

Abstract Interpretation of Resolution-Based Semantics

Patrick Cousot, Radhia Cousot, and Roberto Giacobazzi

ENS & NYU

CNRS & ENS

Università di Verona

October 23, 2009

1

© P. Cousot, R. Cousot, and R. Giacobazzi

Objective

2

Giorgio Levi's Festschrift workshop, Pisa, Italy, October 23, 2009

© P. Cousot, R. Cousot, and R. Giacobazzi

Our objective

- To understand the work of **Giorgio Levi** on the semantics of logic programming languages for static analysis
- By reconstructing the **semantics of Resolution-based/ Logic Programming...**
 - ...by **abstract interpretations** of a concrete semantics
 - ...chosen to be a **branching-time trace-based semantics** (built from a **state transition system**)
- In passing, we get some novel semantics that tackle **impure characteristics of real implementations.**

3

© P. Cousot, R. Cousot, and R. Giacobazzi

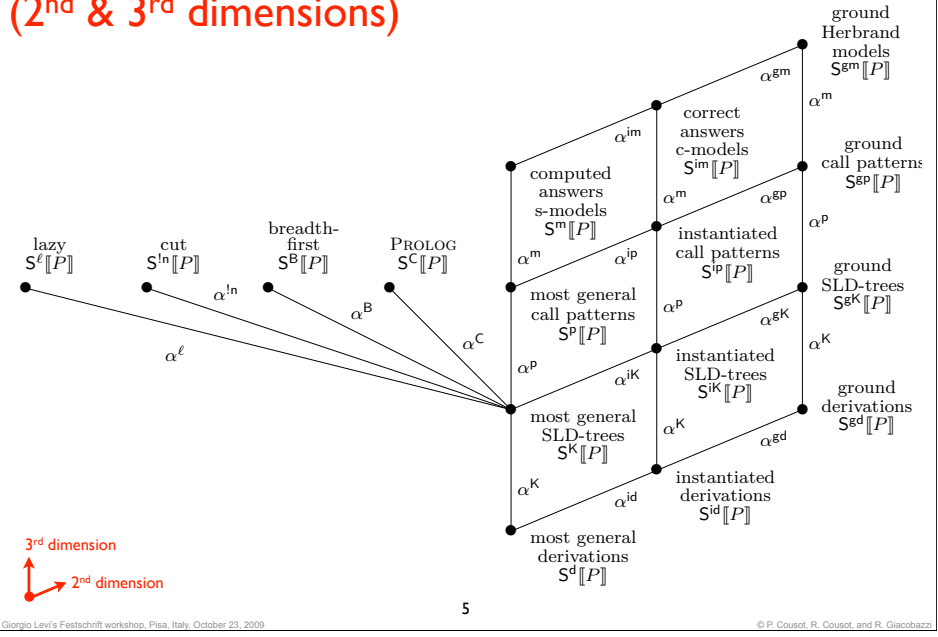
Result

4

Giorgio Levi's Festschrift workshop, Pisa, Italy, October 23, 2009

© P. Cousot, R. Cousot, and R. Giacobazzi

A Hierarchy of Abstractions and Semantics (2nd & 3rd dimensions)



Syntax

Syntax of logic programs

- $f \in \mathbb{F}$ function symbols
- $v \in \mathbb{V}$ variable symbols
- $\vec{v} = v_1, \dots, v_n$ sequences of variables ($\vec{\epsilon}$)
- $T, U, \dots \in \mathbb{T}$ terms built on \mathbb{F} and $\vec{v} \in \mathbb{V}$
- $p \in \mathbb{P}$ predicate symbols
- $A, B \in \mathbb{A}$ atoms built on \mathbb{P} and \mathbb{T}
- $B = B_1 \dots B_n \in \mathbb{B}$ sequences of atoms (\mathcal{E}), body
- $C = A \leftarrow B \in \mathbb{C}$ definite clauses (unit clauses $B = \mathcal{E}$)
- $P \in \mathbb{P}^n \triangleq [0, n[\rightarrow \mathbb{C}$ Prolog programs
- 0: $n(0) \leftarrow$
- 1: $n(s(x)) \leftarrow n(x)$
- $\alpha^l(P) \triangleq \{P_1, \dots, P_n\} \in \mathbb{L}$ abstraction to logic programs
- $\mathbb{G} \triangleq \{p(v) \mid p \in \mathbb{P} \wedge v \in \mathbb{V}\}$ most general atomic goals

