

# Extending propositional separation logic for robustness properties

---

Alessio Mansutti

LSV, CNRS, ENS Paris-Saclay, Université Paris-Saclay, Cachan, France

FSTTCS - December 2018

# Separation logic and program verification

**Hoare calculus** is based on proof rules manipulating Hoare triples.

$$\{\varphi\} C \{\varphi'\}$$

where

- $C$  is a program
- $\varphi$  (precondition) and  $\varphi'$  (postcondition) are assertions in some logical language.

Any (memory) model that satisfies  $\varphi$  will satisfy  $\varphi'$  after being modified by  $C$ .

# Programming languages with pointers

The so-called **rule of constancy**

$$\frac{\{\varphi\} C \{\varphi'\}}{\{\varphi \wedge \psi\} C \{\varphi' \wedge \psi\}}$$

“ $C$  does not mess with  $\psi$ ”

is generally not valid: it is unsound if  $C$  manipulates pointers.

# Programming languages with pointers

The so-called **rule of constancy**

$$\frac{\{\varphi\} C \{\varphi'\}}{\{\varphi \wedge \psi\} C \{\varphi' \wedge \psi\}} \quad \text{“}C \text{ does not mess with } \psi\text{”}$$

is generally not valid: it is unsound if  $C$  manipulates pointers.

Example:

$$\frac{\{\exists u.[x] = u\} [x] \leftarrow 4 \{[x] = 4\}}{\{[y] = 3 \wedge \exists u.[x] = u\} [x] \leftarrow 4 \{[y] = 3 \wedge [x] = 4\}}$$

not true if  $x$  and  $y$  are in aliasing.

## Separation logic (Reynolds'02)

Separation logic add the notion of **separation** ( $*$ ) of a state, so that the **frame rule**

$$\frac{\{\varphi\} C \{\varphi'\} \quad \text{modv}(C) \cap \text{fv}(\psi) = \emptyset}{\{\varphi * \psi\} C \{\varphi' * \psi\}}$$

is valid.

Intuitively, separation means  $([x] = n * [y] = m) \implies x \neq y$

## Separation logic (Reynolds'02)

Separation logic add the notion of **separation** ( $*$ ) of a state, so that the **frame rule**

$$\frac{\{\varphi\} C \{\varphi'\} \quad \text{modv}(C) \cap \text{fv}(\psi) = \emptyset}{\{\varphi * \psi\} C \{\varphi' * \psi\}}$$

is valid.

Intuitively, separation means  $([x] = n * [y] = m) \implies x \neq y$

- **Automatic Verifiers:** Infer, SLAyer, Predator
- **Semi-automatic Verifiers:** Smallfoot, Verifast

Also, see “Why Separation Logic Works” (Pym et al. '18)

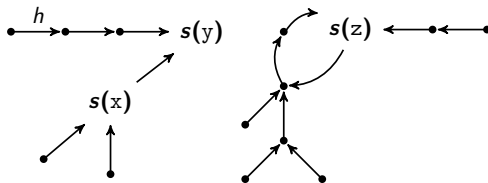
# Memory states

Separation Logic is interpreted over **memory states**  $(s, h)$  where:

■ **store**,  $s : \text{VAR} \rightarrow \text{LOC}$

■ **heap**,  $h : \text{LOC} \rightarrow_{\text{fin}} \text{LOC}$

where  $\text{VAR} = \{x, y, z, \dots\}$  set of (program) variables,  
 $\text{LOC}$  set of locations (typically  $\text{LOC} \cong \mathbb{N} \cong \text{VAR}$ ).



