

**ICTAC 2019**

October 30 – November 4, 2019

Hammamet, Tunisia

# Calculational Design of a Regular Model Checker by Abstract Interpretation

Patrick Cousot

New York University, Courant Institute of Mathematics, Computer Science

[pcousot@cs.nyu.edu](mailto:pcousot@cs.nyu.edu)

[cs.nyu.edu/~pcousot](http://cs.nyu.edu/~pcousot)

# Introduction

# How to design a static analyzer by abstract interpretation [P. Cousot and R. Cousot, 1977, 1979]

- Define the **syntax & semantics** of the language
- Define the **semantic properties** to be analyzed
- Define an **abstraction** of this semantic properties into an **abstract domain** (machine representable subset of the semantic properties)
- Design the static analyzer by **calculational design of the abstraction of the semantics**

# How to design a static analyzer by abstract interpretation [P. Cousot and R. Cousot, 1977, 1979]

- Define the **syntax & semantics** of the language
- Define the **semantic properties** to be analyzed
- Define an **abstraction** of this semantic properties into an **abstract domain** (machine representable subset of the semantic properties)
- Design the static analyzer by **calculational design of the abstraction of the semantics**
- This is illustrated by the design of a **regular model checker**  
("regular" means that the program behaviors are specified using regular expressions [Wolper, 1983])

# Syntax and trace semantics of programs

# Syntax

$x, y, \dots \in \mathcal{V}$   
 $A \in \mathcal{A} ::= 1 \mid x \mid A_1 - A_2$   
 $B \in \mathcal{B} ::= A_1 < A_2 \mid B_1 \text{ nand } B_2$   
 $S \in \mathcal{S} ::=$   
     $x = A ;$   
     $;$   
     $\text{if } (B) S \mid \text{if } (B) S \text{ else } S$   
     $\text{while } (B) S \mid \text{break ;}$   
     $\{ S \}$   
 $SL \in \mathcal{ST} ::= SL S \mid \epsilon$   
 $P \in \mathcal{P} ::= SL$   
 $S \in \mathcal{PC} \triangleq \mathcal{S} \cup \mathcal{ST} \cup \mathcal{P}$

variable ( $\mathcal{V}$  not empty)  
arithmetic expression  
boolean expression  
statement  
    assignment  
    skip  
    conditionals  
    iteration and break  
    compound statement  
statement list  
program  
program component

## Program labelling

Unique labelling to designate (sets of) program points  $\ell \in \mathbb{L}$ :

- $\text{at}[[S]]$  the program point at which execution of  $S$  starts;
- $\text{after}[[S]]$  the program exit point after  $S$ , at which execution of  $S$  is supposed to normally terminate, if ever;
- $\text{escape}[[S]]$  a boolean indicating whether or not the program component  $S$  contains a **break** ; statement escaping out of that component  $S$ ;
- $\text{break-to}[[S]]$  the program point at which execution of the program component  $S$  goes to when a **break** ; statement escapes out of that component  $S$ ;
- $\text{breaks-of}[[S]]$  the set of labels of all **break** ; statements that can escape out of  $S$

## Prefix traces

- Program label:  $\ell \in \mathcal{L}$  (locates next step to be executed in the program)
- Environment:  $\rho \in \mathbb{E}\mathcal{V} \triangleq \mathcal{V} \rightarrow \mathcal{V}$  assigns values  $\rho(x) \in \mathcal{V}$  to variables  $x \in \mathcal{V}$ .
- State:  $\langle \ell, \rho \rangle \in \mathcal{S} \triangleq (\mathcal{L} \times \mathbb{E}\mathcal{V})$
- Trace: finite or infinite sequence  $\pi \in \mathbb{T}^{+\infty}$  of states
- Example:  $\langle \ell_1, \{x \rightarrow 1\} \rangle \langle \ell_2, \{x \rightarrow 2\} \rangle \langle \ell_4, \{x \rightarrow 2\} \rangle$
- Trace concatenation:  $\frown$

$$\begin{array}{ll} \pi_1 \sigma_1 \frown \sigma_2 \pi_2 & \text{undefined if } \sigma_1 \neq \sigma_2 \\ \pi_1 \frown \sigma_2 \pi_2 \triangleq \pi_1 & \text{if } \pi_1 \in \mathbb{S}^\infty \text{ is infinite} \\ \pi_1 \sigma_1 \frown \sigma_1 \pi_2 \triangleq \pi_1 \sigma_1 \pi_2 & \text{if } \pi_1 \in \mathbb{T}^+ \text{ is finite} \end{array}$$

- In pattern matching, we sometimes need the empty trace  $\exists$ . For example if  $\sigma \pi \sigma' = \sigma$  then  $\pi = \exists$  and  $\sigma = \sigma'$ .



## Structural definitions

- Our definitions (semantics, modeled checking, etc) are **structural** *i.e.* by induction on the grammatical program structure

$$\left\{ \begin{array}{l} \widehat{\mathcal{D}}[s] X \triangleq \widehat{\mathcal{T}}[s] \left( \prod_{s' \triangleleft s} \widehat{\mathcal{D}}[s'] \right) X \\ s \in \mathcal{P}_C \end{array} \right.$$

- the transformer  $\widehat{\mathcal{T}}$  uses the results of the immediate components  $s' \triangleleft s$  and involves a fixpoint computation for iteration statements.

## Prefix trace semantics

- A **prefix trace** describes the beginning of a computation
- Evaluation of an arithmetic expression

$$\begin{aligned}\mathcal{A}[[1]]\rho &\triangleq 1 \\ \mathcal{A}[[x]]\rho &\triangleq \rho(x) \\ \mathcal{A}[[A_1 - A_2]]\rho &\triangleq \mathcal{A}[[A_1]]\rho - \mathcal{A}[[A_2]]\rho\end{aligned}\tag{2}$$

- Assignment  $S ::= \ell \ x = A \ ;$  (where  $\text{at}[[S]] = \ell$ )

$$\widehat{\mathcal{S}}_S^*[[S]] = \{\langle \ell, \rho \rangle \mid \rho \in \mathbb{E}\forall\} \cup \{\langle \ell, \rho \rangle \langle \text{after}[[S]], \rho[x \leftarrow \mathcal{A}[[A]]\rho] \rangle \mid \rho \in \mathbb{E}\forall\}\tag{1}$$

## Prefix trace semantics (cont'd)

- Break statement  $S ::= \ell \text{ break ;}$  (where  $\text{at}[[S]] = \ell$ )

$$\mathcal{S}^*[[S]] \triangleq \{\langle \ell, \rho \rangle \mid \rho \in \mathbb{E}\forall\} \cup \{\langle \ell, \rho \rangle \langle \text{break-to}[[S]], \rho \rangle \mid \rho \in \mathbb{E}\forall\} \quad (3)$$

## Prefix trace semantics (cont'd)

- Conditional statement  $S ::= \text{if } \ell \text{ (B) } S_t$  (where  $\text{at}[[S]] = \ell$ )

$$\begin{aligned} \widehat{\mathcal{S}}^*[[S]] \triangleq & \{ \langle \ell, \rho \rangle \mid \rho \in \mathbb{E}\nu \} \\ & \cup \{ \langle \ell, \rho \rangle \langle \text{after}[[S], \rho] \mid \mathcal{B}[[B]]\rho = \text{ff} \} \\ & \cup \{ \langle \ell, \rho \rangle \langle \text{at}[[S_t], \rho] \pi \mid \mathcal{B}[[B]]\rho = \text{tt} \wedge \langle \text{at}[[S_t], \rho] \pi \in \widehat{\mathcal{S}}^*[[S_t]] \} \end{aligned} \tag{4}$$

- If the conditional statement  $S$  is inside an iteration statement, and  $S_t$  has a break, the execution goes on at the  $\text{break-to}[[S]]$  after the iteration.