Substitutions

- $\vartheta, \sigma \in \mathbb{S}$ substitutions (\mathcal{E})
- $\vartheta(T)$ application to a term T
- $\vartheta|_e$ restriction to variables of expression e
- $\vartheta \circ \sigma$ composition
- $\vartheta \preceq \sigma$ pre-order
- $\vartheta \simeq \vartheta'$ equivalence (renaming)
- $\langle \mathbb{S}^\circ / \simeq, \preceq \rangle$ complete lattice of idempotent substitutions up to renaming
- $\mathcal{T}^\circ / \simeq$ similarly for terms up to renaming

Unification

$mgu(\mathcal{T}) = \{\sigma\}$ most general unifier of a set \mathcal{T} of terms
 $\triangleq \emptyset$ not unifiable

$mgu(\mathcal{E})$ most general unifier of a set of equations $\mathcal{E} = \{T_i = U_i \mid i \in \Delta\}$

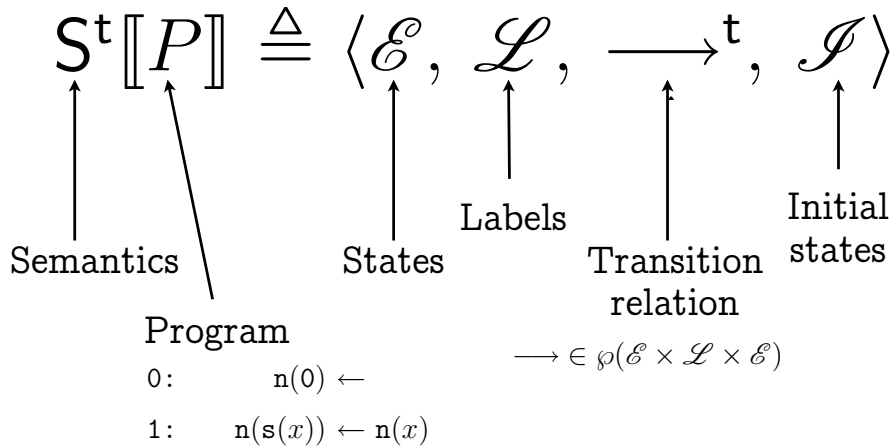
$\uparrow \in \mathbb{S}^\circ / \simeq \times \mathbb{S}^\circ / \simeq \mapsto \mathbb{S}^\circ / \simeq$ parallel composition of idempotent substitutions

9

Operational semantics defined by a labelled transition system

10

Labelled transition system



11

States

states $\eta \in \mathcal{E} \triangleq \mathcal{S} \times \mathbb{S}$

$\eta = \langle \varpi, \vartheta \rangle$

Stack Substitution

Stacks $\varpi \in \mathcal{S} \triangleq \mathcal{K}^+ \triangleq (\mathbb{C}^* \cup \mathcal{M})^+$

- $\lceil \vdash A \rceil$ Initial stack for goal A
 - $\lceil \dashv \square \rceil$ Empty stack final marker
- } markers in \mathcal{M}
- $\mathbb{C}^* \triangleq \{[i:A \leftarrow \mathbf{B}. \mathbf{B}'] \mid i:A \leftarrow \mathbf{B} \mathbf{B}' \in P\}$ specifying the control state of the derivation (\mathbf{B} has been derived while \mathbf{B}' is still to be derived) or a marker \mathcal{M}

12

Initial states

$$\mathcal{I} \triangleq \{ \langle [\vdash A], \vartheta \rangle \mid A \in \mathbb{A} \wedge \vartheta \in \mathbb{S} \}$$

goal $\vartheta(A)$ (most often ϑ is chosen as the empty substitution ε)

13

© P. Cousot, R. Cousot, and R. Giacobazzi

Transition labels

- $(\mathbf{i}:A' \leftarrow B/\sigma)$: apply renamed-apart clause $\mathbf{i}:A' \leftarrow B$ to prove goal A , such that A and A' unify by $\sigma \in \text{mgu}(\vartheta(A), A')$
- $\mathbf{i}:A \leftarrow B$: the proof of B is finished