■ Disjoint heaps:  $\text{dom}(h_1) \cap \text{dom}(h_2) = \emptyset$

■ Sum of disjoint heaps  $(h_1 + h_2) = \text{sum of partial functions}$

# Propositional Separation Logic $SL(*, -*)$

$\varphi := \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \text{emp} \mid \mathbf{x} = \mathbf{y} \mid \mathbf{x} \hookrightarrow \mathbf{y} \mid \varphi_1 * \varphi_2 \mid \varphi_1 -* \varphi_2$

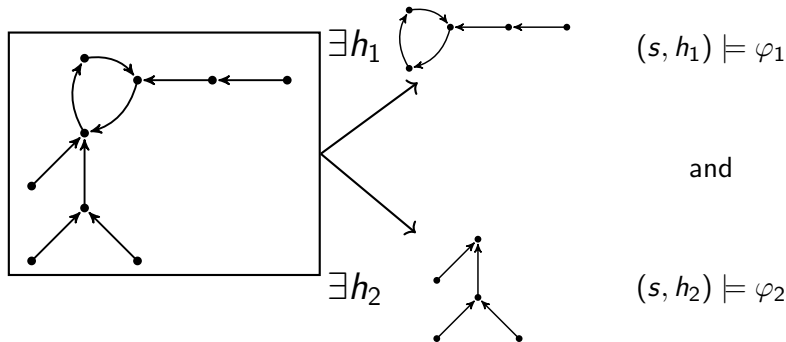
## Semantics

- standard for  $\wedge$  and  $\neg$ ;
- $(s, h) \models \text{emp} \iff \text{dom}(h) = \emptyset$
- $(s, h) \models \mathbf{x} = \mathbf{y} \iff s(\mathbf{x}) = s(\mathbf{y})$
- $(s, h) \models \mathbf{x} \hookrightarrow \mathbf{y} \iff h(s(\mathbf{x})) = s(\mathbf{y}),$  (previously  $[\mathbf{x}] = \mathbf{y}$ )



## Separating conjunction (\*)

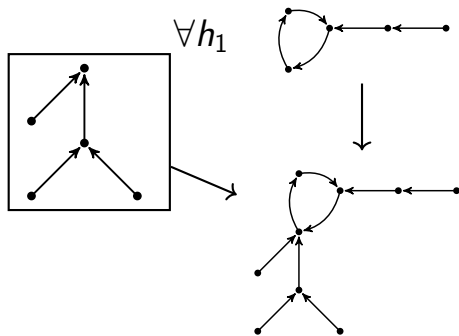
$(s, h) \models \varphi_1 * \varphi_2$  if and only if



There is a way to split the heap into two so that, together with the store, one part satisfies  $\varphi_1$  and the other satisfies  $\varphi_2$ .

## Separating implication ( $\dashv\ast$ )

$(s, h) \models \varphi_1 \dashv\ast \varphi_2$  if and only if



$$\text{dom}(h) \cap \text{dom}(h_1) = \emptyset$$
$$(s, h_1) \models \varphi_1$$



$$(s, h + h_1) \models \varphi_2$$

Whenever a (disjoint) heap that, together with the store, satisfies  $\varphi_1$  is added, the resulting memory state satisfies  $\varphi_2$ .

# Decision Problems

- Hoare proof-system requires to solve classical problems:
  - satisfiability/validity/entailment
  - weakest precondition/strongest postcondition

$$\frac{P \implies P' \quad \{P'\} C \{Q'\} \quad Q' \implies Q}{\{P\} C \{Q\}} \text{consequence rule}$$

- satisfiability is PSPACE-complete for  $SL(*, -*)$

**Note:** entailment and validity reduce to satisfiability for  $SL(*, -*)$ .

## Robustness properties

- **Acyclicity** holds for  $\varphi$  iff every model of  $\varphi$  is acyclic
- **Garbage freedom** holds for  $\varphi$  iff in every model of  $\varphi$ , each memory cell is reachable from a program variable of  $\varphi$

## C. Jansen et al., ESOP'17

**Checking for robustness properties** is  $\text{EXPSPACE}$ -complete for Symbolic Heaps with Inductive Predicates.

- Symbolic Heaps  $\implies$  no negation, no  $\neg*$ , no  $\wedge$  inside  $*$
- Inductive Predicates: akin of Horn clauses where  $*$  replaces  $\wedge$

$$P(\vec{x}) \Leftarrow \exists \vec{z} Q_1 \overset{*}{\wedge} \dots \overset{*}{\wedge} Q_n \quad \text{fv}(Q_i) \subseteq \vec{x}, \vec{z}$$