## Prefix trace semantics (cont'd)

- Statement list  $sl ::= sl' s$  (where  $at[[s]] = after[[sl']]$ )

$$\begin{aligned}\widehat{\mathcal{S}}^*[[sl]] &\triangleq \widehat{\mathcal{S}}^*[[sl']] \cup \widehat{\mathcal{S}}^*[[sl']] \frown \mathcal{S}^*[[s]] \\ \mathcal{S} \frown \mathcal{S}' &\triangleq \{\pi \frown \pi' \mid \pi \in \mathcal{S} \wedge \pi' \in \mathcal{S}' \wedge \pi \frown \pi' \text{ is well-defined}\}\end{aligned}\tag{5}$$

- $\pi' \in \widehat{\mathcal{S}}^*[[s]]$  starts  $at[[s]] = after[[sl']]$  so, by def.  $\frown$ , the trace  $\pi \in \widehat{\mathcal{S}}^*[[sl']]$  must terminate to be able to go on with  $s$ .

## Prefix trace semantics (cont'd)

- Empty statement list  $sl ::= \epsilon$  (where  $at[[sl]] \triangleq after[[sl]]$ )

$$\mathcal{S}^*[[sl]] \triangleq \{\langle at[[sl]], \rho \rangle \mid \rho \in \mathbb{E}\nu\}$$

## Prefix trace semantics (cont'd)

- Iteration statement  $S ::= \text{while } \ell \text{ (B) } S_b$  (where  $\text{at}[[S]] = \ell$ )

$$\widehat{\mathcal{F}}_S^*[[\text{while } \ell \text{ (B) } S_b]] = \text{lfp}^{\subseteq} \mathcal{F}^*[[\text{while } \ell \text{ (B) } S_b]] \quad (6)$$

$$\mathcal{F}_S^*[[\text{while } \ell \text{ (B) } S_b]] X \triangleq \{\langle \ell, \rho \rangle \mid \rho \in \mathbb{E}\nu\} \quad (a)$$

$$\cup \{\pi_2 \langle \ell', \rho \rangle \langle \text{after}[[S]], \rho \rangle \mid \pi_2 \langle \ell', \rho \rangle \in X \wedge \mathcal{B}[[B]] \rho = \text{ff} \wedge \ell' = \ell\} \quad (b)$$

$$\cup \{\pi_2 \langle \ell', \rho \rangle \langle \text{at}[[S_b]], \rho \rangle \cdot \pi_3 \mid \pi_2 \langle \ell', \rho \rangle \in X \wedge \mathcal{B}[[B]] \rho = \text{tt} \wedge \\ \langle \text{at}[[S_b]], \rho \rangle \cdot \pi_3 \in \widehat{\mathcal{F}}_S^*[[S_b]] \wedge \ell' = \ell\} \quad (c)$$

- (a) either the execution observation stop at  $[[\text{while } \ell \text{ (B) } S_b]] = \ell$ , or
  - (b) after a number of iterations, control is back to  $\ell$ , the test is false, and the loop is exited, or
  - (c) after a number of iterations, control is back to  $\ell$ , the test is true, and the loop body is executed
- (This includes the termination of the loop body after  $[[S_b]] = \text{at}[[\text{while } \ell \text{ (B) } S_b]] = \ell$ )

# Maximal trace semantics

- Maximal trace semantics

$$\mathcal{S}^+[[S]] \triangleq \{\pi \langle \ell, \rho \rangle \in \mathcal{S}^*[[S]] \mid (\ell = \text{after}[[S]]) \vee (\text{escape}[[S]] \wedge \ell = \text{break-to}[[S]])\}$$

$$\mathcal{S}^\infty[[S]] \triangleq \lim(\mathcal{S}^*[[S]])$$

- Limit

$$\lim \mathcal{T} \triangleq \{\langle \pi, \pi' \rangle \mid \pi' \in \mathbb{T}^\infty \wedge \forall n \in \mathbb{N} . \langle \pi, \pi'[0..n] \rangle \in \mathcal{T}\}.$$



# Specification of program semantics

## Regular specifications

- We specify execution traces using **regular expressions** where terminals/[meta]-characters are replaced by **local assertions**
- A local assertion  $L : B$  specifies that invariant  $B$  should be true whenever execution reaches a program label  $\ell \in L$  in the set  $L$ .
- $B$  depends on the initial value  $\underline{x}$  of the variables  $x$  and their current value  $x$  at  $\ell$
- Abbreviation:  $? : B \triangleq \mathbb{L} : B$  means that  $B$  holds at any program label  $\ell \in \mathbb{L}$

## Examples of regular specifications

- $(? : x \geq 0)^*$  states that the value of  $x$  is always positive or zero during program execution.
- $(? : x \geq \underline{x})^*$  states that the value of  $x$  is always greater than or equal to its initial value  $\underline{x}$  during execution.
- $(? : x \geq 0)^* \bullet \ell : x == 0 \bullet (? : x < 0)^*$  states that
  - the value of  $x$  should be positive or zero, and next
  - if program point  $\ell$  is ever reached then  $x$  should be 0, and next
  - if computations go on after program point  $\ell$  then  $x$  should be negative afterwards.
- In the literature: Fred Schneider's [security monitors](#) [Schneider, 2000] : monitor the actions of a program, checks the behavior of the program against a given safety specification (and initiate remedial actions)<sup>1,2</sup>

---

<sup>1</sup>use automata equivalent to regular expressions

<sup>2</sup>use actions instead of program labels

]

# Syntax of regular expressions

$L \in \wp(\mathbb{L})$	sets of program labels
$x, y, \dots \in \mathbb{V}$	program variables
$\underline{x}, \underline{y}, \dots \in \underline{\mathbb{V}}$	initial values of variables
$B \in \mathbb{B}$	boolean expressions such that $\text{vars}[B] \subseteq \mathbb{V} \cup \underline{\mathbb{V}}$
$R \in \mathbb{R}$	regular expressions (7)
$R ::= \varepsilon$	empty
$L : B$	invariant $B$ at $L$
$R_1 R_2$ (or $R_1 \bullet R_2$ )	concatenation
$R_1 \mid R_2$	alternative
$R_1^* \mid R_1^+$	zero/one or more occurrences of $R$
$(R_1)$	grouping

## Subsets of regular expressions

$\mathcal{R}_\epsilon$  empty regular expressions

$\mathcal{R}^+$  non-empty regular expressions (used for specifications since no execution is empty)