14

© P. Cousot, R. Cousot, and R. Giacobazzi

Labelled transition relation $\xrightarrow{\ell}^t, \ell \in \mathcal{L}$

- Start from goal $\vartheta(A)$, apply clause $\mathbf{i}:A \leftarrow B$, prove new goal $\sigma \uparrow \vartheta(B)$:

$$\langle [\vdash A], \vartheta \rangle \xrightarrow{(\mathbf{i}:A' \leftarrow B/\sigma)^t} \langle [\vdash \square][\mathbf{i}:A' \leftarrow \mathbf{.}B], \vartheta' \rangle$$

if $\mathbf{i}:A' \leftarrow B \in P, \sigma \in \text{mgu}(\vartheta(A), A'), \vartheta' \in \sigma \uparrow \vartheta$ (2)

- Start from subgoal $\vartheta(B)$, apply clause $\mathbf{j}:B' \leftarrow B''$, prove new goal $\sigma \uparrow \vartheta(B'')$:

$$\langle \varpi[\mathbf{i}:A \leftarrow \mathbf{.}B.BB'], \vartheta \rangle \xrightarrow{(\mathbf{j}:B' \leftarrow B''/\sigma)^t} \langle \varpi[\mathbf{i}:A \leftarrow \mathbf{.}B.BB'][\mathbf{j}:B' \leftarrow \mathbf{.}B''], \vartheta' \rangle$$

if $\mathbf{i}:A \leftarrow \mathbf{.}B.BB', \mathbf{j}:B' \leftarrow B'' \in P, \sigma \in \text{mgu}(\vartheta(B), B'), \vartheta' \in \sigma \uparrow \vartheta$ (3)

Let $\mathbf{i}:A \leftarrow B \in P$ means that $\mathbf{i}:A \leftarrow B$ is a clause of the PROLOG program P renamed/standardized apart using fresh variables

15

© P. Cousot, R. Cousot, and R. Giacobazzi

Labelled transition relation $\xrightarrow{\ell}^t, \ell \in \mathcal{L}$

- Proof of B is finished, go back to previous goal on stack:

$$\langle \varpi[\mathbf{i}:A \leftarrow \mathbf{.}B], \vartheta \rangle \xrightarrow{\mathbf{i}:A \leftarrow B)^t} \langle \varpi, \vartheta \rangle \quad \text{if } \mathbf{i}:A \leftarrow B \in P. \quad (4)$$

16

© P. Cousot, R. Cousot, and R. Giacobazzi

© P. Cousot, R. Cousot, and R. Giacobazzi

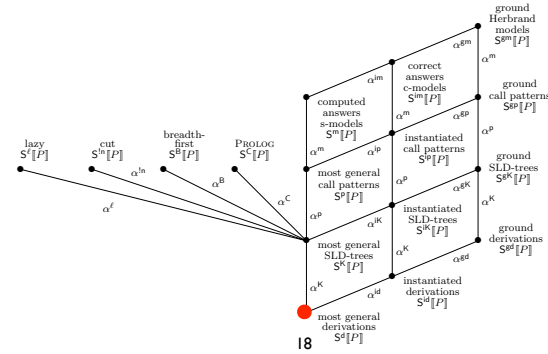
Example:

0:	$n(0) \leftarrow$
1:	$n(s(x)) \leftarrow n(x)$

$$\begin{aligned} & \langle \vdash n(s(s(0))), \varepsilon \rangle && \text{\{initial state\}} \\ \xrightarrow{\langle \vdash n(s(x)) \leftarrow n(x) / \{x \leftarrow s(0)\} \rangle_t} && \text{\{by (2)\}} \\ & \langle \vdash \neg [1 : n(s(x)) \leftarrow n(x)], \{x \leftarrow s(0)\} \rangle \\ \xrightarrow{\langle \vdash 1 : n(s(x')) \leftarrow n(x') / \{x' \leftarrow 0\} \rangle_t} && \text{\{by (3)\}} \\ & \langle \vdash \neg [1 : n(s(x)) \leftarrow n(x)] [1 : n(s(x')) \leftarrow n(x')], \{x \leftarrow s(0), x' \leftarrow 0\} \rangle \\ \xrightarrow{\langle \vdash 0 : n(0) \leftarrow \varepsilon \rangle_t} && \text{\{by (3)\}} \\ & \langle \vdash \neg [1 : n(s(x)) \leftarrow n(x)] [1 : n(s(x')) \leftarrow n(x')] [0 : n(0) \leftarrow \cdot], \\ & \quad \{x \leftarrow s(0), x' \leftarrow 0\} \rangle \\ \xrightarrow{\langle \vdash 0 : n(0) \leftarrow \cdot \rangle_t} && \text{\{by (4)\}} \\ & \langle \vdash \neg [1 : n(s(x)) \leftarrow n(x)] [1 : n(s(x')) \leftarrow n(x')], \{x \leftarrow s(0), x' \leftarrow 0\} \rangle \\ \xrightarrow{\langle \vdash 1 : n(s(x')) \leftarrow n(x') \rangle_t} && \text{\{by (4)\}} \\ & \langle \vdash \neg [1 : n(s(x)) \leftarrow n(x)], \{x \leftarrow s(0), x' \leftarrow 0\} \rangle \\ \xrightarrow{\langle \vdash 1 : n(s(x)) \leftarrow n(x) \rangle_t} && \text{\{by (4)\}} \\ & \langle \vdash \neg [1 : n(s(x)) \leftarrow n(x)], \{x \leftarrow s(0), x' \leftarrow 0\} \rangle \\ \xrightarrow{\langle \vdash 1 : n(s(x)) \leftarrow n(x) \rangle_t} && \text{\{by (4)\}} \\ & \langle \vdash \neg [1 : n(s(x)) \leftarrow n(x)], \{x \leftarrow s(0), x' \leftarrow 0\} \rangle \\ & & \square \end{aligned}$$