### Our Goal

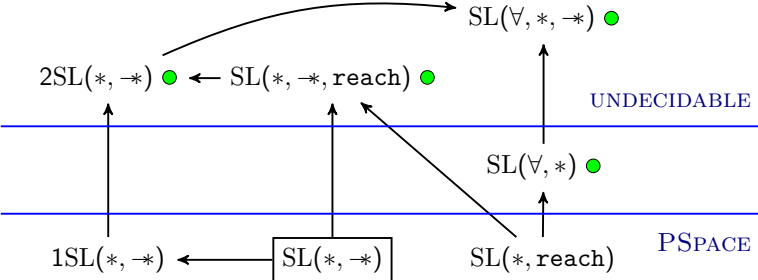
Provide similar results, but for **propositional** separation logic.

# Desiderata

We aim to an extension of propositional separation logic where

- satisfiability, validity and entailment are decidable
- in PSPACE (as propositional separation logic)
- robustness properties reduce to one of these problems

# Known extensions



## SL(\*, -\*) + reachability and one quantified variable

- $(s, h) \models \text{reach}^+(x, y) \iff h^L(s(x)) = s(y)$  for some  $L \geq 1$
- $(s, h) \models \exists u \varphi \iff$  there is  $\ell \in \text{LOC}$  s.t.  $(s[u \leftarrow \ell], h) \models \varphi$

It is only possible to quantify over the variable name  $u$ .

## Robustness properties reduce to entailment

- **Acyclicity:**  $\varphi \models \neg \exists u \text{reach}^+(u, u)$
- **Garbage freedom:**  $\varphi \models \forall u (\text{alloc}(u) \Rightarrow \bigvee_{x \in \text{fv}(\varphi)} \text{reach}(x, u))$

where  $u \notin \text{fv}(\varphi)$  and

- $\text{alloc}(x) \stackrel{\text{def}}{=} x \hookrightarrow x \text{ -* } \perp$
- $\text{reach}(x, y) \stackrel{\text{def}}{=} x = y \vee \text{reach}^+(x, y)$

## Restrictions

The logic  $1SL(*, -*, \text{reach}^+)$  is undecidable. We syntactically restrict the logic so that for each occurrence of  $\text{reach}^+(x, y)$ :

**R1** it is not on the right side of its first  $-*$  ancestor  
(seeing the formula as a tree)

**R2** if  $x = u$  then  $y = u$  (syntactically)

For example, given  $\varphi, \psi$  satisfying these conditions,

- $\text{reach}^+(u, x) * (\varphi -* \psi)$  only satisfies R1
- $\varphi -* (\text{reach}^+(x, u) -* \psi)$  satisfies both R1 and R2
- $\varphi -* (\psi * \text{reach}^+(u, u))$  only satisfies R2

**Note:** robustness properties are expressible in this fragment.

## Results

- 0 Weakening even slightly R1 leads to undecidability
- 1  $1SL_{R1}(*, -*, reach^+)$ : satisfiability is NON-ELEMENTARY (more precisely, TOWER-hard)
- 2  $1SL_{R1}^{R2}(*, -*, reach^+)$ : satisfiability is PSPACE-complete

## Proof Techniques

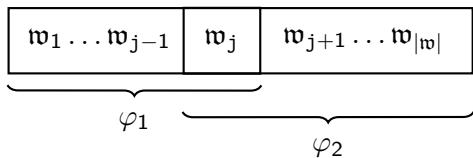
- (1) reduce *Propositional interval temporal logic under locality principle* (PITL) to a logic captured by  $1SL_{R1}(*, -*, reach^+)$
- (2) extend the *test formulae technique* used for  $SL(*, reach)$



# PITL (Moszkowski'83)

$$\varphi := \text{pt} \mid \text{a} \mid \varphi_1 \parallel \varphi_2 \mid \neg \varphi \mid \varphi_1 \wedge \varphi_2$$