$\mathcal{R}^\dagger$  alternative  $|$ -free regular expressions

## Semantics of regular expressions

- The semantics  $\mathcal{S}^r[[R]]$  of a regular expression  $R$  is a relation between
  - an initial environment  $\underline{q}$  (holding the initial values of variables), and
  - the traces  $\pi$  from  $\underline{q}$  satisfying the regular specification  $R$
- Example:
  - $R \triangleq \ell : x = \underline{x} \bullet \ell' : x = \underline{x} + 1$
  - $\mathcal{S}^r[[R]] = \{ \langle \underline{q}, \langle \ell_1, \rho \rangle \langle \ell_2, \rho' \rangle \rangle \mid \rho(x) = \underline{q}(x) \wedge \rho'(x) = \underline{q}(x) + 1 \} []$
  - The program  $\ell_1 x = x + 1 ; \ell_2$  satisfies this specification
  - The program  $\ell_1 x = x + 1 ; \ell_2 x = x + 1 ; \ell_3$  also satisfies this specification (the execution can be longer than the specification)
  - The program  $\ell_1 y = 0 ; \ell_2$  does not satisfy this specification

## Semantics of regular expressions (Cont'd)

### Semantics of boolean expressions

$$\begin{aligned}\mathcal{A}[[1]]_{\underline{e}, \rho} &\triangleq 1 \\ \mathcal{A}[[x]]_{\underline{e}, \rho} &\triangleq \underline{e}(x) \\ \mathcal{A}[[x]]_{\underline{e}, \rho} &\triangleq \rho(x) \\ \mathcal{A}[[A_1 - A_2]]_{\underline{e}, \rho} &\triangleq \mathcal{A}[[A_1]]_{\underline{e}, \rho} - \mathcal{A}[[A_2]]_{\underline{e}, \rho} \\ \mathcal{B}[[A_1 < A_2]]_{\underline{e}, \rho} &\triangleq \mathcal{A}[[A_1]]_{\underline{e}, \rho} < \mathcal{A}[[A_2]]_{\underline{e}, \rho} \\ \mathcal{B}[[B_1 \text{ nand } B_2]]_{\underline{e}, \rho} &\triangleq \mathcal{B}[[B_1]]_{\underline{e}, \rho} \uparrow \mathcal{B}[[B_2]]_{\underline{e}, \rho}\end{aligned}\tag{8}$$

# Semantics of regular expressions (Cont'd)

## Semantics of regular expressions

$$\begin{aligned}\mathcal{S}^r[\varepsilon] &\triangleq \{\langle \underline{q}, \varepsilon \rangle \mid \underline{q} \in \mathbb{E}\mathbf{v}\} & \mathcal{S}^r[\mathbf{R}]^1 &\triangleq \mathcal{S}^r[\mathbf{R}] & (9) \\ \mathcal{S}^r[\mathbf{L} : \mathbf{B}] &\triangleq \{\langle \underline{q}, \langle \ell, \rho \rangle \rangle \mid \ell \in \mathbf{L} \wedge \mathcal{B}[\mathbf{B}]\underline{q}, \rho\} & \mathcal{S}^r[\mathbf{R}]^{n+1} &\triangleq \mathcal{S}^r[\mathbf{R}]^n \circ \mathcal{S}^r[\mathbf{R}] \\ \mathcal{S}^r[\mathbf{R}_1 \mathbf{R}_2] &\triangleq \mathcal{S}^r[\mathbf{R}_1] \circ \mathcal{S}^r[\mathbf{R}_2] & \mathcal{S}^r[\mathbf{R}^*] &\triangleq \bigcup_{n \in \mathbb{N}} \mathcal{S}^r[\mathbf{R}]^n \\ \mathcal{S} \circ \mathcal{S}' &\triangleq \{\langle \underline{q}, \pi \cdot \pi' \rangle \mid \langle \underline{q}, \pi \rangle \in \mathcal{S} \wedge \langle \underline{q}, \pi' \rangle \in \mathcal{S}'\} & \mathcal{S}^r[\mathbf{R}^+] &\triangleq \bigcup_{n \in \mathbb{N} \setminus \{0\}} \mathcal{S}^r[\mathbf{R}]^n \\ \mathcal{S}^r[\mathbf{R}_1 \mid \mathbf{R}_2] &\triangleq \mathcal{S}^r[\mathbf{R}_1] \cup \mathcal{S}^r[\mathbf{R}_2] & \mathcal{S}^r[(\mathbf{R})] &\triangleq \mathcal{S}^r[\mathbf{R}]\end{aligned}$$



# Semantic properties to be analyzed

## Semantics property

- The semantics of program  $P$  satisfies the specification  $R$  (for some initial environment  $\underline{q}$ )
- Traditionally denoted  $P, \underline{q} \models R$
- “satisfies” means the prefix trace semantics of  $P$  is included in that of the specification  $R$  (extended to be long enough)

### Definition 2 (Model checking)

$$P, \underline{q} \models R \triangleq (\{\underline{q}\} \times \widehat{\mathcal{S}}_s^*[P]) \subseteq \text{prefix}(\mathcal{S}^r[R \bullet (? : \text{tt})^*]) \quad \square$$

where

$$\text{prefix}(\Pi) \triangleq \{\langle \underline{q}, \pi \rangle \mid \pi \in \mathcal{S}^+ \wedge \exists \pi' \in \mathcal{S}^* . \langle \underline{q}, \pi \cdot \pi' \rangle \in \Pi\} \quad \text{prefix closure.}$$

the regular specification  $R$  specifies only a prefix of the traces of program  $P$

# Abstraction

# Model checking is an boolean abstraction of the program semantics

$$\alpha_{\underline{\rho}, R}(\Pi) \triangleq (\{\underline{\rho}\} \times \Pi) \subseteq \text{prefix}(\mathcal{S}^r[\mathbb{R} \bullet (? : \text{tt})^*])$$

$$P, \underline{\rho} \models R = \alpha_{\underline{\rho}, R}(\widehat{\mathcal{S}}_S^*[\mathbb{P}])$$

$$\langle \wp(S^+), \subseteq \rangle \xrightleftharpoons[\alpha_{\underline{\rho}, R}]{\gamma_{\underline{\rho}, R}} \langle \mathbb{B}, \Leftrightarrow \rangle$$

# A short digression on regular expressions

## Equivalence of regular expressions

- There are several ways of writing the same regular expression (e.g.  $a^+$  or  $a(a^*)$ )
- Notion of **equivalence**

$$R_1 \approx R_2 \triangleq (\mathcal{S}^r[[R_1]] = \mathcal{S}^r[[R_2]])$$

- Equivalent regular expressions have the same semantics

## Disjunctive normal form of regular expressions

- A regular expression is in **disjunctive normal form** if it is of the form  $(R_1 \mid \dots \mid R_n)$  for some  $n \geq 1$ , in which none of the  $R_i$ , for  $1 \leq i \leq n$ , contains an occurrence of  $\mid$ .
- Kleene's algorithm transforms any regular expression  $R$  into an equivalent disjunctive normal form  $\text{dnf}(R)$  (so  $\text{dnf}(R) \approx R$ )

$$\begin{array}{ll} \text{dnf}(\varepsilon) \triangleq \varepsilon & \text{dnf}(L : B) \triangleq L : B \\ \text{dnf}(R_1 \mid R_2) \triangleq \text{dnf}(R_1) \mid \text{dnf}(R_2) & \text{dnf}(R^+) \triangleq \text{dnf}(RR^*) \\ \text{dnf}(R^*) \triangleq \text{let } R^1 \mid \dots \mid R^n = \text{dnf}(R) \text{ in } ((R^1)^* \dots (R^n)^*)^* & \text{dnf}((R)) \triangleq (\text{dnf}(R)) \\ \text{dnf}(R_1 R_2) \triangleq \text{let } R_1^1 \mid \dots \mid R_1^{n_1} = \text{dnf}(R_1) \text{ and } R_2^1 \mid \dots \mid R_2^{n_2} = \text{dnf}(R_2) \text{ in } \bigvee_{i=1}^{n_1} \bigvee_{j=1}^{n_2} R_1^i R_2^j \end{array}$$

