

# Principles of Abstract Interpretation

## MIT press

### Ch. 46, Points-To Analysis

Patrick Cousot

[pcousot.github.io](http://pcousot.github.io)

[PrAbsInt@gmail.com](mailto:PrAbsInt@gmail.com)    [github.com/PrAbsInt/](https://github.com/PrAbsInt/)

These slides are available at  
<http://github.com/PrAbsInt/slides/slides-46--points-to-analysis-PrAbsInt.pdf>

## Ch. 46, Points-To Analysis

## Pointers in C (not considering `malloc`)

## Pointers in C

- We consider static memory allocation where the memory addresses (where pointer variables of a program can point to) are determined statically by the compiler (as opposed to dynamic memory allocation by malloc in C [DBLP:books/ph/KernighanR88] during program execution).

## Example 46.1

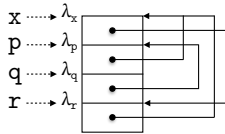
```
#include <stdio.h>
#include <stdlib.h>
int main () {
    int x,*p,**q,*r;
    x = &p;
    p = &x;
    q = &p;
    r = *q;
    *r = &r;
}
```

Where is **p** pointing to on program exit?

## Example 46.1

The points-to analysis of the following C program

```
#include <stdio.h>
#include <stdlib.h>
int main () {
    int x,*p,**q,*r;
    x = &p;
    p = &x;
    q = &p;
    r = *q;
    *r = &r;
}
```



Memory  $\mu$ :

$\mu(\lambda_x)$	$\lambda_r$
$\mu(\lambda_p)$	$\lambda_x$
$\mu(\lambda_q)$	$\lambda_p$
$\mu(\lambda_r)$	$\lambda_x$

should be able to determine that  $p$  points to  $r$  on program exit.

## Pointer or points-to analysis

- Pointer analysis, or points-to analysis consists in automatically inferring where pointer variables of a program can point to.
- Points-to analysis started at the end of the seventies [DBLP:journals/cacm/Barth77; DBLP:conf/popl/Weihl80] and has been extensively studied since then [DBLP:conf/lfp/Hudak86; DBLP:conf/pepm/Deutsch95; DBLP:conf/paste/Hind01; DBLP:journals/csur/KanvarK16].
- Our objective is to explore the formal design of two classical points-to analyzers [Andersen94-Thesis; DBLP:conf/popl/Steensgaard96] as opposed to the resolution of postulated constraints [Andersen94-Thesis] or a postulated type inference system [DBLP:conf/popl/Steensgaard96].

# Syntax of the pointer language, section 46.1



## Pointer syntax

The syntax of programs of section 4.1 is extended as follows.

$p, q, \dots, x, y, \dots$	$\in \mathcal{V}$	variables
$p, q, \dots$	$\in \vec{\mathcal{V}} \subseteq \mathcal{V}$	pointer variables
$A$	$::= \dots$	arithmetic expressions
	$x$	variable
	$\text{NULL}$	null pointer
	$\&x$	address of variable
	$*p$	dereferenced pointer variable
$B$	$::= \dots$	boolean expressions
	$x == y$	equality
	$p == \text{NULL}$	pointer equality to <i>nil</i>
$S$	$::= \dots$	statement
	$x = A ;$	assignment
	$*p = A ;$	assignment to a dereferenced pointer

## Comments on the pointer syntax

- Variables are statically allocated in memory at locations/addresses determined by the compiler and loader. These locations can be values of pointer variables.
- The null pointer `NULL` designates no location at all.
- `&x` is the location of variable `x` in memory.
- `*p` dereferences pointer variable `p`. It is the location contained in the location of variable `p`.
- One can test for nullity `p == NULL` or `x == y` equality of pointer variables.
- Assignment is of an integer or pointer value to a variable location `x = A ;` or an indirect assignment `*p = A ;` to a dereferenced pointer variable.
- Some languages like Pascal [DBLP:journals/acta/Wirth71] do not have such primitives but use them implicitly e.g. when passing parameters by reference (also called by variable) or for array parameters (to avoid a copy of the array).

# Semantics of the pointer language

## Pointer semantic domains, section 46.2

Let us define

$$z \in \mathbb{Z}$$

integers

$$l \in \mathbb{L}$$

enumerable set of locations (addresses) ( $\mathbb{L} \cap \mathbb{Z} = \emptyset$ )

$$\lambda \in \mathbb{L}_v \triangleq \mathbb{V} \rightarrow \mathbb{L}$$

locations of variables ( $\lambda$  injective: unique

location  $\lambda_x$  of variable  $x$  in memory)

$$\text{nil} \notin \mathbb{L}$$

null pointer (value denoted by **NULL**)

$$v \in \mathbb{V} \triangleq \mathbb{Z} \cup \mathbb{L} \cup \{\text{nil}\}$$

values

$$\Omega \notin \mathbb{V}$$

error

$$v \in \mathbb{V}_\Omega \triangleq \mathbb{V} \cup \{\Omega\}$$

values or error

## Pointer semantic domains, section 46.2

- For a given program, the locations  $\lambda$  are fixed by the compiler, linker and loader.
- So we can restrict values to these specific locations.

$l \in \mathbb{L}_\lambda \triangleq \{\lambda_x \in \mathbb{L} \mid x \in \mathbb{V}\}$	set of locations of all variables
$\mu \in \mathbb{M}_\lambda \triangleq \mathbb{L}_\lambda \rightarrow \mathbb{V}$	memories (values of locations)
$v \in \mathbb{V}_\lambda \triangleq \mathbb{Z} \cup \mathbb{L}_\lambda \cup \{\text{nil}\}$	values
$P \in \mathbb{P}_\lambda^{\mathbb{M}} \triangleq \wp(\mathbb{M}_\lambda)$	memory properties

*TODO: Aervoir*

## Comments on the pointer semantic domains

- The compiler and loader initially assign a unique location  $\lambda_x$  to each program variable  $x$
- The memory allocation function  $\lambda \triangleq x \in \mathcal{V} \mapsto \lambda_x \in \mathbb{L}$  is therefore bijective.
- Values returned by expressions are
  - integers,
  - locations,
  - the null pointer `nil` (denoted NULL), or
  - an error  $\Omega$  (e.g. when dereferencing a null pointer which in practice often stops the execution with a segmentation fault (raised after a memory access violation in hardware with memory protection) or bus error (e.g. in misaligned memory access that hardware cannot access)).
- A memory  $\mu$  maps locations  $\lambda_x$  of variables  $x \in \mathcal{V}$  to their content  $\mu(\lambda_x)$  which is the value of  $x$ .

## Comments on the pointer semantic domains

- In case of error, execution is assumed to stop immediately.
- This is contrary to some programming languages such as C [[DBLP:books/ph/KernighanR88](#)] where the program behavior is undefined in case of error.
- In C, errors need not be signaled during execution so the undefined execution may have unpredictable effects like erasing data or changing/destroying the program code.
- Of course it is impossible to define the semantics of such programs.
- So, more restrictively, we consider such undefined behaviors to be an error which immediately stops execution.
- The soundness of the analysis is precisely defined in section 46.12.



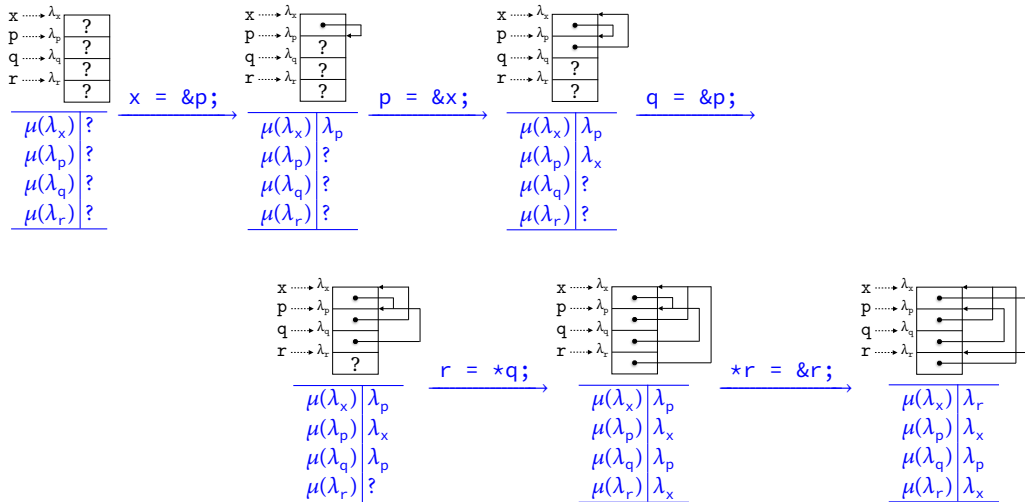
## Invariant properties

- Invariant properties are sets of memories (but not an error since execution stops when the evaluation of a pointer expression is erroneous).