17

Most general maximal terminal derivation semantics of logic programs



18

Transitional Most General Maximal Derivation Semantics

- Maximal traces generated by the transition system starting from most general goals:

$$\begin{aligned} S^d[[P]] &\triangleq \{ \eta_0 \xrightarrow{\ell_0} \eta_1 \dots \eta_{n-1} \xrightarrow{\ell_{n-1}} \eta_n \in \Theta[n+1] \mid n \geq 0 \wedge \\ & \eta_0 = \langle \vdash p(v), \varepsilon \rangle \wedge p \in \mathcal{P} \wedge v \in \mathcal{V} \wedge \forall i \in [0, n-1] : \eta_i \xrightarrow{\ell_i} \eta_{i+1} \wedge \\ & \forall \eta \in \mathcal{S} : \forall \ell \in \mathcal{L} : \neg(\eta_n \xrightarrow{\ell} \eta) \} . \end{aligned}$$

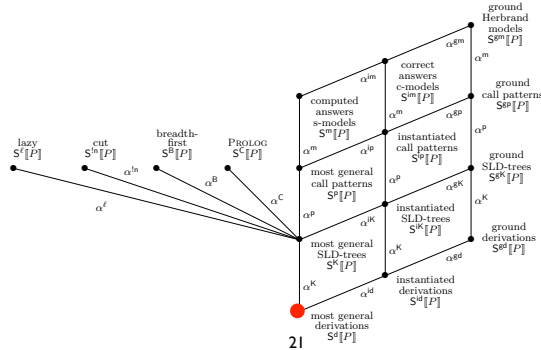
19

Final states

- *answer substitution* states in $\mathcal{E}^{AS} \triangleq \{ \langle \vdash \neg [1 : A \leftarrow B], \vartheta \rangle \mid \vartheta \in \mathcal{S} \}$ for *successful traces*, or
- *finite failure* states in $\mathcal{E}^{FF} \triangleq \{ \langle \varpi [i : A \leftarrow B.BB'], \vartheta \rangle \mid \forall j : B' \leftarrow B'' \in P : \text{mgu}(\vartheta(B), B') = \emptyset \}$ for *failing traces*.

20

Most general maximal terminal derivation semantics of logic programs in fixpoint form



21

Transitional Most General Maximal Derivation Semantics in Fixpoint Form

Theorem 20 $S^d[[P]] = lfp^{\subseteq} \hat{F}^d[[\bar{P}]]$.

$\hat{F}^d[[P]] \in \wp(\Theta) \mapsto \wp(\Theta)$

$$\hat{F}^d[[P]] \triangleq \lambda \Theta \cdot \bigcup_{i:A \leftarrow B \in P, p \in \mathbb{D}, v \in \mathbb{V}, \vartheta \in mgu(p(v), A)} \langle \{ \vdash p(v) \}, \varepsilon \rangle \xrightarrow{\{i:A \leftarrow B/\vartheta\}} \hat{F}^d[i:A \leftarrow \cdot B] \vartheta \Theta \quad (9)$$