## [Brzozowski, 1964] derivative of regular expressions

- a string of the form  $a\sigma$  (starting with the symbol  $a$ ) matches an expression  $R$  iff the suffix  $\sigma$  matches the *derivative*  $D_a(R)$  (also denoted  $a^{-1}R$ )
- Given a non-empty and alternative-free regular expression  $R \in \mathbb{R}^+ \cap \mathbb{R}^\dagger$ , we define  $\text{fstnxt}(R) = \langle L : B, R' \rangle$  such that
  - $L : B$  recognizes the first state of sequences of states recognized by  $R$ ;
  - the derivative  $R'$  recognizes sequences of states after the first state of sequences of states recognized by  $R$ .

$$\text{fstnxt}(L : B) \triangleq \langle L : B, \varepsilon \rangle \tag{10}$$

$$\text{fstnxt}(R_1 R_2) \triangleq \text{fstnxt}(R_2) \quad \text{if } R_1 \in \mathbb{R}_\varepsilon$$

$$\text{fstnxt}(R_1 R_2) \triangleq \text{let } \langle R_1^f, R_1^n \rangle = \text{fstnxt}(R_1) \text{ in } \left( [ R_1^n \in \mathbb{R}_\varepsilon \text{ ? } \langle R_1^f, R_2 \rangle \text{ : } \langle R_1^f, R_1^n \bullet R_2 \rangle ] \right) \quad \text{if } R_1 \notin \mathbb{R}_\varepsilon$$

$$\text{fstnxt}(R^+) \triangleq \text{let } \langle R^f, R^n \rangle = \text{fstnxt}(R) \text{ in } \left( [ R^n \in \mathbb{R}_\varepsilon \text{ ? } \langle R^f, R^* \rangle \text{ : } \langle R^f, R^n \bullet R^* \rangle ] \right)$$

$$\text{fstnxt}((R)) \triangleq \text{fstnxt}(R)$$



# Calculational design of the abstract interpreter (I)

# Methodology

- Apply the abstraction function

$$\alpha_{\underline{Q},R}(\Pi) \triangleq (\{\underline{Q}\} \times \Pi) \subseteq \text{prefix}(\mathcal{S}^r[\mathbb{R} \bullet (? : \text{tt})^*])$$

to the semantics

$$\widehat{\mathcal{S}}_S^*[\mathbb{S}]$$

of program components  $S$

- by structural induction on the program components  $S$

# Methodology

- Problem:  $\alpha_{\underline{\rho}, R}(\widehat{\mathcal{F}}_S^*[[S]])$  is not structurally inductive on  $S$
- Counter-example:

$$\alpha_{\underline{\rho}, R}(\widehat{\mathcal{F}}_S^*[[S_1; S_2]]) = f_{\underline{\rho}, R}(\alpha_{\underline{\rho}, R_1}(\widehat{\mathcal{F}}_S^*[[S_1]]), \alpha_{\underline{\rho}, R_2}(\widehat{\mathcal{F}}_S^*[[S_2]]))$$

where  $R = R_1 R_2$ ,  $R_1$  specifies  $S_1$ , and  $R_2$  specifies  $S_2$

How do we get  $R_1$  and  $R_2$ ???

- Solution: use a more refined abstraction
  - Checking that  $S$  satisfies the beginning  $R_1$  of  $R$
  - Returns the remaining  $R_2$  of  $R$  at the end of  $S$

$$\begin{aligned} \alpha_{\underline{\rho}, R}(\widehat{\mathcal{F}}_S^*[[S_1; S_2]]) &= \text{let } \langle b_1, R_2 \rangle = \alpha_{\underline{\rho}, R}(\widehat{\mathcal{F}}_S^*[[S_1]]) \text{ in} \\ &\quad \text{let } \langle b_2, R_3 \rangle = \alpha_{\underline{\rho}, R_2}(\widehat{\mathcal{F}}_S^*[[S_2]]) \text{ in} \\ &\quad \langle b_1 \wedge b_2, R_3 \rangle \end{aligned}$$

# Structural regular model- checking abstraction

## Definition 2 of regular model checking

- We first consider the case of  $\perp$ -free regular expressions
- Trace model checking abstraction ( $\underline{\rho} \in \mathbb{E}\mathbb{V}$  is an initial environment and  $R \in \mathbb{R}^+ \cap \mathbb{R}^\dagger$  is a non-empty and  $\perp$ -free regular expression):

$$\mathcal{M}^t \langle \underline{\rho}, \varepsilon \rangle \pi \triangleq \langle \text{tt}, \varepsilon \rangle \quad (11)$$

$$\mathcal{M}^t \langle \underline{\rho}, R \rangle \exists \triangleq \langle \text{tt}, R \rangle$$

$$\mathcal{M}^t \langle \underline{\rho}, R \rangle \pi \triangleq \text{let } \langle \ell_1, \rho_1 \rangle \pi' = \pi \text{ and } \langle L : B, R' \rangle = \text{fstnxt}(R) \text{ in} \quad \pi \neq \exists$$

$$(\langle \underline{\rho}, \langle \ell_1, \rho_1 \rangle \rangle \in \mathcal{S}^r \llbracket L : B \rrbracket \text{ ? } \mathcal{M}^t \langle \underline{\rho}, R' \rangle \pi' \text{ : } \langle \text{ff}, R' \rangle )$$