## Reachability semantics of the pointer language, section 46.4

- We should extend the trace semantics and then abstract it into the reachability semantics of 19
- For brevity, we directly and postulate the reachability semantics  $\widehat{\mathcal{S}}^{\vec{e}}$  of the pointer manipulating statements.
- Of course it could be obtained by abstraction as shown in chapter 20, “Computational Design of the Forward Reachability Semantics” from an extension of the prefix trace semantics of chapter 6, “Structural Deductive Stateless Prefix Trace Semantics” to include pointers.

## Example 46.5 of execution trace of example 46.1



## Reachability semantics of the pointer language, section 46.4

- The reachability semantics

$$\hat{\mathcal{S}}_{\lambda}^{\vec{\ell}} \llbracket S \rrbracket \mathcal{P}_0^{\ell}$$

defines the memory states reachable at program point  $\ell \in \text{labx} \llbracket S \rrbracket$  of program component  $S$  when its execution starts in an initial memory state satisfying precondition  $\mathcal{P}_0$ .

- The reachability semantics  $\hat{\mathcal{S}}_{\lambda}^{\vec{\ell}}$  is relative to a given assignment  $\lambda$  of a location  $\lambda_x$  to each variable  $x \in \text{vars} \llbracket S \rrbracket$  of program component  $S$ .
- The forward reachability semantics  $\hat{\mathcal{S}}^{\vec{\ell}} \llbracket S \rrbracket$  of a program component  $S$  in chapter 19 was operating on environments mapping variables to values,

$$\hat{\mathcal{S}}^{\vec{\ell}} \llbracket S \rrbracket \in \wp(\text{Env}^{\ell}) \rightarrow \text{labx} \llbracket S \rrbracket \rightarrow \wp(\text{Env}^{\ell})$$

## Reachability semantics of the pointer language, section 46.4

- The forward reachability semantics is now operating on memories (for a fixed allocation  $\lambda$  of locations to variables),

$$\hat{\mathcal{S}}_{\lambda}^{\vec{e}}[[S]] \in \wp(M_{\lambda}^e) \rightarrow \text{labx}[[S]] \rightarrow \wp(M_{\lambda}^e)$$

- So the value of a variable  $x$  which was  $\rho(x)$  for environments  $\rho \in \mathbb{E}_v$  is now  $\mu(\lambda_x)$  for memories  $\mu \in M_{\lambda}$  with given memory allocation  $\lambda$ .
- By using symbolic addresses, the analysis is independent of the particular memory allocation  $\lambda$  determined at compile, linking, and loading time.
- We first extend the semantics of arithmetic and boolean expressions in section 3.6 to allow for pointer operations and comparison.

## Definition of the semantics of arithmetic and pointer expressions

$$\mathcal{A} \in \lambda \in \mathbb{L}_v \rightarrow \mathcal{A} \rightarrow \mathbb{M}_\lambda \rightarrow \mathbb{V}_\lambda \cup \{\Omega\} \quad (46.2)$$

$$\mathcal{A}_\lambda[[1]] \mu \triangleq 1$$

$$\mathcal{A}_\lambda[[x]] \mu \triangleq \mu(\lambda_x)$$

$$\mathcal{A}_\lambda[[A_1 - A_2]] \mu \triangleq (\mathcal{A}_\lambda[[A_1]] \mu \in \mathbb{Z} \wedge \mathcal{A}_\lambda[[A_2]] \mu \in \mathbb{Z} ? \mathcal{A}_\lambda[[A_1]] \mu - \mathcal{A}_\lambda[[A_2]] \mu : \Omega)$$

$$\mathcal{A}_\lambda[[\text{NULL}]] \mu \triangleq \text{nil}$$

$$\mathcal{A}_\lambda[[\&x]] \mu \triangleq \lambda_x$$

$$\mathcal{A}_\lambda[[*p]] \mu \triangleq (\mu(\lambda_p) \in \mathbb{Z} \cup \{\text{nil}\} ? \Omega : \mu(\mu(\lambda_p)))$$

## Comments on the semantics of arithmetic and pointer expressions

- The evaluation of an expression may return an error  $\Omega$ .
- `NULL` denotes the null pointer `nil`.
- `&x` denotes the location of the variable `x`.
- It is assumed that the locations  $\lambda$  allocated by the compiler for different variables cannot overlap and be erroneous so  $\lambda \triangleq x \in \mathcal{V} \mapsto \lambda_x \in \mathbb{L}$  is bijective.
- The value of a variable `x` is obtained by looking at its location  $\lambda_x$  and then at the value  $\mu(\lambda_x)$  in the memory at that location.
- For `*p` there is one more level of indirection  $\mu(\mu(\lambda_p))$  unless `p` is not a pointer in which case this is a runtime error.
- Otherwise the definition is similar to (3.4), except that errors are propagated.

## Definition of the semantics of boolean expressions

$$\mathcal{B} \in \lambda \in \mathbb{L}_v \rightarrow \mathcal{B} \rightarrow \mathbb{M}_\lambda \rightarrow \mathbb{B} \cup \{\Omega\}$$

$$\mathcal{B}_\lambda[A_1 < A_2] \mu \triangleq (\mathcal{A}_\lambda[A_1] \mu \in \mathbb{Z} \wedge \mathcal{A}_\lambda[A_2] \mu \in \mathbb{Z} ? \mathcal{A}_\lambda[A_1] \mu < \mathcal{A}_\lambda[A_2] \mu : \Omega)$$

$$\mathcal{B}_\lambda[p == \text{NULL}] \mu \triangleq (\mu(\lambda_p) = \text{nil})$$

$$\mathcal{B}_\lambda[x == y] \mu \triangleq (\mu(\lambda_x) = \mu(\lambda_y))$$

$$\mathcal{B}_\lambda[B_1 \text{ nand } B_2] \mu \triangleq (\mathcal{B}_\lambda[B_1] \mu \neq \Omega \wedge \mathcal{B}_\lambda[B_2] \mu \neq \Omega ? \mathcal{B}_\lambda[B_1] \mu \uparrow \mathcal{B}_\lambda[B_2] \mu : \Omega)$$



# Definition of the reachability of an arithmetic or pointer assignment statement

$S ::= x = A ;$

$$\widehat{\mathcal{S}}^{\vec{\ell}} \in \lambda \in \mathbb{L}_V \rightarrow \mathcal{S} \rightarrow \wp(\mathbb{M}_{\lambda}^{\vec{\ell}}) \rightarrow \mathbb{L} \rightarrow \wp(\mathbb{M}_{\lambda}^{\vec{\ell}}) \quad (46.3)$$

$$\begin{aligned} \widehat{\mathcal{S}}_{\lambda}^{\vec{\ell}}[\![S]\!] \mathcal{P}_0^{\ell} = & \left( \ell = \text{at}[\![S]\!] \text{ ? } \mathcal{P}_0 \right. \\ & \left. \parallel \ell = \text{after}[\![S]\!] \text{ ? } \text{assign}_{\lambda}^{\vec{\ell}}[\![x, A]\!] \mathcal{P}_0 \right. \\ & \left. : \emptyset \right) \end{aligned}$$

$$\text{assign}_{\lambda}^{\vec{\ell}}[\![x, A]\!] \mathcal{P}_0 \triangleq \{ \mu[\lambda_x \leftarrow \mathcal{A}_{\lambda}[\![A]\!] \mu] \mid \mu \in \mathcal{P}_0 \wedge \mathcal{A}_{\lambda}[\![A]\!] \mu \neq \Omega \}$$

$$\text{assign}_{\lambda}^{\vec{\ell}}[\![x, A]\!] \mathcal{P}_0 \triangleq \{ \langle \mu_0, \mu[\lambda_x \leftarrow \mathcal{A}_{\lambda}[\![A]\!] \mu] \rangle \mid \langle \mu_0, \mu \rangle \in \mathcal{P}_0 \wedge \mathcal{A}_{\lambda}[\![A]\!] \mu \neq \Omega \}$$

## Comments on the semantics of arithmetic and pointer expressions

- The semantics (19.12) of an assignment statement  $S ::= x = A ;$  maps the initial states to themselves on entry and to themselves as modified by the assignment on exit.
- An arithmetic or pointer assignment modifies the memory (but not the allocation of locations to variables which are chosen statically by the compiler and loader hence immutable during execution).
- Moreover, if the evaluation of the expression  $A$  returns an error  $\Omega$  then the execution stops so there is no reachable successor state (so the postcondition is  $\emptyset$ ).
- The difference with (19.12) is that the value of a variable  $x$  has to be fetched in memory  $\mu$  at the location  $\lambda_x$  for that variable so  $\rho(x) = \mu(\lambda_x)$  and that execution stops on error (there was no error in (19.12) but e.g. integer overflow could be handled similarly by stopping execution).

