}, \text{null}) \text{ AND}$

$\text{le}(\text{zero}, \llbracket E_2 \rrbracket_{\text{TERM}}) \text{ AND}$

$\text{lt}(\llbracket E_2 \rrbracket_{\text{TERM}}, \text{length}(\llbracket E_1 \rrbracket_{\text{TERM}}))$

$\llbracket E_1[E_2 \rightarrow E_3] \rrbracket_{DF} = \llbracket E_1 \rrbracket_{DF} \text{ AND } \llbracket E_2 \rrbracket_{DF} \text{ AND } \llbracket E_3 \rrbracket_{DF} \text{ AND}$

$\! \text{eq}(\llbracket E_1 \rrbracket_{\text{TERM}}, \text{null}) \text{ AND}$

$\text{le}(\text{zero}, \llbracket E_2 \rrbracket_{\text{TERM}}) \text{ AND}$

$\text{lt}(\llbracket E_2 \rrbracket_{\text{TERM}}, \text{length}(\llbracket E_1 \rrbracket_{\text{TERM}}))$