## Example

- $\pi = \langle \ell_1, \rho_1 \rangle \pi'$  with  $\pi' = \langle \ell_2, \rho_2 \rangle \ni$  with  $\rho_2 = \rho_1[x \leftarrow \rho_1(x) + 1]$  is a trace of  $\widehat{\mathcal{S}}_{\mathcal{S}}^*[[\ell_1 \ x = x + 1 \ ; \ell_2]]$
- $R_1 = ? : x = \underline{x} \bullet ? : x = \underline{x} + 1 \bullet ? : x = \underline{x} + 3$
- $\text{fstnxt}(R_1) = \langle \mathcal{L} : x = \underline{x}, R_2 \rangle$  with  $R_2 = ? : x = \underline{x} + 1 \bullet ? : x = \underline{x} + 3$
- $\text{fstnxt}(R_2) = \langle \mathcal{L} : x = \underline{x} + 1, R_3 \rangle$  with  $R_3 = ? : x = \underline{x} + 3$
- $\mathcal{M}^t \langle \underline{\rho}, R_3 \rangle \ni = \langle \text{tt}, \varepsilon \rangle$
- $\langle \underline{\rho}, \langle \ell_2, \rho_2 \rangle \rangle \in \mathcal{S}^r[[\mathcal{L} : x = \underline{x} + 1]] = \rho_2(x) = \underline{\rho}(\underline{x}) + 1$
- $\mathcal{M}^t \langle \underline{\rho}, R_2 \rangle \pi' = ([ \langle \underline{\rho}, \langle \ell_2, \rho_2 \rangle \rangle \in \mathcal{S}^r[[\mathcal{L} : x = \underline{x} + 1]] \ ? \ \mathcal{M}^t \langle \underline{\rho}, R_3 \rangle \ni \circ \langle \text{ff}, R_3 \rangle ] = [ \rho_2(x) = \underline{\rho}(\underline{x}) + 1 \ ? \ \langle \text{tt}, \varepsilon \rangle \circ \langle \text{ff}, R_3 \rangle ])$
- $\langle \underline{\rho}, \langle \ell_1, \rho_1 \rangle \rangle \in \mathcal{S}^r[[\mathcal{L} : x = \underline{x}]] = \rho_1(x) = \underline{\rho}(\underline{x})$
- $\mathcal{M}^t \langle \underline{\rho}, R_1 \rangle \pi \hat{=} ([ \langle \underline{\rho}, \langle \ell_1, \rho_1 \rangle \rangle \in \mathcal{S}^r[[\mathcal{L} : x = \underline{x}]] \ ? \ \mathcal{M}^t \langle \underline{\rho}, R_2 \rangle \pi' \circ \langle \text{ff}, R_2 \rangle ] = [ \rho_1(x) = \underline{\rho}(\underline{x}) \ ? \ \mathcal{M}^t \langle \underline{\rho}, R_2 \rangle \pi' \circ \langle \text{ff}, R_2 \rangle ] = [ \rho_1(x) = \underline{\rho}(\underline{x}) \ ? \ [ \rho_2(x) = \underline{\rho}(\underline{x}) + 1 \ ? \ \langle \text{tt}, \varepsilon \rangle \circ \langle \text{ff}, R_3 \rangle ] \circ \langle \text{ff}, R_2 \rangle ] \leftarrow$  if ff we could also return the counter-example  $\pi$

## Definition 2 of regular model checking (Cont'd)

- Set of traces model checking abstraction (for an  $\perp$ -free regular expression  $R \in \mathcal{R}^+$ ):

$$\mathcal{M}^+(\underline{q}, R)\Pi \triangleq \{ \langle \pi, R' \rangle \mid \pi \in \Pi \wedge \langle \text{tt}, R' \rangle = \mathcal{M}^+(\underline{q}, R)\pi \} \quad (12)$$

This abstraction is a Galois connection

$$\langle \wp(\mathcal{S}^+), \subseteq \rangle \xrightleftharpoons[\mathcal{M}^+(\underline{q}, R)]{\gamma_{\mathcal{M}^+(\underline{q}, R)}} \langle \wp(\mathcal{S}^+ \times \mathcal{R}^+), \subseteq \rangle \quad \text{for } R \in \mathcal{R}^+ \text{ in (12)} \quad (16)$$

- Program component  $S \in \mathcal{PC}$  model checking (for an  $\perp$ -free regular expression  $R \in \mathcal{R}^+$ ):

$$\mathcal{M}^+[[S]](\underline{q}, R) \triangleq \mathcal{M}^+(\underline{q}, R)(\widehat{\mathcal{S}}_s^*[[S]]) \quad (13)$$

## Definition 2 of regular model checking (Cont'd)

- We now consider the general case by decomposition into  $\mid$ -free regular expressions
- **Set of traces model checking** (for regular expression  $R \in \mathcal{R}$ ):

$$\mathcal{M} \langle \underline{q}, R \rangle \Pi \triangleq \text{let } (R_1 \mid \dots \mid R_n) = \text{dnf}(R) \text{ in} \quad (14)$$

$$\bigcup_{i=1}^n \{ \pi \mid \exists R' \in \mathcal{R} . \langle \pi, R' \rangle \in \mathcal{M}^+ \langle \underline{q}, R_i \rangle \Pi \}$$

This abstraction is a Galois connection

$$\langle \wp(S^+), \subseteq \rangle \xrightleftharpoons[\mathcal{M} \langle \underline{q}, R \rangle]{\gamma_{\mathcal{M} \langle \underline{q}, R \rangle}} \langle \wp(S^+), \subseteq \rangle \quad \text{for } R \in \mathcal{R} \text{ in (14)} \quad (17)$$

- **Model checking of a program component**  $S \in \mathcal{P}\mathcal{C}$  (for regular expression  $R \in \mathcal{R}$ ):

$$\mathcal{M} \llbracket S \rrbracket \langle \underline{q}, R \rangle \triangleq \mathcal{M} \langle \underline{q}, R \rangle (\widehat{\mathcal{S}}_S^* \llbracket S \rrbracket) \quad (15)$$



## Definition 2 of regular model checking (Cont'd)

- Back to boolean model-checking

$$\langle \wp(\mathbb{S}^+), \subseteq \rangle \xrightleftharpoons[\alpha_{\mathcal{M}(\underline{Q}, R)}]{\gamma_{\mathcal{M}(\underline{Q}, R)}} \langle \mathbb{B}, \Leftrightarrow \rangle \quad (18)$$

where  $\alpha_{\mathcal{M}(\underline{Q}, R)}(X) \triangleq (\{\underline{q}\} \times X) \subseteq \mathcal{M}(\underline{Q}, R)(X)$

**Theorem 4 (Model checking soundness ( $\Leftarrow$ ) and completeness ( $\Rightarrow$ ))**

$$P, \underline{q} \models R \Leftrightarrow \alpha_{\mathcal{M}(\underline{Q}, R)}(\widehat{\mathcal{S}}_S^*[P]) \quad \square$$

Note that we can prove soundness/completeness from the specification of the model-checking algorithm (still to be designed)

# Structural model checking

- We have solved the non-inductiveness problem!

$$\left[ \text{Lemma 5 } \mathcal{M}^t \langle \underline{q}, R \rangle (\pi \cdot \pi') = \langle \text{tt}, R' \rangle \Leftrightarrow (\exists R'' \in R . \mathcal{M}^t \langle \underline{q}, R \rangle (\pi) = \langle \text{tt}, R'' \rangle \wedge \mathcal{M}^t \langle \underline{q}, R'' \rangle (\pi') = \langle \text{tt}, R' \rangle). \quad \square \right.$$

- Structural model checking

$$\left\{ \begin{array}{l} \widehat{\mathcal{M}} [S] \langle \underline{q}, R \rangle \triangleq \widehat{\mathcal{F}} [S] \left( \prod_{S' \triangleleft S} \widehat{\mathcal{M}} [S'] \right) \langle \underline{q}, R \rangle \\ S \in Pc \end{array} \right.$$

The  $S' \triangleleft S$  are the immediate components of program component  $S$ . By calculus,

$$\left[ \text{Theorem 6 } \widehat{\mathcal{M}} [S] \langle \underline{q}, R \rangle = \mathcal{M} [S] \langle \underline{q}, R \rangle. \quad \square \right.$$