# Definition of the reachability of a dereferenced pointer assignment statement

$S ::= *p = A ;$

$$\widehat{\mathcal{S}}^{\vec{\ell}} \in \lambda \in \mathbb{L}_v \rightarrow \mathcal{S} \rightarrow \wp(\mathbb{M}_\lambda^{\vec{\ell}}) \rightarrow \mathbb{L} \rightarrow \wp(\mathbb{M}_\lambda^{\vec{\ell}}) \quad (46.4)$$

$$\begin{aligned} \widehat{\mathcal{S}}_\lambda^{\vec{\ell}}[[S]] \mathcal{P}_0 \ell &= (\ell = \text{at}[[S]] \text{ ? } \mathcal{P}_0 \\ &\quad \parallel \ell = \text{after}[[S]] \text{ ? } \text{assign}_\lambda^{\vec{\ell}}[[*p, A]] \mathcal{P}_0 \\ &\quad : \emptyset) \end{aligned}$$

$$\text{assign}_\lambda^{\vec{r}}[[*p, A]] \mathcal{P}_0 \triangleq \{\mu[\mu(\lambda_p) \leftarrow \mathcal{A}_\lambda[A] \mu] \mid \mu \in \mathcal{P}_0 \wedge \mu(\lambda_p) \in \mathbb{L} \wedge \mathcal{A}_\lambda[A] \mu \neq \Omega\}$$

$$\begin{aligned} \text{assign}_\lambda^{\vec{R}}[[*p, A]] \mathcal{P}_0 &\triangleq \{\langle \mu_0, \mu[\mu(\lambda_p) \leftarrow \mathcal{A}_\lambda[A] \mu] \rangle \mid \langle \mu_0, \mu \rangle \in \mathcal{P}_0 \wedge \mu(\lambda_p) \in \mathbb{L} \wedge \\ &\quad \mathcal{A}_\lambda[A] \mu \neq \Omega\} \end{aligned}$$

## Comments on the reachability semantics of a dereferenced pointer assignment statement $S ::= *p = A ;$

- The assignment to a dereferenced pointer  $*p = A ;$  assigns the value of expression  $A$  to the memory pointed to by pointer  $p$ .
- This is an error if  $p$  does not contain a pointer i.e. is an integer or `nil` so that  $\mu(\lambda_p) \notin \mathbb{L}$ .
- An error can also be raised by the evaluation of expression  $A$  (e.g. by dereferencing a null pointer).
- In case of error, the execution stops so there is no reachable successor state.

## Reachability semantics of a test B

$$\text{test}^{\vec{e}}, \overline{\text{test}}^{\vec{e}} \in \lambda \in \mathbb{L}_v \rightarrow \mathcal{B} \rightarrow \wp(\mathcal{M}_\lambda^{\vec{e}}) \rightarrow \wp(\mathcal{M}_\lambda^{\vec{e}}) \quad (46.6)$$

$$\text{test}_\lambda^{\vec{r}}[\![B]\!] \mathcal{P}_0 \triangleq \{\mu \in \mathcal{P}_0 \mid \mathcal{B}_\lambda[\![B]\!] \mu = \text{tt}\}$$

$$\text{test}_\lambda^{\vec{R}}[\![B]\!] \mathcal{P}_0 \triangleq \{\langle \mu_0, \mu \rangle \in \mathcal{P}_0 \mid \mathcal{B}_\lambda[\![B]\!] \mu = \text{tt}\}$$

$$\overline{\text{test}}_\lambda^{\vec{r}}[\![B]\!] \mathcal{P}_0 \triangleq \{\mu \in \mathcal{P}_0 \mid \mathcal{B}_\lambda[\![B]\!] \mu = \text{ff}\}$$

$$\overline{\text{test}}_\lambda^{\vec{R}}[\![B]\!] \mathcal{P}_0 \triangleq \{\langle \mu_0, \mu \rangle \in \mathcal{P}_0 \mid \mathcal{B}_\lambda[\![B]\!] \mu = \text{ff}\}$$

## Comments on the reachability semantics of a test $B$

- The reachability test transformers of (19.16) are adapted to the memory model. In case of error  $\mathcal{B} \llbracket B \rrbracket \lambda \mu = \Omega \notin \{\mathbf{tt}, \mathbf{ff}\}$ , the execution stops so there is no reachable successor state.
- Observe that, by our definition of reachability, the fatal error  $\Omega$  is never reachable by our assumption that execution stop in case of error. To improve error reporting, it would also be possible to keep errors in the reachable states (even with e.g. a program point to better track the origin of errors).

## Reachability semantics of all other program components

- The reachability semantics  $\hat{\mathcal{S}}_{\lambda}^{\vec{e}} \mathcal{P}_0^{\ell}$  of all other program components in (19.12) is modified in the same way, replacing environments  $\rho$  by memories  $\mu$  for the given location assignment  $\lambda$  and stopping execution when expressions (e.g. boolean conditions in conditional or iteration statements) return an error.
- For example the reachability semantics (19.22) and (19.23) of conditional statements and (19.16) of an iteration statement remain the same but for the definition of  $\text{test}^{\vec{e}} \llbracket B \rrbracket$  by (46.6)
- Similarly for  $\overline{\text{test}}^{\vec{e}} \llbracket B \rrbracket$ . In case the evaluation  $\mathcal{B} \llbracket B \rrbracket \lambda \mu$  of the boolean expression  $B$  returns an error  $\Omega$ ,  $\text{test}^{\vec{e}} \llbracket B \rrbracket X^{(\ell)}$  is the empty set  $\emptyset$ , so by theorem 19.36,  $\hat{\mathcal{S}}^{\vec{e}} \llbracket S_b \rrbracket (\text{test}^{\vec{e}} \llbracket B \rrbracket X^{(\ell)})^{\ell'}$  is also the empty set  $\emptyset$  i.e. there are no further reachable states since the trace execution is stopped by the error.

## Abstract interpreters for the pointer language, section 46.5



## Objective

- We extend the flow sensitive static analysis of section 21.2 and the flow insensitive static analysis of section 45.3 to handle the pointer statements of section 46.1.

## Flow sensitive abstract interpreter for the pointer language

- The abstract interpreter of section 21.2 is designed for an abstract domain

$$\mathbb{D}^{\alpha} \triangleq \langle \mathbb{P}^{\alpha}, \sqsubseteq^{\alpha}, \perp^{\alpha}, \sqcup^{\alpha}, \text{assign}_{\lambda}^{\alpha}[\![x, A]\!], \text{test}^{\alpha}[\![B]\!], \overline{\text{test}}^{\alpha}[\![B]\!] \rangle$$

is unchanged

- but for the concretization  $\gamma_{\lambda} \in \mathbb{P}_{\lambda}^{\alpha} \longrightarrow \wp(\mathbb{M}_{\lambda}^{\ell})$  to the memory collecting domain  $\wp(\mathbb{M}_{\lambda}^{\ell})$  instead of the environment collecting domain  $\wp(\mathbb{E}^{\ell})$ .
- and (21.7) becomes

$$\mathbb{D}_{\lambda}^{\alpha} \triangleq \langle \mathbb{P}_{\lambda}^{\alpha}, \sqsubseteq_{\lambda}^{\alpha}, \perp_{\lambda}^{\alpha}, \sqcup_{\lambda}^{\alpha}, \text{assign}_{\lambda}^{\alpha}[\![x, A]\!], \text{assign}_{\lambda}^{\alpha}[\![*p, A]\!], \text{test}_{\lambda}^{\alpha}[\![B]\!], \overline{\text{test}}_{\lambda}^{\alpha}[\![B]\!] \rangle$$