$\hat{F}^d[i:A \leftarrow B.B'] \in \mathbb{S} \mapsto \wp(\Theta) \mapsto \wp(\Theta)$

$$\hat{F}^d[i:A \leftarrow B.BB'] \triangleq \lambda \vartheta \cdot \lambda \Theta \cdot \quad (10)$$

$$\{ \langle \{ \vdash \square [i:A \leftarrow B.BB'] \}, \{ \vdash \square [i:A \leftarrow BB.B'] \}, \vartheta \rangle \uparrow^d \eta \xrightarrow{\ell} \langle \varpi, \vartheta' \rangle \} \uparrow \theta \mid \eta \xrightarrow{\ell} \langle \varpi, \vartheta' \rangle \in \Theta . B' \wedge \sigma \in mgu(B, B') \wedge \theta \in \hat{F}^d[i:A \leftarrow BB.B'] (\vartheta \uparrow \sigma \uparrow \vartheta'^3) \Theta \}$$

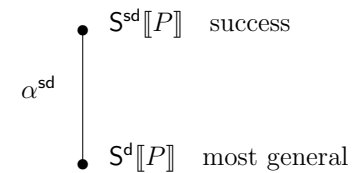
$$\hat{F}^d[i:A \leftarrow B.] \triangleq \lambda \vartheta \cdot \lambda \Theta \cdot \{ \langle \{ \vdash \square [i:A \leftarrow B.] \}, \vartheta \rangle \xrightarrow{i:A \leftarrow B} \langle \{ \vdash \square \}, \vartheta \rangle \} \quad (11)$$

22

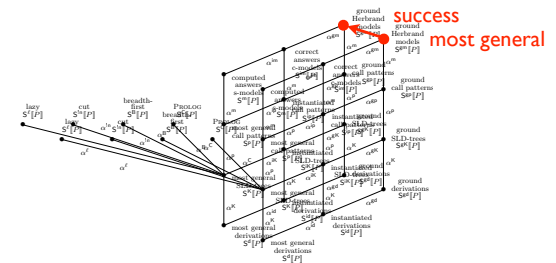
Abstractions of the trace semantics

23

1st dimension: Partial correctness Abstractions

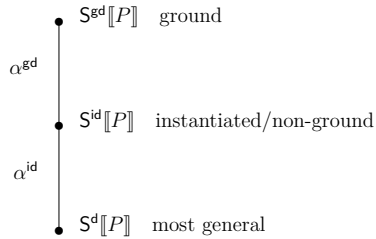


The *success abstraction* eliminates finite failures

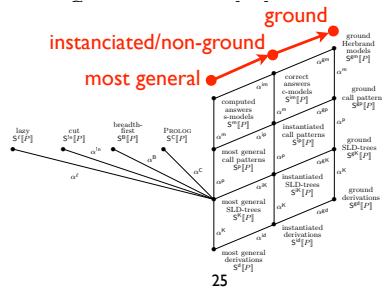


24

2nd dimension: Instantiation Abstractions



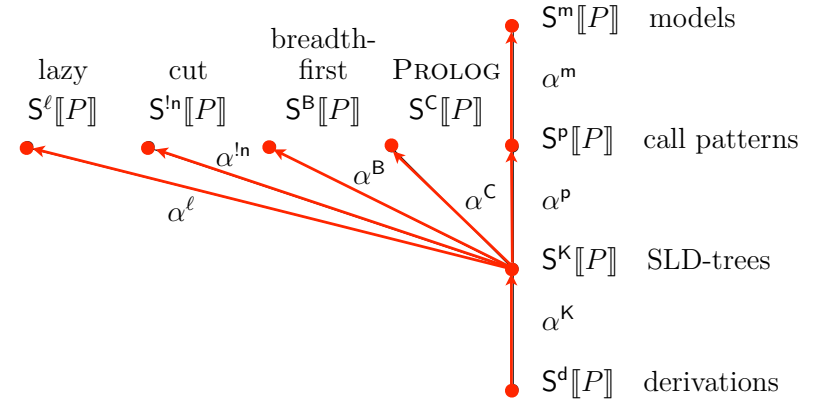
The *derivation ground instantiation abstraction* maps derivations for non-ground goals to derivations for ground instantiations of these goals.