# Calculational design of the structural model-checking abstract interpreter (II)

## Computational design

$$\begin{aligned} & \mathcal{M}[\underline{S}] \langle \underline{q}, R \rangle \\ \triangleq & \mathcal{M} \langle \underline{q}, R \rangle (\widehat{\mathcal{S}}_S^*[\underline{S}]) && \{(15)\} \\ = & \mathcal{M} \langle \underline{q}, R \rangle (\widehat{\mathcal{T}}_S[\underline{S}] (\prod_{S' \triangleleft S} \widehat{\mathcal{S}}_S^*[S'])) \langle \underline{q}, R \rangle \\ & \{ \text{by structural definition } \widehat{\mathcal{S}}_S^*[S] = \widehat{\mathcal{T}}_S[S] (\prod_{S' \triangleleft S} \widehat{\mathcal{S}}_S^*[S']) \text{ of the stateful prefix} \\ & \text{trace semantics in Section 2} \} \\ = & \dots \{ \text{calculus to expand definitions, rewrite and simplify formulæ by algebraic laws} \} \\ = & \widehat{\mathcal{T}}_{\mathcal{M}}[S] (\prod_{S' \triangleleft S} \mathcal{M}[S']) \langle \underline{q}, R \rangle \\ & \{ \text{by calculational design to commute the model checking abstraction on the} \\ & \text{result to the model checking of the arguments of } \widehat{\mathcal{S}}_S^*[S] \} \\ = & \widehat{\mathcal{T}}_{\mathcal{M}}[S] (\prod_{S' \triangleleft S} \widehat{\mathcal{M}}[S']) \langle \underline{q}, R \rangle && \{ \text{ind. hyp.} \} \\ \triangleq & \widehat{\mathcal{M}}[S] \langle \underline{q}, R \rangle && \{ \text{by defining } \widehat{\mathcal{M}}[S] \triangleq \widehat{\mathcal{T}}_{\mathcal{M}}[S] (\prod_{S' \triangleleft S} \widehat{\mathcal{M}}[S']) \} \end{aligned}$$

## Computational design

For iteration statements,  $\widehat{\mathcal{T}}[s](\prod_{s' \triangleleft s} \widehat{\mathcal{S}}_s^*[s']) \langle \varrho, \mathbb{R} \rangle$  is a fixpoint, and this proof involves the fixpoint transfer theorem [P. Cousot and R. Cousot, 1979, Th. 7.1.0.4 (3)] based on the commutation of the concrete and abstract transformer with the abstraction.

**Theorem 7 (exact least fixpoint abstraction in a complete lattice)** Assume that  $\langle \mathcal{C}, \sqsubseteq, \perp, \top, \sqcup, \sqcap \rangle$  and  $\langle \mathcal{A}, \preceq, 0, 1, \vee, \wedge \rangle$  are complete lattices,  $f \in \mathcal{C} \xrightarrow{\alpha} \mathcal{C}$  is increasing,  $\langle \mathcal{C}, \sqsubseteq \rangle \xrightleftharpoons[\alpha]{\gamma} \langle \mathcal{A}, \preceq \rangle$ ,  $\bar{f} \in \mathcal{A} \xrightarrow{\gamma} \mathcal{A}$  is increasing, and  $\alpha \circ f = \bar{f} \circ \alpha$  (*commutation property*). Then  $\alpha(\text{lfp}^{\sqsubseteq} f) = \text{lfp}^{\preceq} \bar{f}$ .  $\square$

# Structural regular model checking of an empty specification $\varepsilon$

$$\mathcal{M}^+[\underline{s}]\langle \underline{q}, \varepsilon \rangle$$

$$\triangleq \mathcal{M}^+\langle \underline{q}, \varepsilon \rangle (\widehat{\mathcal{F}}_s^*[\underline{s}]) \quad \{(13)\}$$

$$\triangleq \{ \langle \pi, \varepsilon' \rangle \mid \pi \in \widehat{\mathcal{F}}_s^*[\underline{s}] \wedge \langle \text{tt}, \varepsilon' \rangle = \mathcal{M}^t \langle \underline{q}, \varepsilon \rangle \pi \} \quad \{(12)\}$$

$$\triangleq \{ \langle \pi, \varepsilon' \rangle \mid \pi \in \widehat{\mathcal{F}}_s^*[\underline{s}] \wedge \langle \text{tt}, \varepsilon' \rangle = \langle \text{tt}, \varepsilon \rangle \} \quad \{ \mathcal{M}^t \langle \underline{q}, \varepsilon \rangle \pi \triangleq \langle \text{tt}, \varepsilon \rangle \text{ by (11)} \}$$

$$= \{ \langle \pi, \varepsilon \rangle \mid \pi \in \widehat{\mathcal{F}}_s^*[\underline{s}] \} \quad \{\text{def. } =\}$$

$$\triangleq \widehat{\mathcal{M}}^+[\underline{s}]\langle \underline{q}, \varepsilon \rangle$$

## Definition 3 (Structural model checking)

- Model checking an empty temporal specification  $\varepsilon$ .

$$\widehat{\mathcal{M}}^+[\underline{s}]\langle \underline{q}, \varepsilon \rangle \triangleq \{ \langle \pi, \varepsilon \rangle \mid \pi \in \widehat{\mathcal{F}}_s^*[\underline{s}] \} \quad (20)$$

# Structural regular model checking of programs $P ::= \text{sl}$

$$\begin{aligned}
 & \mathcal{M} \llbracket P \rrbracket \langle \underline{q}, R \rangle \\
 \triangleq & \mathcal{M} \langle \underline{q}, R \rangle (\widehat{\mathcal{S}}_s^* \llbracket P \rrbracket) \quad \{ (15) \} \\
 \triangleq & \text{let } (R_1 \mid \dots \mid R_n) = \text{dnf}(R) \text{ in } \bigcup_{i=1}^n \{ \pi \mid \exists R' \in \mathcal{R} . \langle \pi, R' \rangle \in \mathcal{M}^+ \langle \underline{q}, R_i \rangle (\widehat{\mathcal{S}}_s^* \llbracket P \rrbracket) \} \quad \{ (14) \} \\
 = & \text{let } (R_1 \mid \dots \mid R_n) = \text{dnf}(R) \text{ in } \bigcup_{i=1}^n \{ \pi \mid \exists R' \in \mathcal{R} . \langle \pi, R' \rangle \in \mathcal{M}^+ \langle \underline{q}, R_i \rangle (\widehat{\mathcal{S}}_s^* \llbracket \text{sl} \rrbracket) \} \\
 & \quad \quad \quad \{ \widehat{\mathcal{S}}_s^* \llbracket P \rrbracket = \widehat{\mathcal{S}}_s^* \llbracket \text{sl} \rrbracket \} \\
 = & \text{let } (R_1 \mid \dots \mid R_n) = \text{dnf}(R) \text{ in } \bigcup_{i=1}^n \{ \pi \mid \exists R' \in \mathcal{R} . \langle \pi, R' \rangle \in \widehat{\mathcal{M}}^+ \langle \underline{q}, R_i \rangle (\widehat{\mathcal{S}}_s^* \llbracket \text{sl} \rrbracket) \} \\
 & \quad \quad \quad \{ \text{ind. hyp.} \} \\
 = & \text{let } (R_1 \mid \dots \mid R_n) = \text{dnf}(R) \text{ in } \bigcup_{i=1}^n \{ \pi \mid \exists R' \in \mathcal{R} . \langle \pi, R' \rangle \in \widehat{\mathcal{M}}^+ \llbracket \text{sl} \rrbracket \langle \underline{q}, R_i \rangle \} \quad \{ (13) \} \\
 \triangleq & \widehat{\mathcal{M}} \llbracket \text{sl} \rrbracket \langle \underline{q}, R \rangle \quad \{ (19) \}
 \end{aligned}$$