*Flow-sensitive abstract semantics of an assignment statement  $S ::= x = A ;$*

$$\begin{aligned} \widehat{\mathcal{S}}^\alpha &\in \lambda \in \mathbb{L}_v \rightarrow \mathcal{S} \rightarrow \mathbb{P}_\lambda^\alpha \rightarrow \mathbb{L} \rightarrow \mathbb{P}_\lambda^\alpha \\ \widehat{\mathcal{S}}_\lambda^\alpha \overline{P} \ell &\triangleq (\ell = \text{at}[\![S]\!] \text{ ? } \overline{P} \\ &\quad \parallel \ell = \text{after}[\![S]\!] \text{ ? } \text{assign}_\lambda^\alpha[\![x, A]\!] \lambda \overline{P} \\ &\quad \text{: } \perp^\alpha) \end{aligned} \tag{46.8}$$

where  $\text{assign}_\lambda^\alpha[\![x, A]\!] \in \mathbb{L} \rightarrow \mathbb{P}^\alpha \rightarrow \mathbb{P}^\alpha$  satisfies  $\text{assign}_\lambda^{\vec{\theta}}[\![x, A]\!] \circ \gamma_\lambda \subseteq \gamma_\lambda \circ \text{assign}_\lambda^\alpha[\![x, A]\!]$ .

*Flow-sensitive abstract semantics of a dereferenced pointer assignment statement*  $S ::= *p = A ;$

$$\begin{aligned} \widehat{\mathcal{S}}^{\alpha} &\in \lambda \in \mathbb{L}_V \rightarrow \mathcal{S} \rightarrow \mathbb{P}_{\lambda}^{\alpha} \rightarrow \mathbb{L} \rightarrow \mathbb{P}_{\lambda}^{\alpha} & (46.9) \\ \widehat{\mathcal{S}}_{\lambda}^{\alpha} \llbracket S \rrbracket \overline{P}^{\ell} &\triangleq ( \ell = \text{at} \llbracket S \rrbracket \text{ ? } \overline{P} \\ &\quad \parallel \ell = \text{after} \llbracket S \rrbracket \text{ ? } \text{assign}_{\lambda}^{\alpha} \llbracket *p, A \rrbracket \overline{P} \\ &\quad \circ \perp_{\lambda}^{\alpha} ) \end{aligned}$$

where  $\text{assign}_{\lambda}^{\alpha} \llbracket *p, A \rrbracket \in \mathbb{P}_{\lambda}^{\alpha} \rightarrow \mathbb{P}_{\lambda}^{\alpha}$  satisfies  $\text{assign}_{\lambda}^{\vec{\alpha}} \llbracket *p, A \rrbracket \circ \gamma_{\lambda} \subseteq \gamma_{\lambda} \circ \text{assign}_{\lambda}^{\alpha} \llbracket *p, A \rrbracket$ .

- The abstract semantics of conditional and iteration statements is unchanged except for the use of (46.6) for the collecting semantics of tests.

**Theorem (46.10) Well-definedness of the reachability semantics and the abstract interpreter** The abstract interpreter  $\widehat{\mathcal{S}}_\lambda^\alpha[[S]]$  for the abstract domain  $\mathbb{D}_\lambda^\alpha$  assumed to be well-defined by definition 21.1 (and assuming  $\text{assign}_\lambda^\alpha[[x, A]]$ ,  $\text{test}_\lambda^\alpha[[B]]$ , and  $\overline{\text{test}}_\lambda^\alpha[[B]]$  to be continuous) is well-defined and continuous for any program component  $rS \in \mathcal{PC}$ .

**Proof** The language extension of section 46.1 preserves the well-definedness of the abstract interpreter in theorem 21.16, that of the reachability semantics in corollary ??, as well as the soundness of the abstract interpreted theorem 27.4 (in particular for the collecting semantics (46.3) and (46.4)). □

## Flow insensitive abstract interpreter for the pointer language

- As in chapter 45, “Flow-Insensitive Static Analysis,” the flow insensitive abstract semantics for the pointer language is obtained by the  $\alpha_g$ -abstraction of its flow insensitive semantics.

*Flow-insensitive semantics of an assignment statement  $S ::= x = A ;$*

$$\begin{aligned} \widehat{\mathcal{S}}^{\text{gf}\alpha} &\in \lambda \in \mathbb{L}_v \rightarrow \mathcal{S} \rightarrow \mathbb{P}_\lambda^\alpha \rightarrow \mathbb{P}_\lambda^\alpha \\ \widehat{\mathcal{S}}_\lambda^{\text{gf}\alpha} \llbracket S \rrbracket \overline{P} &\triangleq \overline{P} \sqcup_\lambda^\alpha \text{assign}_\lambda^\alpha \llbracket x, A \rrbracket \overline{P} \end{aligned} \quad (46.12)$$

**Proof of (46.12)**

$$\begin{aligned} &\alpha_g(\widehat{\mathcal{S}}^\alpha \llbracket x = A ; \rrbracket \lambda) \overline{P} \\ = &\alpha_g(\widehat{\mathcal{S}}^\alpha \llbracket x = A ; \rrbracket \lambda \overline{P}) && \text{\{ def. (45.1) of } \alpha_g \}} \\ = &\alpha_g((\ell = \text{at} \llbracket S \rrbracket \text{ ? } \overline{P} \parallel \ell = \text{after} \llbracket S \rrbracket \text{ ? } \text{assign}_\lambda^\alpha \llbracket x, A \rrbracket \overline{P} \circ \perp^\alpha)) && \text{\{ def. (46.8) of } \widehat{\mathcal{S}}^\alpha \}} \\ = &\overline{P} \sqcup_\lambda^\alpha \text{assign}_\lambda^\alpha \llbracket x, A \rrbracket \overline{P} && \text{\{ def. (45.1) of } \alpha_g \text{ and } \text{labx} \llbracket S \rrbracket = \{\text{at} \llbracket S \rrbracket, \text{after} \llbracket S \rrbracket\} \}} \\ \triangleq &\widehat{\mathcal{S}}_\lambda^{\text{gf}\alpha} \llbracket S \rrbracket \lambda \overline{P} && \text{\{ as defined in (46.12) \}} \quad \square \end{aligned}$$

By a similar proof using (45.1) and (46.9), we get (46.13).

*Flow-insensitive semantics of a dereferenced pointer assignment statement  $S ::= *p = A ;$*

$$\begin{aligned} \hat{\mathcal{S}}^{\text{gf}\alpha} &\in \lambda \in \mathbb{L}_v \rightarrow \mathcal{S} \rightarrow \mathbb{P}_\lambda^\alpha \rightarrow \mathbb{P}_\lambda^\alpha \\ \hat{\mathcal{S}}_\lambda^{\text{gf}\alpha} \llbracket S \rrbracket \bar{P} &= \bar{P} \sqcup_\lambda^\alpha \text{assign}_\lambda^\alpha \llbracket *p, A \rrbracket \bar{P} \end{aligned} \quad (46.13)$$



theorem 45.14 becomes

**Theorem (46.14)** The flow insensitive abstract semantics  $\widehat{\mathcal{S}}_{\lambda}^{\text{gr}} \llbracket S \rrbracket$  on a well-defined abstract domain of definition 21.1 is well-defined.

## Cartesian abstract interpreters for the pointer language, section 46.5

## Cartesian abstract domain for the pointer language

- Most existing points-to static analyzers are cartesian.
- Cartesian points-to analyzers can express  $x$  and  $y$  may point only to the locations of  $u$  or  $v$  but not that  $x$  and  $y$  may only point to the same location either  $u$  or  $v$ .
- For such cartesian analyzers, chapter 28, “Abstract Cartesian Semantics” is most relevant and the cartesian semantics can be easily extended to handle the pointer statements of section 46.1.