25

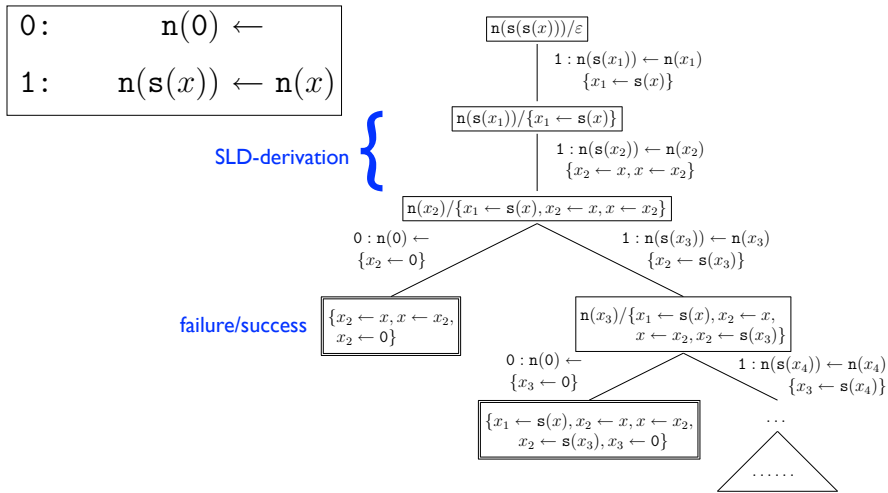
3rd dimension: Computational Information Abstractions

- Abstract away the information provided by a computation



26

SLD trees



27

SLD abstraction

- The SLD-abstraction collects the nodes of the SLD-tree from the states of traces.
- The SLD-trees are built from traces by grouping their common prefixes in the order of the PROLOG program clauses.

$$\alpha^K(\llbracket \vdash A \rrbracket, \vartheta) \triangleq \frac{\leftarrow \vartheta(A)/\vartheta}{\leftarrow \langle \alpha^K(\llbracket \varpi, \vartheta \rrbracket), \vartheta \rangle}$$

$$\alpha^K(\llbracket \varpi, \vartheta \rrbracket) \triangleq \frac{\leftarrow \langle \alpha^K(\llbracket \varpi, \vartheta \rrbracket), \vartheta \rangle}{\leftarrow \langle \alpha^K(\llbracket \varpi, \vartheta \rrbracket), \vartheta \rangle}$$

$$\alpha^K(\llbracket \varpi[i:A \leftarrow B.BB'] \rrbracket, \vartheta) \triangleq \vartheta(BB')\alpha^K(\llbracket \varpi, \vartheta \rrbracket)$$

$$\alpha^K(\llbracket \lceil \square \rrbracket, \vartheta \rrbracket) \triangleq \varepsilon$$

$$\alpha^K(\Theta) \triangleq \{ \alpha^K(\eta) \mid i_1 : \ell_1 \alpha^K(\Theta_1); \dots; i_n : \ell_n \alpha^K(\Theta_n) \mid \eta \in \mathcal{E} \wedge i_1 < \dots < i_n \wedge \Theta.\eta = \bigcup_{k=1}^n \Theta_k \wedge \forall k \in [1, n] : \Theta_k = \{ \theta \mid \eta \xrightarrow{i_k : \ell_k} \theta \in \Theta.\eta \} \neq \emptyset \}$$

$$\alpha^K(\{ \theta \mid \eta \xrightarrow{i:C} \theta \in \Theta \}) \cup \{ \langle \vartheta \rangle \mid \exists \vartheta : \langle \lceil \square \rrbracket, \vartheta \rangle \in \Theta \}$$

28

Call-patterns abstractions

- The *call-patterns abstraction* collects the goal, call-patterns and the answer substitution for each derivation, including those leading to finite failures

$$\begin{aligned} \alpha^P(\langle \xi_i, i \in \Delta \rangle) &\triangleq \bigcup \{ \alpha^P(\xi_i) \mid i \in \Delta \} && \text{SLD derivation forest} \\ \alpha^P(\overline{\leftarrow A/\sigma} \llbracket i_1 : A_1 \leftarrow B_1/\vartheta_1 \xi_1; \dots; i_n : A_n \leftarrow B_n/\vartheta_n \xi_n \rrbracket) &\triangleq && \text{SLD tree} \\ &\alpha^{FP}(\overline{\leftarrow A/\sigma} \llbracket i_1 : A_1 \leftarrow B_1/\vartheta_1 \xi_1; \dots; i_n : A_n \leftarrow B_n/\vartheta_n \xi_n \rrbracket)(\sigma(A)) \\ \alpha^{FP}(\overline{\leftarrow BB/\sigma} \llbracket i_1 : A_1 \leftarrow B_1/\vartheta_1 \xi_1; \dots; i_n : A_n \leftarrow B_n/\vartheta_n \xi_n \rrbracket)A' &\triangleq \\ &\{ \langle \sigma(A'), \sigma(B) \rangle \} \cup \bigcup_{i=1}^n \alpha^{FP}(\xi_i)(A') \\ \alpha^{FP}(\overline{\leftarrow B/\sigma} \llbracket \rrbracket)A' &\triangleq \emptyset && \text{failure} \\ \alpha^{FP}(\overline{\leftarrow \sigma} \llbracket \rrbracket)A' &\triangleq \{ \langle \sigma(A'), [\neg] \rangle \} && \text{success.} \end{aligned}$$