# Structural regular model checking of programs $P ::= S\downarrow$ (Cont'd)

## Definition 3 (Structural model checking, contn'd)

- Model checking a program  $P ::= S\downarrow$  for a temporal specification  $R \in \mathcal{R}$  with alternatives.

$$\widehat{\mathcal{M}} \llbracket P \rrbracket \langle \underline{q}, R \rangle \triangleq \text{let } (R_1 \mid \dots \mid R_n) = \text{dnf}(R) \text{ in} \quad (19)$$
$$\bigcup_{i=1}^n \{ \pi \mid \exists R' \in \mathcal{R} . \langle \pi, R' \rangle \in \widehat{\mathcal{M}}^+ \llbracket S\downarrow \rrbracket \langle \underline{q}, R_i \rangle \}$$



# Structural regular model checking of assignments $S ::= \ell \ x = A ;$

## Definition 3 (Structural model checking, contr'n'd)

- Model checking an assignment statement  $S ::= \ell \ x = A ;$

$$\widehat{\mathcal{M}}^+[[S]]\langle \underline{q}, R \rangle \triangleq \text{let } \langle L : B, R' \rangle = \text{fstnxt}(R) \text{ in} \quad (23)$$

$$\{ \langle \langle \text{at}[[S]], \rho \rangle, R' \rangle \mid \langle \underline{q}, \langle \text{at}[[S]], \rho \rangle \rangle \in \mathcal{S}^r[[L : B]] \} \quad (a)$$

$$\cup \{ \langle \langle \text{at}[[S]], \rho \rangle \langle \text{after}[[S]], \rho[x \leftarrow \mathcal{A}[[A]]\rho] \rangle, \varepsilon \rangle \mid R' \in \mathcal{R}_\varepsilon \wedge \quad (b)$$

$$\langle \underline{q}, \langle \text{at}[[S]], \rho \rangle \rangle \in \mathcal{S}^r[[L : B]] \}$$

$$\cup \{ \langle \langle \text{at}[[S]], \rho \rangle \langle \text{after}[[S]], \rho[x \leftarrow \mathcal{A}[[A]]\rho] \rangle, R'' \rangle \mid R' \notin \mathcal{R}_\varepsilon \wedge \quad (c)$$

$$\langle \underline{q}, \langle \text{at}[[S]], \rho \rangle \rangle \in \mathcal{S}^r[[L : B]] \wedge \langle L' : B', R'' \rangle = \text{fstnxt}(R') \wedge$$

$$\langle \underline{q}, \langle \text{after}[[S]], \rho[x \leftarrow \mathcal{A}[[A]]\rho] \rangle \rangle \in \mathcal{S}^r[[L' : B']] \}$$

# Structural regular model checking of assignments $S ::= \ell \ x = A ;$ (Cont'd)

$$\begin{aligned}
 & \mathcal{M}^+[[S]] \langle \underline{q}, R \rangle \\
 = & \{ \langle \pi, R' \rangle \mid \pi \in \widehat{\mathcal{S}}_s^*[[S]] \wedge \langle \text{tt}, R' \rangle = \mathcal{M}^t \langle \underline{q}, R \rangle \pi \} && \text{\textcircled{13} and \textcircled{12}} \\
 = & \{ \langle \pi, R' \rangle \mid \pi \in \{ \langle \ell, \rho \rangle \mid \rho \in \mathbb{E}v \} \cup \{ \langle \ell, \rho \rangle \langle \text{after}[[S]], \rho[x \leftarrow v] \rangle \mid \rho \in \mathbb{E}v \wedge v = \mathcal{A}[[A]]\rho \wedge \langle \text{tt}, R' \rangle = \mathcal{M}^t \langle \underline{q}, R \rangle \pi \} \} && \text{\textcircled{1}} \\
 = & \{ \langle \langle \ell, \rho \rangle, R' \rangle \mid \rho \in \mathbb{E}v \wedge \langle \text{tt}, R' \rangle = \mathcal{M}^t \langle \underline{q}, R \rangle \langle \ell, \rho \rangle \} \cup \\
 & \{ \langle \langle \ell, \rho \rangle \langle \text{after}[[S]], \rho[x \leftarrow v] \rangle, R' \rangle \mid \rho \in \mathbb{E}v \wedge v = \mathcal{A}[[A]]\rho \wedge \langle \text{tt}, R' \rangle = \mathcal{M}^t \langle \underline{q}, R \rangle \langle \ell, \rho \rangle \langle \text{after}[[S]], \rho[x \leftarrow v] \rangle \} && \text{\textcircled{def. } \cup \text{ and } \in} \\
 = & \{ \langle \langle \ell, \rho \rangle, R' \rangle \mid \langle \text{tt}, R' \rangle = \text{let } \langle L : B, R'' \rangle = \text{fstnxt}(R) \text{ in } ( \langle \underline{q}, \langle \ell, \rho \rangle \rangle \in \mathcal{S}^r[[L : B]] \text{ ? } \langle \text{tt}, R'' \rangle \text{ : } \langle \text{ff}, R' \rangle ) \} \cup \\
 & \{ \langle \langle \ell, \rho \rangle \langle \text{after}[[S]], \rho[x \leftarrow v] \rangle, R' \rangle \mid v = \mathcal{A}[[A]]\rho \wedge \langle \text{tt}, R' \rangle = \text{let } \langle L : B, R'' \rangle = \text{fstnxt}(R) \text{ in } ( \langle \underline{q}, \langle \ell, \rho \rangle \rangle \in \mathcal{S}^r[[L : B]] \text{ ? } \mathcal{M}^t \langle \underline{q}, R'' \rangle \langle \text{after}[[S]], \rho[x \leftarrow v] \rangle \text{ : } \langle \text{ff}, R'' \rangle ) \} && \text{\textcircled{11}} \\
 = & \dots
 \end{aligned}$$