- interpreted on finite non-empty words over a finite alphabet  $\Sigma$
- $w \models \text{pt} \iff |w| = 1$
- $w \models \text{a} \iff w$  headed by  $\text{a}$  (locality principle)
- $w \models \varphi_1 \parallel \varphi_2 \iff w[1 : j] \models \varphi_1$  and  $w[j : |w|] \models \varphi_2$   
for some  $j \in [1, |w|]$



- Satisfiability is decidable, but NON-ELEMENTARY

## Auxiliary Logic on Trees (ALT)

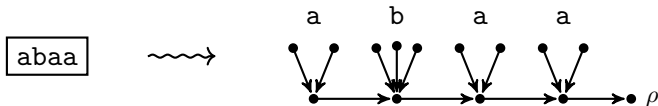
$$\varphi := \varphi_1 \wedge \varphi_2 \mid \neg\varphi \mid \varphi_1 * \varphi_2 \mid \exists u \varphi \mid T(u) \mid G(u)$$

- interpreted on acyclic memory states
- one *special* location: the root  $\rho$  of a tree
- $(s, h) \models T(u)$  iff  $s(u) \in \text{dom}(h)$  and it does reach  $\rho$
- $(s, h) \models G(u)$  iff  $s(u) \in \text{dom}(h)$  and it does not reach  $\rho$
- $\exists u \varphi$  and  $\varphi_1 * \varphi_2$  as before

**Note:** ALT is captured by  $1\text{SL}_{R1}(*, \neg*, \text{reach}^+)$ .

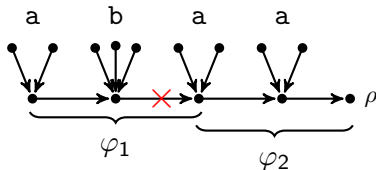
# Reducing PITL to ALT

- Easy to encode words as acyclic memory states



- Set of models encoding words can be characterised in ALT
- However, difficult to translate  $\varphi_1 \mid \varphi_2$ :

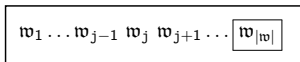
ALT cannot express properties about the set of locations in  $\text{dom}(h)$  that do not reach  $\rho$ , apart from its size



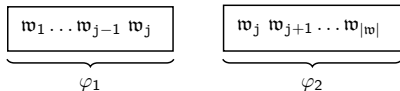
After the **cut**, left side does not reach  $\rho$  anymore.

# Reducing PITL to ALT: alternative semantics for PITL

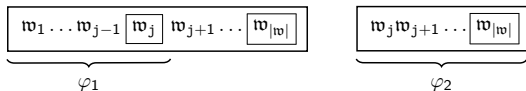
- $\boxed{a}$  marked representation of a



- $\varphi \mid \psi$  on standard semantics:



- $\varphi \mid \psi$  on marked semantics (can be simulated in ALT)



1 ALT and  $1SL_{R1}(*, \neg*, \text{reach}^+)$  are NON-ELEMENTARY

2 ALT is decidable in TOWER, as it is captured by  $SL(\forall, *)$

$1SL_{R1}^{R2}(*, -*, \text{reach}^+)$  is in PSPACE

~~$1SL_{R1}^{R2}(*, -*, reach^+)$  is in PSPACE~~

Test Formulae “technique”

## Test formulae example on a Toy Logic

$$\varphi := \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 * \varphi_2 \mid \exists u \varphi \mid \text{alloc}(u) \mid u \xrightarrow{2} u$$

where  $(s, h) \models u \xrightarrow{2} u$  iff  $h(s(u)) = \ell \neq s(u)$  and  $h(\ell) = s(u)$ .