# Cartesian abstraction for the pointer language

- We have  $\mathbb{M}_\lambda \triangleq \mathbb{L}_\lambda \rightarrow \mathbb{V}_\lambda$ .
- Assume we abstract value properties by

$$\langle \wp(\mathbb{V}_\lambda), \subseteq \rangle \xrightleftharpoons[\alpha_\lambda^\square]{\gamma_\lambda^\square} \langle \mathbb{P}_\lambda^\square, \sqsubseteq_\lambda^\square \rangle. \quad (46.16)$$

- Then the cartesian abstraction is

$$\langle \wp(\mathbb{L}_\lambda \rightarrow \mathbb{V}_\lambda), \subseteq \rangle \xrightleftharpoons[\dot{\alpha}_\lambda^\times]{\dot{\gamma}_\lambda^\times} \langle \mathbb{L}_\lambda \rightarrow \wp(\mathbb{V}_\lambda), \subseteq \rangle \xrightleftharpoons[\dot{\alpha}_\lambda^\square]{\dot{\gamma}_\lambda^\square} \langle \mathbb{L}_\lambda \rightarrow \mathbb{P}_\lambda^\square, \sqsubseteq_\lambda^\square \rangle$$

where

$$\dot{\alpha}_\lambda^\times(P) \triangleq l \in \mathbb{L}_\lambda \mapsto \{\mu(l) \mid \mu \in P\} \quad (46.17)$$

is the cartesian abstraction and  $\dot{\alpha}_\lambda^\square(P) \triangleq l \in \mathbb{L}_\lambda \mapsto \alpha_\lambda^\square(P(l))$  is the pointwise extension of  $\alpha_\lambda^\square$ .

- By composition of these Galois connections, we get

$$\langle \wp(\mathbb{L}_\lambda \rightarrow \mathbb{V}_\lambda), \subseteq \rangle \xrightleftharpoons[\dot{\alpha}_\lambda^{\times\square}]{\dot{\gamma}_\lambda^{\times\square}} \langle \dot{\mathbb{P}}_\lambda^\square, \sqsubseteq^\square \rangle \quad \text{where} \quad \dot{\mathbb{P}}_\lambda^\square \triangleq \mathbb{L}_\lambda \rightarrow \mathbb{P}_\lambda^\square$$

with  $\dot{\alpha}_\lambda^{\times\square}(P) \triangleq \dot{\alpha}_\lambda^\square \circ \dot{\alpha}_\lambda^\times(P) = l \in \mathbb{L}_\lambda \mapsto \alpha_\lambda^\square(\{\mu(l) \mid \mu \in P\})$ .

## Cartesian abstract domain

- The cartesian abstract domain

$$\mathbb{D}^{\alpha} \triangleq \langle \mathbb{P}^{\alpha}, \sqsubseteq^{\alpha}, \perp^{\alpha}, \top^{\alpha}, \sqcup^{\alpha}, \sqcap^{\alpha}, 1^{\alpha}, \theta^{\alpha}, \theta^{\alpha_1}, \Theta^{\alpha}, \overline{\Theta}^{\alpha} \rangle \quad (28.42)$$

is extended to include abstract pointer operations

$$\mathbb{D}_{\lambda}^{\alpha} \triangleq \langle \mathbb{P}_{\lambda}^{\alpha}, \sqsubseteq_{\lambda}^{\alpha}, \perp_{\lambda}^{\alpha}, \top_{\lambda}^{\alpha}, \sqcup_{\lambda}^{\alpha}, \sqcap_{\lambda}^{\alpha}, 1_{\lambda}^{\alpha}, \theta_{\lambda}^{\alpha}, \theta_{\lambda}^{\alpha_1}, \star_{\lambda}^{\alpha}, \Theta^{\alpha}, \overline{\Theta}^{\alpha}, \text{NULL}_{\lambda}^{\alpha}, \lambda_{\lambda}^{\alpha}, \star_{\lambda}^{\alpha} \rangle \quad (46.18)$$

with well-definedness conditions  $\text{NULL}_{\lambda}^{\alpha} \in \mathbb{P}_{\lambda}^{\alpha}$ ,  $\lambda^{\alpha}, \star^{\alpha} \in \mathbb{L}_V \rightarrow V \rightarrow \mathbb{P}_{\lambda}^{\alpha}$ .

- We must then calculate structural definitions of  $\mathcal{A}_{\lambda}^{\alpha}[[A]] \overline{P}$ ,  $\mathcal{A}_{\lambda}^{\alpha_1}[[1]] \chi \overline{P}$ ,  $\text{assign}_{\lambda}^{\alpha}[[x, A]]$ ,  $\text{assign}_{\lambda}^{\alpha}[[\star p, A]]$ ,  $\text{test}_{\lambda}^{\alpha}[[B]]$ , and  $\overline{\text{test}}_{\lambda}^{\alpha}[[B]]$  to extend (28.17), (28.30), and (28.38) for the pointer language of section 46.1.

## Cartesian forward reachability of an arithmetic or pointer expressions

- The cartesian reachability abstraction of expressions is

$$\mathcal{A}_\lambda^\alpha[A] \in (\mathbb{L}_\lambda \rightarrow \mathbb{P}_\lambda^\alpha) \rightarrow \mathbb{P}_\lambda^\alpha$$

such that

$$\mathcal{A}_\lambda^\alpha[A] \bar{P} \dot{\sqsubseteq}_\lambda^\alpha \alpha_\lambda^\alpha(\{\mathcal{A}_\lambda[A] \mu \mid \mu \in \gamma_\lambda^{\times\alpha}(\bar{P})\})$$

- To define  $\mathcal{A}_\lambda^\alpha[A]$  structurally, we extend (28.17) as follows (in case  $\gamma_\lambda^\alpha(\perp^\alpha) = \emptyset$ , the smash of the cartesian property  $\bigoplus_\lambda^\alpha$  ensures strictness by a test for the infimum  $\perp_\lambda^\alpha$ )

# Cartesian abstract forward reachability semantics of an arithmetic or pointer expression A

The abstract cartesian semantics of an expression  $A$  and an assignment statement  $x = A$  ; is defined as follows ( $\overline{Q} \in \mathbb{P}_\lambda^\alpha, \overline{P} \in \mathbb{L}_\lambda \rightarrow \mathbb{P}_\lambda^\alpha$ )

$$\begin{aligned}
 \bigodot_\lambda^\alpha(\overline{Q}) \overline{P} &\triangleq \overline{Q} && \text{if } \gamma_\lambda^\alpha(\perp_\lambda^\alpha) \neq \emptyset \quad \text{smash} \\
 &\triangleq (\overline{P} = \perp_\lambda^\alpha ? \perp_\lambda^\alpha : \overline{Q}) && \text{if } \gamma_\lambda^\alpha(\perp_\lambda^\alpha) = \emptyset \\
 \mathcal{A}^\alpha &\in \lambda \in \mathcal{A} \rightarrow \mathbb{L}_\lambda \rightarrow \dot{\mathbb{P}}^\alpha \rightarrow \mathbb{V}_\lambda \quad \text{where} \quad \dot{\mathbb{P}}^\alpha \triangleq \mathbb{L}_\lambda \rightarrow \mathbb{P}_\lambda^\alpha && (46.19) \\
 \mathcal{A}_\lambda^\alpha[1] \overline{P} &\triangleq \bigodot_\lambda^\alpha(1_\lambda^\alpha) \overline{P} && 1_\lambda^\alpha \triangleq_\lambda^\alpha \alpha_\lambda^\alpha(\{1\}) \\
 \mathcal{A}_\lambda^\alpha[x] \overline{P} &\triangleq \overline{P}(\lambda_x) \\
 \mathcal{A}_\lambda^\alpha[A_1 - A_2] \overline{P} &\triangleq \bigodot_\lambda^\alpha(\mathcal{A}_\lambda^\alpha[A_1] \overline{P} \ominus_\lambda^\alpha \mathcal{A}_\lambda^\alpha[A_2] \overline{P}) \overline{P} \\
 &\text{where } \overline{P} \ominus_\lambda^\alpha \overline{Q} \triangleq_\lambda^\alpha \alpha_\lambda^\alpha(\{x - y \mid x \in \gamma_\lambda^\alpha(\overline{P}) \cap \mathbb{Z} \wedge y \in \gamma_\lambda^\alpha(\overline{Q}) \cap \mathbb{Z}\})
 \end{aligned}$$

$$\begin{array}{lll}
\mathcal{A}_\lambda^\alpha[\text{NULL}] \bar{P} & \triangleq & \bigoplus_\lambda^\alpha(\text{NULL}_\lambda^\alpha) \bar{P} & \text{NULL}_\lambda^\alpha \dot{\triangleq}_\lambda^\alpha \alpha_\lambda^\alpha(\{\text{nil}\}) \\
\mathcal{A}_\lambda^\alpha[\&x] \bar{P} & \triangleq & \bigoplus_\lambda^\alpha(\lambda_\lambda^\alpha[x]) \bar{P} & \lambda_\lambda^\alpha[x] \dot{\triangleq}_\lambda^\alpha \alpha_\lambda^\alpha(\{\lambda_x\}) \\
\mathcal{A}_\lambda^\alpha[*p] \bar{P} & \triangleq & \bigoplus_\lambda^\alpha(*_\lambda^\alpha[p] \bar{P}) \bar{P} & \\
\end{array}$$

where  $*_\lambda^\alpha[p] \bar{P} \dot{\triangleq}_\lambda^\alpha \alpha_\lambda^\alpha(\{\mu(\mu(\lambda_p)) \mid \mu \in \dot{\gamma}_\lambda^{\times\alpha}(\bar{P}) \wedge \mu(\lambda_p) \in \mathbb{L}\})$