$$\begin{aligned}
&= \{ \langle \langle \ell, \rho \rangle, R' \rangle \mid \langle L : B, R' \rangle = \text{fstnxt}(R) \wedge \langle \underline{\varrho}, \langle \ell, \rho \rangle \rangle \in \mathcal{S}^r \llbracket L : B \rrbracket \} \cup \\
&\quad \{ \langle \langle \ell, \rho \rangle \langle \text{after} \llbracket S \rrbracket, \rho[x \leftarrow v] \rangle, R' \rangle \mid v = \mathcal{A} \llbracket A \rrbracket \rho \wedge \exists R'' \in \mathcal{R} . \langle L : B, R'' \rangle = \text{fstnxt}(R) \wedge \langle \underline{\varrho}, \langle \ell, \rho \rangle \rangle \in \mathcal{S}^r \llbracket L : B \rrbracket \wedge (R'' \in \mathcal{R}_\varepsilon \text{ ? } \text{tt} : \mathcal{M}^t \langle \underline{\varrho}, R'' \rangle \langle \text{after} \llbracket S \rrbracket, \rho[x \leftarrow v] \rangle = \langle \text{tt}, R' \rangle) \} \\
&\hspace{20em} \{ \text{def. = and } \mathcal{M}^t \langle \underline{\varrho}, \varepsilon \rangle \pi \triangleq \langle \text{tt}, \varepsilon \rangle \text{ by (11)} \} \\
&= \{ \langle \langle \ell, \rho \rangle, R' \rangle \mid \langle L : B, R' \rangle = \text{fstnxt}(R) \wedge \langle \underline{\varrho}, \langle \ell, \rho \rangle \rangle \in \mathcal{S}^r \llbracket L : B \rrbracket \} \cup \\
&\quad \{ \langle \langle \ell, \rho \rangle \langle \text{after} \llbracket S \rrbracket, \rho[x \leftarrow v] \rangle, R' \rangle \mid v = \mathcal{A} \llbracket A \rrbracket \rho \wedge \exists R'' \in \mathcal{R} . \langle L : B, R'' \rangle = \text{fstnxt}(R) \wedge \langle \underline{\varrho}, \langle \ell, \rho \rangle \rangle \in \mathcal{S}^r \llbracket L : B \rrbracket \wedge (R'' \in \mathcal{R}_\varepsilon \text{ ? } \text{tt} : \text{let } \langle L' : B', R''' \rangle = \text{fstnxt}(R'') \text{ in } \langle \underline{\varrho}, \langle \text{after} \llbracket S \rrbracket, \rho[x \leftarrow v] \rangle \rangle \in \mathcal{S}^r \llbracket L' : B' \rrbracket) \} \\
&\hspace{20em} \{ (11) \} \\
&= \text{let } \langle L : B, R' \rangle = \text{fstnxt}(R) \text{ in} \\
&\quad \{ \langle \langle \ell, \rho \rangle, R' \rangle \mid \langle \underline{\varrho}, \langle \ell, \rho \rangle \rangle \in \mathcal{S}^r \llbracket L : B \rrbracket \} \\
&\quad \cup \{ \langle \langle \ell, \rho \rangle \langle \text{after} \llbracket S \rrbracket, \rho[x \leftarrow v] \rangle, \varepsilon \rangle \mid v = \mathcal{A} \llbracket A \rrbracket \rho \wedge \langle \underline{\varrho}, \langle \ell, \rho \rangle \rangle \in \mathcal{S}^r \llbracket L : B \rrbracket \wedge R' \in \mathcal{R}_\varepsilon \} \\
&\quad \cup \{ \langle \langle \ell, \rho \rangle \langle \text{after} \llbracket S \rrbracket, \rho[x \leftarrow v] \rangle, R'' \rangle \mid v = \mathcal{A} \llbracket A \rrbracket \rho \wedge \langle \underline{\varrho}, \langle \ell, \rho \rangle \rangle \in \mathcal{S}^r \llbracket L : B \rrbracket \wedge R' \notin \mathcal{R}_\varepsilon \wedge \text{let } \langle L' : B', R'' \rangle = \text{fstnxt}(R') \text{ in } \langle \underline{\varrho}, \langle \text{after} \llbracket S \rrbracket, \rho[x \leftarrow v] \rangle \rangle \in \mathcal{S}^r \llbracket L' : B' \rrbracket \} \\
&\hspace{20em} \{ \text{def. } \cup \} \\
&= \widehat{\mathcal{M}}^+ \llbracket S \rrbracket \langle \underline{\varrho}, R \rangle \hspace{10em} \{ (23) \} \quad \square
\end{aligned}$$

# Structural regular model checking of a statement list $sl ::= sl' s$

## Definition 3 (Structural model checking, contn'd)

- Model checking a **statement list**  $sl ::= sl' s$

$$\widehat{\mathcal{M}}^+[[sl]]\langle \underline{q}, R \rangle \triangleq \widehat{\mathcal{M}}^+[[sl']]\langle \underline{q}, R \rangle \cup \{ \langle \pi \cdot \langle \text{at}[[S]], \rho \rangle \cdot \pi', R'' \rangle \mid \langle \pi \cdot \langle \text{at}[[S]], \rho \rangle, R' \rangle \in \widehat{\mathcal{M}}^+[[sl']]\langle \underline{q}, R \rangle \wedge \langle \langle \text{at}[[S]], \rho \rangle \cdot \pi', R'' \rangle \in \widehat{\mathcal{M}}^+[[S]]\langle \underline{q}, R' \rangle \} \quad (21)$$

# Structural regular model checking of iterations $S ::= \text{while}^\ell (B) S_b$

## Definition 3 (Structural model checking, contr'n'd)

- Model checking an iteration statement  $S ::= \text{while}^\ell (B) S_b$

$$\widehat{\mathcal{M}}^+[[S]]\langle \underline{e}, R \rangle \triangleq \text{lfp}^\subseteq (\widehat{\mathcal{F}}^+[[S]]\langle \underline{e}, R \rangle) \quad (26)$$

$$\widehat{\mathcal{F}}^+[[S]]\langle \underline{e}, R \rangle X \triangleq \dots\dots\dots$$

# Scalability

# Convergence

- In practice, the set  $\mathcal{S}$  of states must be assumed to be **finite** (and very small) and encoded symbolically
- Regular expressions may be replaced by **finite automata**
- Nevertheless, model-checking in general, and regular model checking in particular, **does not scale**
- **Convergence acceleration** methods (widening, narrowing, and duals) must be used (trivial example: bounded model checking limits the length of traces to an arbitrary length  $n$ )

# Liveness



# Liveness

- If the set of states is **finite**, this is safety
- Otherwise, abstraction is needed, BUT liveness is not preserved by over-approximation and under-approximation is difficult in infinite systems
- In general liveness in the finite abstract homomorphic transition does NOT imply liveness in the infinite concrete transition system, and
- non-liveness in the infinite concrete transition system does NOT imply non-liveness in the finite abstract transition system
- Our solution: **variant functions**.

# Conclusion

# Conclusion

- We have shown that a model-checker is an abstract interpretation of a program semantics [P. Cousot and R. Cousot, 2000]
- So the model-checker can be formally constructed by calculational design
- This provides a machine checkable [Jourdan, Laporte, Blazy, Leroy, and Pichardie, 2015] formal proof of soundness (and completeness) of the model-checker
- Soundness does not seem to be a preoccupation of the model-checking community!
- A computation tool (better than  $\text{\LaTeX}$  editing, grep, and copy-paste) would be very helpful
- Pave the way for further non trivial abstractions (beyond the homomorphic abstractions)

# Bibliography

## References I

- Brzozowski, Janusz A. (1964). “Derivatives of Regular Expressions”. *J. ACM* 11.4, pp. 481–494 (32).
- Cousot, Patrick and Radhia Cousot (1977). “Abstract Interpretation: A Unified Lattice Model for Static Analysis of Programs by Construction or Approximation of Fixpoints”. In: *POPL. ACM*, pp. 238–252 (3, 4).
- (1979). “Systematic Design of Program Analysis Frameworks”. In: *POPL. ACM Press*, pp. 269–282 (3, 4, 45).
  - (2000). “Temporal Abstract Interpretation”. In: *POPL. ACM*, pp. 12–25 (59).
- Jourdan, Jacques-Henri, Vincent Laporte, Sandrine Blazy, Xavier Leroy, and David Pichardie (2015). “A Formally-Verified C Static Analyzer”. In: *POPL. ACM*, pp. 247–259 (59).
- Schneider, Fred B. (2000). “Enforceable security policies”. *ACM Trans. Inf. Syst. Secur.* 3.1, pp. 30–50 (19).

## References II

Wolper, Pierre (1983). “Temporal Logic Can Be More Expressive”. *Information and Control* 56.1/2, pp. 72–99 (3, 4).

The End, Thank you