Some formulae:

- $\#\text{loops}(2) \geq \beta \stackrel{\text{def}}{=} \overbrace{\exists u u \xrightarrow{2} u * \dots * \exists u u \xrightarrow{2} u}^{\beta-1 \text{ times } *}$
- $H_1 \stackrel{\text{def}}{=} \exists u \text{alloc}(u) \wedge \neg(\exists u \text{alloc}(u) * \exists u \text{alloc}(u))$
- $\text{rem} \geq 0 \stackrel{\text{def}}{=} \top$
- $\text{rem} \geq \beta+1 \stackrel{\text{def}}{=} \exists u : \text{alloc}(u) \wedge \neg u \xrightarrow{2} u \wedge ((\text{alloc}(u) \wedge H_1) * \text{rem} \geq \beta)$

# Test Formulae

- 1 Design an equivalence relation on models, based on the satisfaction of atomic predicates (test formulae), e.g.

$$\#loops(2) \geq \beta \qquad \text{rem} \geq \beta$$

- 2 **Show that any formula of our logic is equivalent to a Boolean combination of test formulae**, e.g.

$$\#loops(2) \geq 3 * \#loops(2) \geq 5 \iff \#loops(2) \geq 8$$

- 3 Prove small-model property for the logic of test formulae.



# (1) Designing Test Formulae

- Fix  $\alpha \in \mathbb{N}^+$
- Let  $\text{Test}(\alpha)$  be the **finite** set of predicates:  
 $\{\# \text{loops}(2) \geq \beta, \text{rem} \geq \gamma \mid \beta \in [1, \mathcal{L}(\alpha)], \gamma \in [1, \mathcal{G}(\alpha)]\}$   
for some functions  $\mathcal{L}$  and  $\mathcal{G}$  in  $[\mathbb{N} \rightarrow \mathbb{N}]$

Indistinguishability relation  $(s, h) \approx_\alpha (s', h')$

for every  $T \in \text{Test}(\alpha)$ ,  $(s, h) \models T$  iff  $(s', h') \models T$

**Note:**  $\alpha$  is related to the number of occurrences of  $*$  and  $\rightarrow^*$  in a formula of separation logic.

## (2) \* elimination Lemma

We want to design  $\text{Test}(\alpha)$  so that the following result holds

### Hypothesis:

- $(s, h) \approx_{\alpha} (s', h')$
- $\alpha_1, \alpha_2 \in \mathbb{N}^+$  s.t.  $\alpha_1 + \alpha_2 = \alpha$
- $h_1 + h_2 = h$

**Thesis:** there are  $h'_1, h'_2$  s.t.

- $h'_1 + h'_2 = h'$
- $(s, h_1) \approx_{\alpha_1} (s', h'_1)$
- $(s, h_2) \approx_{\alpha_2} (s', h'_2)$

**Note:** it can be restated as an EF-style game. Spoiler splits  $\alpha$  and  $h$ , Duplicator has to mimic the split on  $h'$  so that  $\approx$  still holds.

## (2) \* elimination Lemma

We want to design  $\text{Test}(\alpha)$  so that the following result holds

### Hypothesis:

- $(s, h) \approx_\alpha (s', h')$

- $\alpha$

- $h$

$$\left\{ \begin{array}{l} \# \text{loops}(2) \geq \beta, \quad \beta \in [1, \mathcal{L}(\alpha)] \\ \text{rem} \geq \gamma \quad \gamma \in [1, \mathcal{G}(\alpha)] \end{array} \right\}$$

### Thesis:

- $h$

- (s find  $\mathcal{L}$  and  $\mathcal{G}$  so that lemma holds.

- $(s, h_2) \approx_{\alpha_2} (s', h'_2)$

**Note:** it can be restated as an EF-style game. Spoiler splits  $\alpha$  and  $h$ , Duplicator has to mimic the split on  $h'$  so that  $\approx$  still holds.

## Finding $\mathcal{G}$ for $\text{rem} \geq \gamma$ formulae

Given  $h = h_1 + h_2$ , every location not in a loop of size 2 of  $h$  cannot be in a loop of size 2 of  $h_1$  or  $h_2$ . Then  $\mathcal{G}$  must satisfy

$$\mathcal{G}(\alpha) \geq \max_{\substack{\alpha_1, \alpha_2 \in \mathbb{N}^+ \\ \alpha_1 + \alpha_2 = \alpha}} (\mathcal{G}(\alpha_1) + \mathcal{G}(\alpha_2))$$