## Remark 46.20

- By choosing  $\alpha_\lambda^\square$  to be the identity, we get the definition of cartesian reachability  $\mathcal{A}_\lambda^\times[A] \bar{P} \triangleq \{\mathcal{A}_\lambda[A] \mu \mid \mu \in \dot{\gamma}_\lambda^\times(\bar{P})\} \in (\mathbb{L}_\lambda \rightarrow \wp(\mathbb{V}_\lambda) \rightarrow \wp(\mathbb{V}_\lambda))$  as an instance of  $\mathcal{A}_\lambda^\square[A] \bar{P}$  with
  - $\dot{P}_\lambda^\square \triangleq \wp(\mathbb{L}_\lambda \rightarrow \mathbb{V}_\lambda)$ ,
  - $\dot{I}_\lambda^\times \triangleq \emptyset$ ,
  - $\dot{\perp}_\lambda^\times \triangleq \emptyset$ ,
  - $1_\lambda^\times \triangleq \{1\}$ ,
  - $P \ominus_\lambda^\times Q = \{x - y \mid x \in P \cap \mathbb{Z} \wedge y \in Q \cap \mathbb{Z}\}$ ,
  - $\text{NULL}_\lambda^\times \triangleq \{\text{nil}\}$ ,
  - $\lambda_\lambda^\times[x] \triangleq \{\lambda_x\}$ , and
  - $\star_\lambda^\times[p] P \triangleq \{\mu(\mu(\lambda_p)) \mid \mu \in P \wedge \mu(\lambda_p) \in \mathbb{L}\}$ .

Then lemma 28.14 remains valid, as follows

**Lemma (46.21)**

$$\forall P \in \wp(\mathbb{L}_\lambda \rightarrow \mathbb{V}_\lambda) . \{ \mathcal{A}_\lambda[A] \mu \mid \mu \in P \wedge \mathcal{A}_\lambda[A] \mu \neq \Omega \} \subseteq \mathcal{A}_\lambda^\times[A](\dot{\alpha}_\lambda^\times(P)).$$

*TODO: A revoir*

## Proof of lemma 46.21

$$\begin{aligned}
 & \text{— } \{ \mathcal{A}[\text{NULL}] \lambda \mu \mid \mu \in P \wedge \mathcal{A}[\text{NULL}] \lambda \mu \neq \Omega \} \\
 &= \{ P = \emptyset \text{ ? } \emptyset : \{\text{nil}\} \} \quad \quad \quad \{ \text{def. (46.2) of } \mathcal{A}[\text{NULL}] \lambda \mu \triangleq \text{nil} \neq \Omega \} \\
 &= \{ \dot{\alpha}_\times(P) = \dot{\emptyset} \text{ ? } \emptyset : \{\text{nil}\} \} \quad \quad \quad \{ \text{def. } \dot{\alpha}_\times(P) \triangleq \lambda \mapsto \{ \mu(\lambda) \mid \mu \in P \} \text{ so } P = \emptyset \Leftrightarrow \dot{\alpha}_\times(P) = \dot{\emptyset} \} \\
 &= {}^\times(\text{NULL}^\times) \dot{\alpha}_\times(P) \quad \quad \quad \{ \text{def. } {}^\times \text{ and } \text{NULL}^\times \triangleq \{\text{nil}\} \} \\
 &= \mathcal{A}_\lambda^\square[\text{NULL}] \dot{\alpha}_\times(P) \quad \quad \quad \{ \text{def. (46.19) of } \mathcal{A}_\lambda^\square \}
 \end{aligned}$$

$$\begin{aligned}
 & \text{— } \{ \mathcal{A}[\text{*p}] \lambda \mu \mid \mu \in P \wedge \mathcal{A}[\text{*p}] \lambda \mu \neq \Omega \} \\
 &= \{ \{ \mu(\lambda_p) \in \mathbb{Z} \cup \{\text{nil}\} \text{ ? } \Omega : \mu(\mu(\lambda_p)) \} \mid \mu \in P \wedge \mu(\lambda_p) \notin \mathbb{Z} \cup \{\text{nil}\} \} \\
 & \quad \quad \quad \{ \text{def. (46.2) of } \mathcal{A}[\text{*p}] \lambda \mu \} \\
 &= \{ \mu(\mu(\lambda_p)) \mid \mu \in P \wedge \mu(\lambda_p) \in \mathbb{L} \} \quad \quad \quad \{ \text{def. values } \mathbb{V} \triangleq \mathbb{Z} \cup \mathbb{L} \cup \{\text{nil}\} \} \\
 &\subseteq \{ \mu(\mu(\lambda_p)) \mid \mu \in \dot{\gamma}_\times(\dot{\alpha}_\times(P)) \wedge \mu(\lambda_p) \in \mathbb{L} \}
 \end{aligned}$$

{cartesian abstraction Galois connection  $\langle \wp(\mathbb{L} \rightarrow \mathbb{V}), \subseteq \rangle \xrightleftharpoons[\alpha_x]{\gamma_x} \langle \mathbb{L} \rightarrow \wp(\mathbb{V}), \subseteq \rangle$  and exercise 11.35.3}

$$= \llbracket \dot{\alpha}_x(P) = \dot{\emptyset} \text{ ? } \emptyset : \{\mu(\mu(\lambda_p)) \mid \mu \in \dot{\gamma}_x(\dot{\alpha}_x(P)) \wedge \mu(\lambda_p) \in \mathbb{L}\} \rrbracket \quad \{\dot{\gamma}_x(\dot{\alpha}_x(\dot{\emptyset})) = \emptyset\}$$

$$= {}^\times(\star_\lambda^\times \llbracket p \rrbracket \dot{\alpha}_x(P)) \dot{\alpha}_x(P) \quad \{\text{def. } {}^\times \text{ and } \star_\lambda^\times \llbracket p \rrbracket\}$$

$$= \mathcal{A}_\lambda^\times \llbracket A \rrbracket (\dot{\alpha}_x(P)) \quad \{\text{def. (46.19) of } \mathcal{A}_\lambda^\times \llbracket \star p \rrbracket \bar{P} \triangleq {}^\times(\star_\lambda^\times \llbracket p \rrbracket \bar{P}) \bar{P}\}$$

The proof of the other cases is similar. □

# Cartesian forward reachability of a pointer assignment

*Cartesian abstract forward reachability semantics of assignment  $x = A$  ;*

$$\text{assign}_\lambda^\alpha[[x, A]] \bar{P} \triangleq \bar{P}[\lambda_x \leftarrow \mathcal{A}_\lambda^\alpha[[A]] \bar{P}] \quad (46.22)$$

*Cartesian abstract forward reachability semantics of assignment  $*p = A$  ;*

$$\begin{aligned} \text{assign}_\lambda^\alpha[[*p, A]] \bar{P} &\triangleq \lambda_y \mapsto \bar{P}(\lambda_y) \sqcup^\alpha ( \lambda_\lambda^\alpha[[y]] \sqsubseteq^\alpha \bar{P}(\lambda_p) ? \mathcal{A}^\alpha[[A]] \bar{P} : \perp^\alpha ) \\ &\triangleq \lambda_y \mapsto \bar{P}(\lambda_y) \sqcup^\alpha \mathcal{A}^\alpha[[A]] \bar{P} \end{aligned}$$

if  $\alpha_\lambda^\alpha(\{\lambda_x\}) = \lambda_\lambda^\alpha[[x]]$   
otherwise

