

Undecidability of Semi-unification on a Napkin

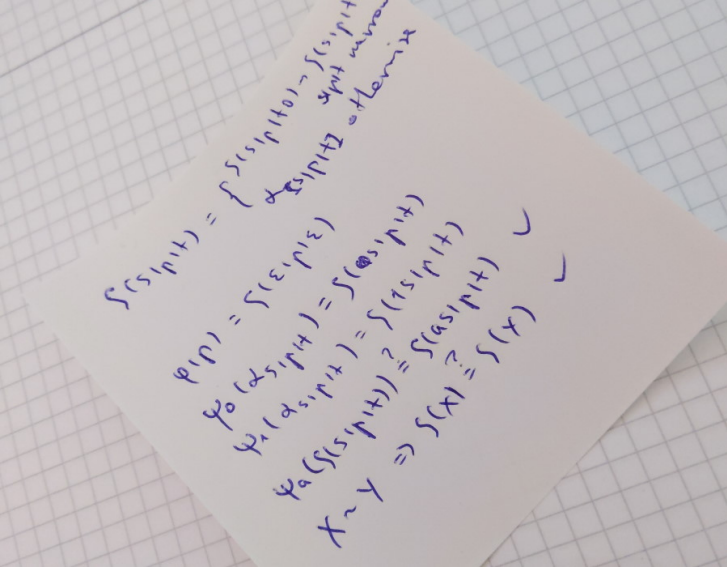
Andrej Dudenhefner

Saarland University, Saarbrücken, Germany

FSCD 2020

2020-07-02, Paris, France

Undecidability of Semi-unification on a Napkin



Semi-unification

Definition (Terms \mathbb{T})

$\mathbb{T} \ni \sigma, \tau ::= \alpha \mid \sigma \rightarrow \tau$ where α ranges over variables \mathbb{V}

- Semi-unification \sim first-order **unification** combined with **matching**

Problem (Semi-unification)

Given inequalities $\mathcal{I} = \{\sigma_1 \leq \tau_1, \dots, \sigma_n \leq \tau_n\}$,
is there a substitution $\varphi : \mathbb{V} \rightarrow \mathbb{T}$ such that
for each inequality $(\sigma \leq \tau) \in \mathcal{I}$
there is a substitution $\psi : \mathbb{V} \rightarrow \mathbb{T}$ such that
 $\psi(\varphi(\sigma)) = \varphi(\tau)$?

Theorem ([Kfoury, Tiuryn, and Urzyczyn 1993a])

Semi-unification is undecidable.

Semi-unification Occurrences

- Type inference in polymorphic functional programming
[Leiß 1989; Kfoury, Tiuryn, and Urzyczyn 1993b; Henglein 1993]
- Type inference in polymorphic logic programming
[Mycroft and O'Keefe 1984]
- System F type checking
[Wells 1999]
- Loop detection in term rewriting
[Purdom 1987]
- Program flow analysis
[Fähndrich, Rehof, and Das 2000]
- Natural language processing
[Dörre and Rounds 1990]
- ...

Semi-unification Example

Example (Composed Iteration)

- `iter2 :: Nat -> (a -> b) -> (b -> a) -> a -> a`
- `iter2 0 f g x = x`
- `iter2 1 f g x = g (f x)`
- `iter2 2 f g x = g (f (g (f x)))`
- ...

In Haskell

```
iter2 0 f g x = x
iter2 n f g x = g (iter2 (n-1) g f (f x))
```

has type

```
iter2 :: Nat -> (a -> a) -> (a -> a) -> a -> a
```

where types of `f` and `g` are **unified**.

Semi-unification Example

Example (Composed Iteration)

```
iter2 0 f g x = x
```

```
iter2 n f g x = g (iter2 (n-1) g f (f x))
```

- **Parametric polymorphism:** *monomorphic* recursive calls

\rightsquigarrow find substitution φ such that

$$\varphi(\text{Nat} \rightarrow (\alpha \rightarrow \beta) \rightarrow (\beta \rightarrow \alpha) \rightarrow \alpha \rightarrow \alpha)$$

$$= \varphi(\text{Nat} \rightarrow (\beta \rightarrow \alpha) \rightarrow (\alpha \rightarrow \beta) \rightarrow \beta \rightarrow \beta)$$

$\rightsquigarrow \varphi = \{\alpha \mapsto a, \beta \mapsto a\}$

- **Recursive polymorphism:** *instantiated* recursive calls

\rightsquigarrow find substitutions φ, ψ such that

$$\psi(\varphi(\text{Nat} \rightarrow (\alpha \rightarrow \beta) \rightarrow (\beta \rightarrow \alpha) \rightarrow \alpha \rightarrow \alpha))$$

$$= \varphi(\text{Nat} \rightarrow (\beta \rightarrow \alpha) \rightarrow (\alpha \rightarrow \beta) \rightarrow \beta \rightarrow \beta)$$

$\rightsquigarrow \varphi = \{\alpha \mapsto \alpha, \beta \mapsto \beta\}, \psi = \{\alpha \mapsto \beta, \beta \mapsto \alpha\}$

Different ψ for individual recursive calls \rightsquigarrow **semi-unification**

Semi-unification Undecidability

Original Proof Synopsis.