## Finding $\mathcal{L}$ for $\#\text{loops}(2) \geq \beta$ formulae

Take  $h = h_1 + h_2$ . Given a loop of size 2 of  $h$ , two cases:

- both locations of the loop are in the same heap ( $h_1$  or  $h_2$ );
- one location of the loop is in  $h_1$  and the other is in  $h_2$ .

$$\mathcal{L}(\alpha) \geq \max_{\substack{\alpha_1, \alpha_2 \in \mathbb{N}^+ \\ \alpha_1 + \alpha_2 = \alpha}} (\mathcal{L}(\alpha_1) + \mathcal{L}(\alpha_2) + \mathcal{G}(\alpha_1) + \mathcal{G}(\alpha_2))$$

## Finding $\mathcal{L}$ and $\mathcal{G}$

We have the inequalities

$$\mathcal{G}(1) \geq 1 \quad \mathcal{G}(\alpha) \geq \max_{\substack{\alpha_1, \alpha_2 \in \mathbb{N}^+ \\ \alpha_1 + \alpha_2 = \alpha}} (\mathcal{G}(\alpha_1) + \mathcal{G}(\alpha_2))$$

$$\mathcal{L}(1) \geq 1 \quad \mathcal{L}(\alpha) \geq \max_{\substack{\alpha_1, \alpha_2 \in \mathbb{N}^+ \\ \alpha_1 + \alpha_2 = \alpha}} (\mathcal{L}(\alpha_1) + \mathcal{L}(\alpha_2) + \mathcal{G}(\alpha_1) + \mathcal{G}(\alpha_2))$$

Which admit  $\mathcal{G}(\alpha) = \alpha$  and  $\mathcal{L}(\alpha) = \frac{1}{2}\alpha(\alpha + 3) - 1$  as a solution.

An indistinguishability relation built on the set

$$\left\{ \begin{array}{l} \# \text{loops}(2) \geq \beta, \\ \text{rem} \geq \gamma \end{array} \left| \begin{array}{l} \beta \in \left[ 1, \frac{1}{2}\alpha(n+3) - 1 \right] \\ \gamma \in [1, \alpha] \end{array} \right. \right\}$$

satisfy the \* elimination Lemma.

### (3) Test formulae, after $*$ elimination

**Hypothesis:** Two family of test formulae, such that

- captures the atomic predicates of the Toy Logic
- satisfies the  $*$  elimination Lemma (and  $\exists$  elimination Lemma)

**Thesis:** for every formulae  $\varphi$  of Toy Logic,  
by taking  $\alpha \geq |\varphi|$  we have

- If  $(s, h) \approx_\alpha (s, h')$  then we have  $(s, h) \models \varphi$  iff  $(s, h') \models \varphi$ .
- $\varphi$  is equivalent to a Boolean combination of test formulae.

### Small-model property

- 1 Small-model property for Boolean combination of test formulae carries over to Toy Logic.
- 2 All bounds are polynomial  $\implies$  test formulae in PSPACE
- 3 Toy Logic is in PSPACE

# $1SL_{R1}^{R2}(*, \neg*, \text{reach}^+)$ is in PSPACE

$$\pi := x = y \mid x \hookrightarrow y \mid \text{emp} \mid \underline{\mathcal{A} \neg* \mathcal{C}} \text{ (R1)}$$

$$\mathcal{C} := \pi \mid \mathcal{C} \wedge \mathcal{C} \mid \neg \mathcal{C} \mid \exists u \mathcal{C} \mid \mathcal{C} * \mathcal{C}$$

$$\mathcal{A} := \pi \mid \underline{\text{reach}^+(v_1, v_2)} \mid \mathcal{A} \wedge \mathcal{A} \mid \neg \mathcal{A} \mid \exists u \mathcal{A} \mid \mathcal{A} * \mathcal{A}$$