The cartesian assignment theorem 28.15 becomes (by calculational design)

**Theorem (cartesian assignment for the pointer language)**

$$\begin{aligned}\dot{\alpha}_{\lambda}^{\times} \circ \text{assign}_{\lambda}^{\vec{r}}[\![x, A]\!] &\subseteq \text{assign}_{\lambda}^{\times}[\![x, A]\!] \circ \dot{\alpha}_{\lambda}^{\times} \\ \dot{\alpha}_{\lambda}^{\times} \circ \text{assign}_{\lambda}^{\vec{r}}[\![*p, A]\!] &\subseteq \text{assign}_{\lambda}^{\times}[\![*p, A]\!] \circ \dot{\alpha}_{\lambda}^{\times}\end{aligned}$$

theorem ?? becomes

**Theorem (46.24, Abstract cartesian assignment)**

If  $1^\times \sqsubseteq^\times \gamma_\lambda^\alpha(1^\alpha)$ ,  $\Theta^\times$  is  $\sqsubseteq^\times$ -argumentwise increasing, and  $\forall X, Y \in \mathbb{P}_\lambda^\alpha . \gamma_\lambda^\alpha(X) \Theta^\times \gamma_\lambda^\alpha(Y) \sqsubseteq^\times \gamma_\lambda^\alpha(X \Theta^\alpha Y)$  then

$$\begin{aligned} \mathcal{A}_\lambda^\times[A] \circ \dot{\gamma}_\alpha &\sqsubseteq^\times \gamma_\lambda^\alpha \circ \mathcal{A}_\lambda^\alpha[A] \\ \text{assign}_\lambda^\times[x, A] \circ \dot{\gamma}_\alpha &\sqsubseteq^\times \dot{\gamma}_\alpha \circ \text{assign}_\lambda^\alpha[x, A] \\ \text{assign}_\lambda^\times[\star p, A] \circ \dot{\gamma}_\alpha &\sqsubseteq^\times \dot{\gamma}_\alpha \circ \text{assign}_\lambda^\alpha[\star p, A] \end{aligned}$$



## Cartesian backward accessibility of an arithmetic or pointer expressions

The cartesian backward accessibility abstraction of expressions is

$$\mathcal{A}_\lambda^{\text{B}} \llbracket A \rrbracket \in \mathbb{P}_\lambda^{\text{B}} \rightarrow (\mathbb{L}_\lambda \rightarrow \mathbb{P}_\lambda^{\text{B}}) \rightarrow (\mathbb{L}_\lambda \rightarrow \mathbb{P}_\lambda^{\text{B}}),$$

$$\mathcal{A}_\lambda^{\text{B}} \llbracket A \rrbracket \chi P \triangleq \alpha_\lambda^{\text{B}}(\{\mu \in \gamma_\lambda^{\text{B}}(\overline{P}) \mid \mathcal{A}_\lambda \llbracket A \rrbracket \mu \in \gamma_\lambda^{\text{B}}(\chi)\}).$$

We extend (28.30)

*Cartesian backward accessibility semantics of an arithmetic or pointer expression A*

$$\begin{aligned}
 \mathcal{A}_\lambda^{\mathfrak{A}_1} \llbracket A \rrbracket &\in \mathbb{P}_\lambda^{\mathfrak{A}} \rightarrow (\mathbb{L}_\lambda \rightarrow \mathbb{P}_\lambda^{\mathfrak{A}}) \rightarrow (\mathbb{L}_\lambda \rightarrow \mathbb{P}_\lambda^{\mathfrak{A}}) & (46.25) \\
 \mathcal{A}_\lambda^{\mathfrak{A}_1} \llbracket 1 \rrbracket \chi \bar{P} &\triangleq (1^{\mathfrak{A}} \sqcap^{\mathfrak{A}} \chi \neq \perp^{\mathfrak{A}} ? \bar{P} : \perp^{\mathfrak{A}}) \\
 \mathcal{A}_\lambda^{\mathfrak{A}_1} \llbracket \lambda_x \rrbracket \chi \bar{P} &\triangleq (\bar{P}(\lambda_x) \sqcap^{\mathfrak{A}} \chi \neq \perp^{\mathfrak{A}} ? \bar{P}[\lambda_x \leftarrow \bar{P}(\lambda_x) \sqcap^{\mathfrak{A}} \chi] : \perp^{\mathfrak{A}}) \\
 \mathcal{A}_\lambda^{\mathfrak{A}_1} \llbracket A_1 - A_2 \rrbracket \chi \bar{P} &\triangleq \text{let } \langle \chi_1, \chi_2 \rangle = \Theta_\lambda^{\mathfrak{A}_1} \chi \langle \mathcal{A}_\lambda^{\mathfrak{A}_1} \llbracket A_1 \rrbracket \bar{P}, \mathcal{A}_\lambda^{\mathfrak{A}_1} \llbracket A_2 \rrbracket \bar{P} \rangle \text{ in} \\
 &\quad \mathcal{A}_\lambda^{\mathfrak{A}_1} \llbracket A_1 \rrbracket \chi_1 \bar{P} \dot{\sqcap}^{\mathfrak{A}} \mathcal{A}_\lambda^{\mathfrak{A}_1} \llbracket A_2 \rrbracket \chi_2 \bar{P} \\
 \text{where } \Theta_\lambda^{\mathfrak{A}_1} P \langle Q_1, Q_2 \rangle &\triangleq_\lambda^{\mathfrak{A}} \langle \alpha_\lambda^{\mathfrak{A}}(\{v_1 \in \gamma_\lambda^{\mathfrak{A}}(Q_1) \mid \exists v_2 \in \gamma_\lambda^{\mathfrak{A}}(Q_2) . (v_1 - v_2) \in \gamma_\lambda^{\mathfrak{A}}(P)\}), \\
 &\quad \alpha_\lambda^{\mathfrak{A}}(\{v_2 \in \gamma_\lambda^{\mathfrak{A}}(Q_2) \mid \exists v_1 \in \gamma_\lambda^{\mathfrak{A}}(Q_1) . (v_1 - v_2) \in \gamma_\lambda^{\mathfrak{A}}(P)\}) \rangle
 \end{aligned}$$

$$\mathcal{A}_\lambda^{\boxplus_1} \llbracket \text{NULL} \rrbracket \chi \bar{P} \triangleq (\text{NULL}_\lambda^{\boxplus} \sqcap^{\boxplus} \chi \neq \perp^{\boxplus} \text{? } \bar{P} \text{: } \perp^{\boxplus})$$

$$\mathcal{A}_\lambda^{\boxplus_1} \llbracket \&x \rrbracket \chi \bar{P} \triangleq (\lambda_\lambda^{\boxplus} \llbracket \lambda_x \rrbracket \sqcap^{\boxplus} \chi \neq \perp^{\boxplus} \text{? } \bar{P}[\lambda_x \leftarrow \lambda_\lambda^{\boxplus} \llbracket x \rrbracket \sqcap^{\boxplus} \chi] \text{: } \perp^{\boxplus})$$

$$\mathcal{A}_\lambda^{\boxplus_1} \llbracket *p \rrbracket \chi \bar{P} \triangleq \star^{\boxplus_1} \llbracket p \rrbracket \lambda \chi \bar{P}$$

$$\text{where } \star^{\boxplus_1} \llbracket p \rrbracket \lambda \chi \bar{P} \triangleq_{\lambda} \lambda_z \mapsto \alpha_\lambda^{\boxplus}(\{\mu(\lambda_z) \in \gamma_\lambda^{\boxplus}(\bar{P}(\lambda_z)) \mid \mu(\mu(\lambda_p)) \in \chi \cap \gamma_\lambda^{\boxplus}(\bar{P}(\mu(\lambda_p)))\})$$