29

The model abstraction

- The *model abstraction* collects answers in the call patterns

$$\alpha^m(K) \triangleq \{ A \in \mathbb{A} \mid \langle A, [\neg] \rangle \in K \}$$

30

The PROLOG abstraction

- The *PROLOG abstraction* abstracts a forest $\langle \xi_i, i \in \Delta \rangle$ of SLD-trees $\xi_i, i \in \Delta$ into the set of execution traces corresponding to a depth-first traversal of these SLD-trees ξ_i (as in the PROLOG interpreter).

- SLD-trees may have infinite branches so the execution sequence, defined by transfinite recursion, may be transfinite (and is truncated to ω by PROLOG interpreters, which is a further abstraction).

$$\begin{aligned} \alpha^C(\langle \xi_i, i \in \Delta \rangle) &\triangleq \langle \alpha^C(\xi_i), i \in \Delta \rangle \\ \alpha^C(\overline{\leftarrow B/\sigma} \llbracket i_1 : A_1 \leftarrow B_1/\vartheta_1 \xi_1; \dots; i_n : A_n \leftarrow B_n/\vartheta_n \xi_n \rrbracket) &\triangleq \\ &\overline{\leftarrow B/\sigma} i_1 : A_1 \leftarrow B_1/\vartheta_1 \alpha^C(\xi_1) \dots i_n : A_n \leftarrow B_n/\vartheta_n \alpha^C(\xi_n) \\ \alpha^C(\overline{\leftarrow B/\sigma} \llbracket \rrbracket) &\triangleq \epsilon \\ \alpha^C(\overline{\leftarrow \sigma} \llbracket \rrbracket) &\triangleq \sigma \end{aligned}$$

31

Fixpoint abstract semantics

32

Abstract semantics

1. Define an **abstraction** of the trace semantics
2. Constructively derive the **abstract semantics in fixpoint form** (by proving commutation and applying the exact fixpoint transfer theorem)

33

Computational design of the abstract fixpoint semantics

- The trace semantics is in fixpoint form $S^d[[P]] = \text{lfp}^{\subseteq} \hat{F}^d[[P]]$
- So, by abstraction, the abstract fixpoint semantics also have a fixpoint definition
- **Example: Fixpoint s-semantics**

Theorem 24 (G. Levi et al.) $S^s[[P]] = \text{lfp}^{\subseteq} \hat{F}^s[[P]]$.

Let us define the *bottom-up call-patterns transformer* $\hat{F}^s[[P]] \in \wp(\mathbb{A}) \mapsto \wp(\mathbb{A})$ for a PROLOG program $P \in \mathbb{P}$ as

$$\hat{F}^s[[P]] \triangleq \lambda \mathcal{A} \cdot \bigcup_{i:A \leftarrow B \in P} \{\vartheta(A) \mid \vartheta \in \hat{F}^s[i:A \leftarrow \cdot B] \mathcal{A} \setminus \{\varepsilon\}\} \quad (12)$$

where the *clause transformer* $\hat{F}^s[i:A \leftarrow B.B'] \in \wp(\Theta) \mapsto \wp(\mathbb{S}) \mapsto \wp(\mathbb{S})$ is defined as