where (R2) if  $v_1 = u$  then  $v_2 = u$

## Not so easy...

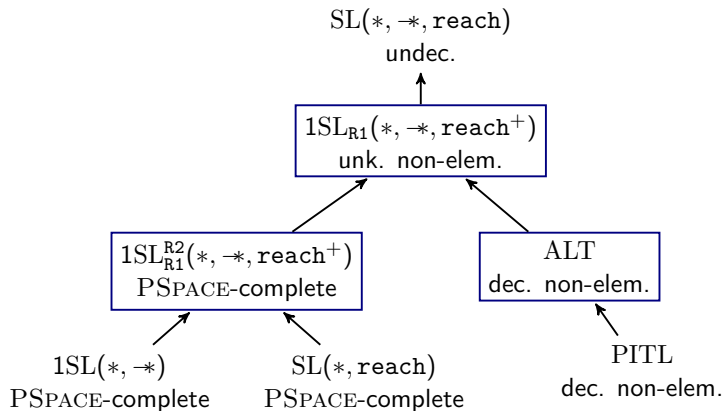
- Find the right set of test formulae that capture the logic
- Asymmetric  $\mathcal{A} \neg* \mathcal{C}$ .
  - two indistinguishability relation, two sets of test formulae
  - two  $*$  and two  $\exists$  elimination Lemmata
  - $\neg*$  elimination Lemma that glues the two relations

If you like bounds:  $\text{Test}(X, \alpha)$  for the  $\mathcal{A}$  fragment

$$\left\{ \begin{array}{l} v_1 = v_2, \text{sees}_X(v_1, v_2) \geq \beta^{\downarrow} \\ \#\text{loop}_X(\beta) \geq \beta^{\circ}, \#\text{loop}_X^{\uparrow} \geq \beta^{\circ} \\ \#\text{pred}_X^A(x) \geq \beta, \text{size}_X^A \geq \beta \\ u \in \text{sees}_X(v_1, v_2) \geq (\overleftarrow{\beta}, \overrightarrow{\beta}) \\ u = v_1, u \in \text{loop}_X(\beta), u \in \text{loop}_X^{\uparrow} \\ u \in \text{pred}_X^A(x), u \in \text{size}_X^A \end{array} \right\} \left| \begin{array}{l} \beta^{\downarrow} \in [1, \frac{1}{6}(\alpha+1)(\alpha+2)(\alpha+3)] \\ \beta^{\circ} \in [1, \frac{1}{2}\alpha(\alpha+3) - 1], \beta \in [1, \alpha] \\ \overleftarrow{\beta} \in [1, \frac{1}{6}\alpha(\alpha+1)(\alpha+2) + 1] \\ \overrightarrow{\beta} \in [1, \frac{1}{2}\alpha(\alpha+3)] \\ x \in X, v_1, v_2 \in \mathcal{A}\text{TERM}_X \end{array} \right.$$

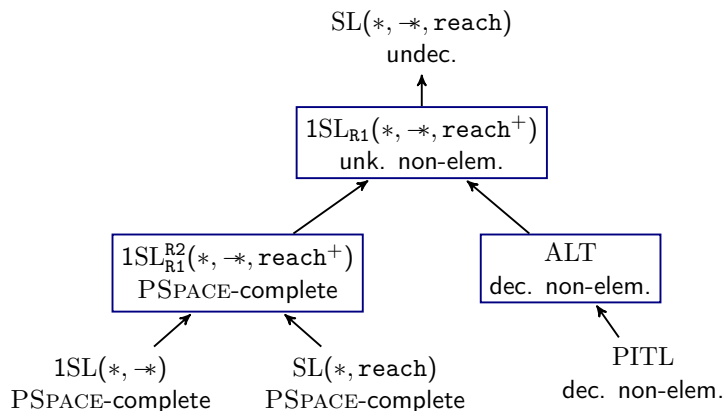


# Recap



- $1SL_{R1}^{R2}(*, -*, reach^+)$  strictly generalise other PSPACE-complete extensions of propositional separation logic
- Can be used to check for robustness properties

# Recap



- $ALT$  seems to be an interesting tool for reductions, as it is a fragment or it is easily captured by many logics in TOWER e.g.  $QCTL(U)$ ,  $MSL(\diamond, \langle U \rangle, *)$ ,  $2SL(*)$