## Remark 46.26

- By choosing  $\alpha_\lambda^{\mathfrak{a}}$  to be the identity, we get the definitions of
  - $\dot{\subseteq}^\times \triangleq \dot{\subseteq}$ ,
  - $\Theta^{\mathfrak{a}_1} P \langle Q_1, Q_2 \rangle \triangleq \langle \{v_1 \in Q_1 \mid \exists v_2 \in Q_2 . (v_1 - v_2) \in P\}, \{v_2 \in Q_2 \mid \exists v_1 \in Q_1 . (v_1 - v_2) \in P\} \rangle$  as an instance of  $\Theta_\lambda^{\mathfrak{a}_1}$ , and
  - $\mathcal{A}_\lambda^{\times_1} \llbracket A \rrbracket$  as an instance of  $\mathcal{A}_\lambda^{\mathfrak{a}_1} \llbracket A \rrbracket$
 in (46.25).
- Then  $\mathcal{A}_\lambda^{\times_1} \llbracket A \rrbracket \chi$  is a lower closure operator on  $\mathcal{V} \rightarrow \wp(\mathbb{V}_\lambda)$  and  $\mathcal{A}_\lambda^{\times_1} \llbracket A \rrbracket \chi$  is  $\dot{\subseteq}$ -increasing in  $\chi$ .

theorem 28.29 becomes

**Theorem (46.27)**  $\dot{\alpha}_\lambda^\times \circ \mathcal{A}_\lambda^{\times_1} \llbracket A \rrbracket \chi \subseteq \mathcal{A}_\lambda^{\times_1} \llbracket A \rrbracket \chi \circ \dot{\alpha}_\lambda^\times.$

For the pointer semantics, theorem 28.31 becomes

**Theorem (46.29)** If  $\gamma_\lambda^\alpha \in \langle \mathbb{P}_\lambda^\alpha, \sqsubseteq^\alpha \rangle \xrightarrow{\eta} \langle \mathbb{P}^\times, \sqsubseteq^\times \rangle$  preserves finite meets, is strict ( $\gamma_\lambda^\alpha(\perp^\alpha) = \emptyset$ ),  $1 \in \gamma_\lambda^\alpha(1^\alpha)$ , the hypotheses of theorem 46.24 hold,  $(\Theta^{\times_1} \chi)$  is argumentwise increasing, and  $\Theta^{\times_1} \gamma_\lambda^\alpha(\chi) \langle \gamma_\lambda^\alpha(P_1), \gamma_\lambda^\alpha(P_2) \rangle \subseteq \dot{\gamma}_\alpha(\Theta_\lambda^{\alpha_1} \chi \langle P_1, P_2 \rangle)$  then

$$\mathcal{A}_\lambda^{\times_1} \llbracket A \rrbracket \gamma_\lambda^\alpha(\chi) \circ \dot{\gamma}_\alpha \subseteq \dot{\gamma}_\alpha \circ \mathcal{A}_\lambda^{\alpha_1} \llbracket A \rrbracket \chi.$$

## Cartesian forward reachability of tests

For tests, (28.38) is completed by pointer equality tests.

*Cartesian abstract forward reachability semantics of an arithmetic or pointer boolean expressions*

$$\begin{aligned} \text{test}^\alpha, \overline{\text{test}}^\alpha &\in \lambda \in \mathbb{L} \rightarrow \mathcal{B} \rightarrow \dot{\mathbb{P}}_\lambda^\alpha \rightarrow \dot{\mathbb{P}}_\lambda^\alpha \quad \text{where} \quad \dot{\mathbb{P}}_\lambda^\alpha \triangleq \mathbb{L}_\lambda \rightarrow \mathbb{P}_\lambda^\alpha & (46.30) \\ \text{test}_\lambda^\alpha \llbracket p == \text{NULL} \rrbracket \overline{P} &\triangleq ( \overline{P}(\lambda_p) \sqcap^\alpha \text{NULL}_\lambda^\alpha \neq \text{NULL}_\lambda^\alpha \text{ ? } \perp^\alpha : \overline{P}[\lambda_p \leftarrow \text{NULL}_\lambda^\alpha] ) \\ \overline{\text{test}}_\lambda^\alpha \llbracket p == \text{NULL} \rrbracket \overline{P} &\triangleq ( \overline{P}(\lambda_p) \sqsubseteq^\alpha \text{NULL}_\lambda^\alpha \text{ ? } \perp^\alpha : \overline{P} ) \\ \text{test}_\lambda^\alpha \llbracket x == y \rrbracket \overline{P} &\triangleq \text{let } a = \overline{P}(\lambda_x) \sqcap^\alpha \overline{P}(\lambda_y) \text{ in } ( a = \perp^\alpha \text{ ? } \perp^\alpha : \overline{P}[\lambda_x \leftarrow a][\lambda_y \leftarrow a] ) \\ \overline{\text{test}}_\lambda^\alpha \llbracket x == y \rrbracket \overline{P} &\triangleq ( \overline{P}(\lambda_x) \sqsubseteq^\alpha \text{NULL}_\lambda^\alpha \wedge \overline{P}(\lambda_y) \sqsubseteq^\alpha \text{NULL}_\lambda^\alpha \text{ ? } \perp^\alpha : \overline{P} ) \end{aligned}$$

## Remark 46.31

Following Remark 46.20, by choosing  $\alpha_\lambda^\square$  to be the identity, we get

$\text{test}^\times \in \mathbb{L}_\lambda \rightarrow \mathcal{B} \rightarrow (\mathbb{L}_\lambda \rightarrow \wp(\mathbb{V}_\lambda)) \rightarrow (\mathbb{L}_\lambda \rightarrow \wp(\mathbb{V}_\lambda))$  where  $\text{test}_\lambda^\times \llbracket \mathcal{B} \rrbracket / \overline{\text{test}_\lambda^\times \llbracket \mathcal{B} \rrbracket}$  is the instance of  $\text{test}_\lambda^\square \llbracket \mathcal{B} \rrbracket / \overline{\text{test}_\lambda^\square \llbracket \mathcal{B} \rrbracket}$  with

- $\dot{\mathbb{P}}^\square \triangleq \wp(\mathbb{L} \rightarrow \mathbb{V}_\lambda),$
- $\perp^\times \triangleq \emptyset,$
- $\dot{\perp}^\times \triangleq \emptyset,$
- $\sqcap^\square \triangleq \cap,$
- $\sqsubseteq^\square \triangleq \subseteq,$  and
- $\text{NULL}^\times \triangleq \{\text{nil}\}.$



theorem 28.35 is extended as

**Theorem (46.32, cartesian test)**

$$\dot{\alpha}_{\lambda}^{\times} \circ \text{test}_{\lambda}^{\vec{r}}[[B]] \subseteq \text{test}_{\lambda}^{\times}[[B]] \circ \dot{\alpha}_{\lambda}^{\times}$$

and

$$\dot{\alpha}_{\lambda}^{\times} \circ \overline{\text{test}_{\lambda}^{\vec{r}}[[B]]} \subseteq \overline{\text{test}_{\lambda}^{\times}[[B]]} \circ \dot{\alpha}_{\lambda}^{\times}$$

With static pointers, theorem 28.39 on abstract cartesian tests becomes

**Theorem (46.33, abstract cartesian test)** If

- $\Theta^\times$  is argumentwise increasing,
- $\Theta^\times \langle \gamma_\lambda^\alpha(X), \gamma_\lambda^\alpha(Y) \rangle \dot{\subseteq}^\times \dot{\gamma}_\alpha(\Theta^\alpha \langle X, Y \rangle)$ , and
- $\gamma_\lambda^\alpha$  preserves finite glbs

then

$$\text{test}_\lambda^\times \llbracket B \rrbracket \circ \dot{\gamma}_\alpha \dot{\subseteq}^\times \dot{\gamma}_\alpha \circ \text{test}_\lambda^\alpha \llbracket B \rrbracket$$

and

$$\overline{\text{test}}_\lambda^\times \llbracket B \rrbracket \circ \dot{\gamma}_\alpha \dot{\subseteq}^\times \dot{\gamma}_\alpha \circ \overline{\text{test}}_\lambda^\alpha \llbracket B \rrbracket$$

The soundness proof theorem 28.44 can be extended as follows.

**Theorem (46.34)      Well-definedness and soundness of the cartesian static analysis**

The cartesian abstract domain  $\mathbb{D}^\sharp$ , pointwise extension of  $\mathbb{D}^\flat$  in (46.18) for  $\mathbb{L}$  is well-defined according to definition 21.1 so the flow-sensitive abstract interpreter  $\hat{\mathcal{S}}^\sharp \llbracket S \rrbracket$  and the flow-insensitive abstract interpreter  $\hat{\mathcal{S}}_\lambda^{\text{gs}} \llbracket S \rrbracket$  are well-defined and a sound abstraction of the assertional forward reachability semantics  $\mathcal{S}^r \llbracket S \rrbracket$  of section 46.4 for any program component  $S \in \mathcal{PC}$ .