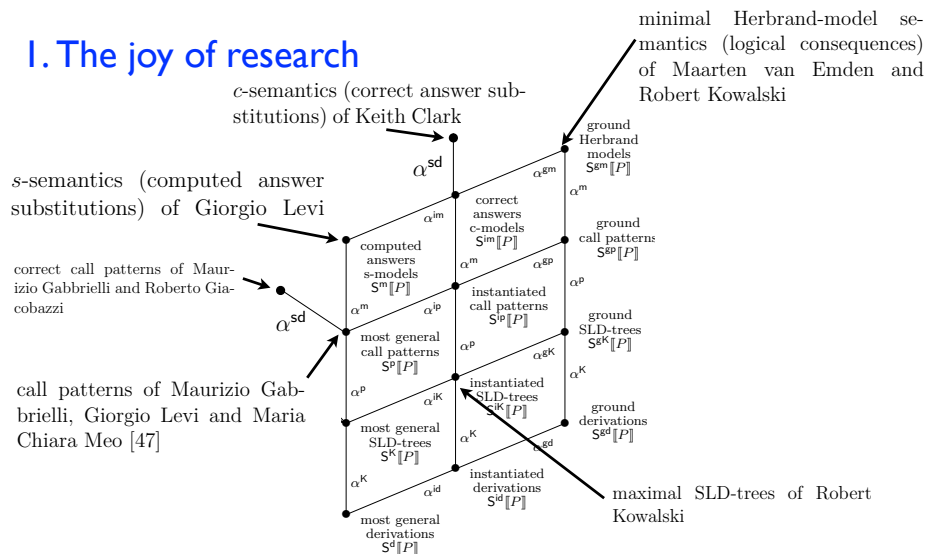
$$\hat{F}^s[i:A \leftarrow B.BB'] \triangleq \lambda \mathcal{A} \cdot \lambda \mathcal{S} \cdot \{\vartheta' \mid B' \in \mathcal{A} \wedge \sigma \in \text{mgu}(B, B') \wedge \vartheta \in \mathcal{S} \wedge \vartheta' \in \hat{F}^s[i:A \leftarrow BB.B'] \mathcal{A} (\vartheta \uparrow \sigma)\} \quad (13)$$

$$\hat{F}^s[i:A \leftarrow B] \triangleq \lambda \mathcal{A} \cdot \lambda \mathcal{S} \cdot \mathcal{S}. \quad (14)$$

34

Conclusion

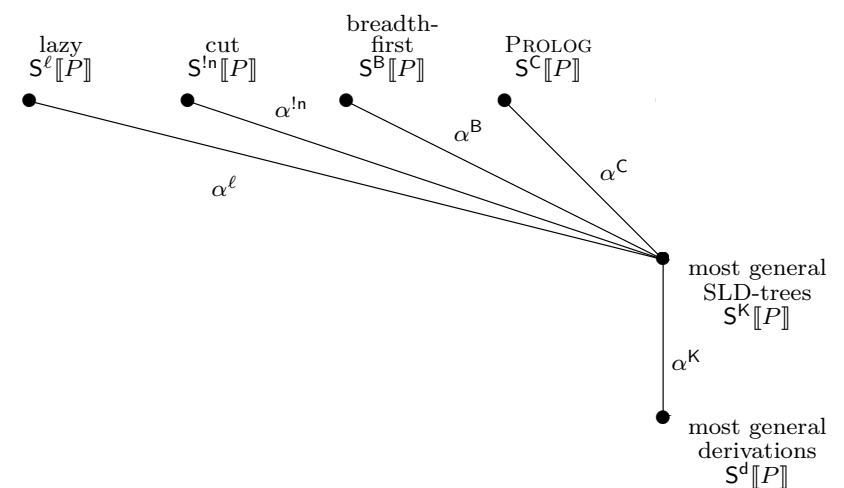
1. The joy of research



35

Conclusion (cont'd)

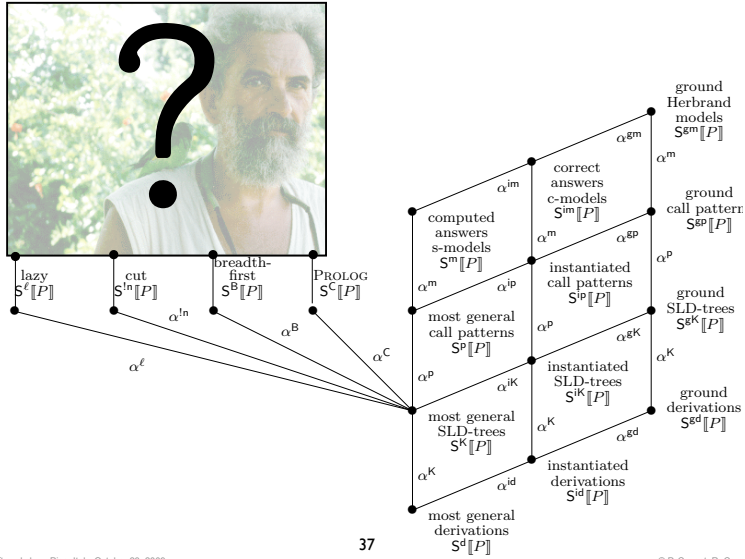
2. Life is hard!



36

Conclusion (cont'd)

3. Future work for Giorgio



37

Thank you



38