- Turing machine immortality** [Hooper 1966]
(is there an non-terminating configuration?)
 - \leq **Turing machine uniform boundedness**
(is the number of reachable configurations uniformly bounded?)
 - \leq **Symmetric intercell Turing machine uniform boundedness**
(as above; returning to potential past configurations)
 - \leq **Path equation derivability**
(reachability in a tailored rewriting system)
 - \leq **Unification constraint normalization**
(halting in a tailored redex contraction system)
 - \leq **Semi-unification**
- Uses excluded middle and König's lemma



Semi-unification Undecidability

New Proof Synopsis.

Turing machine immortality [Hooper 1966]

(is there an non-terminating configuration?)

≤ **Stack machine uniform boundedness**

(is the number of reachable configurations uniformly bounded?)

≤ **Semi-unification**

- First step uses fan theorem (Brouwer's intuitionism)
- Second step is fully constructive (axiom-free Coq, 1500 loc)



Simple Stack Machine

Definition (Simple Stack Machine)

Instruction: $ap \longrightarrow qb$ or $pb \longrightarrow aq$

where p, q are *states* and $a, b \in \{0, 1\}$ are *symbols*

Simple stack machine: list of instructions \mathcal{M}

Configuration: $s|p|t$

where p is a state and $s, t \in \{0, 1\}^*$ are words

Step relation:

$sa|p|t \longrightarrow_{\mathcal{M}} s|q|bt$ if $(ap \longrightarrow qb) \in \mathcal{M}$

$s|p|bt \longrightarrow_{\mathcal{M}} sa|q|t$ if $(pb \longrightarrow aq) \in \mathcal{M}$

- Simple stack machine \sim space-bounded intercell Turing machine

Problem (Uniform Boundedness)

Given a simple stack machine \mathcal{M} ,

is there an $n \in \mathbb{N}$ such that for any configuration X we have

$|\{Y \mid X \xrightarrow{*}_{\mathcal{M}} Y\}| \leq n?$

Simple Stack Machine Properties

- Mechanization-friendly
(specification 30 loc)
- No infinite tape
(linear automaton)
- Decidable reachability and termination
(every run operates in bounded space)
- Undecidable uniform boundedness
(similar to **Turing machine immortality** \leq **Turing machine uniform boundedness**)

Simple Semi-unification

Definition (Simple Constraint)

Simple constraint: $a|\alpha| \epsilon \doteq \epsilon|\beta|b$

where $\alpha \in \mathbb{V}$ and $a, b \in \{0, 1\}$

Model: $(\varphi, \psi_0, \psi_1) \models a|\alpha| \epsilon \doteq \epsilon|\beta|b$ if either

- ▶ $b = 0$ and $\psi_a(\varphi(\alpha)) \rightarrow \tau = \varphi(\beta)$ for some τ
- ▶ $b = 1$ and $\sigma \rightarrow \psi_a(\varphi(\alpha)) = \varphi(\beta)$ for some σ

Definition (Simple Semi-unification)

Given a finite set \mathcal{C} of simple constraints,

are there substitutions $\varphi, \psi_0, \psi_1 : \mathbb{V} \rightarrow \mathbb{T}$ such that for all constraints $C \in \mathcal{C}$ we have $(\varphi, \psi_0, \psi_1) \models C$?

- Undecidable fragment of semi-unification

Not Uniformly Bounded Example

Example (Not Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \longrightarrow p1\}$$

Not Uniformly Bounded Example

Example (Not Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \longrightarrow p1\}$$

$$000|p| \in$$

Not Uniformly Bounded Example

Example (Not Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \rightarrow p1\}$$

$$000|p| \epsilon \rightarrow_{\mathcal{M}} 00|p| 1$$

Not Uniformly Bounded Example

Example (Not Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \rightarrow p1\}$$

$$000|p|\epsilon \xrightarrow{\mathcal{M}} 00|p|1 \xrightarrow{\mathcal{M}} 0|p|11$$

Not Uniformly Bounded Example

Example (Not Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \rightarrow p1\}$$

$$000|p|\epsilon \xrightarrow{\mathcal{M}} 00|p|1 \xrightarrow{\mathcal{M}} 0|p|11 \xrightarrow{\mathcal{M}} \epsilon|p|111$$

Not Uniformly Bounded Example

Example (Not Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \rightarrow p1\}$$

$$000|p|\epsilon \xrightarrow{\mathcal{M}} 00|p|1 \xrightarrow{\mathcal{M}} 0|p|11 \xrightarrow{\mathcal{M}} \epsilon|p|111$$

$\rightsquigarrow 0^n|p|\epsilon$ reaches $n + 1$ distinct configurations

\rightsquigarrow no *uniform* bound on number of reachable configurations

Not Uniformly Bounded Example

Example (Not Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \longrightarrow p1\}$$

$$000|p|\epsilon \longrightarrow_{\mathcal{M}} 00|p|1 \longrightarrow_{\mathcal{M}} 0|p|11 \longrightarrow_{\mathcal{M}} \epsilon|p|111$$

$\rightsquigarrow 0^n|p|\epsilon$ reaches $n + 1$ distinct configurations

\rightsquigarrow no *uniform* bound on number of reachable configurations

- $0p \longrightarrow p1 \rightsquigarrow 0|p|\epsilon \doteq \epsilon|p|1$

Example (Unsolvable Constraints)

$$\mathcal{C} = \{0|p|\epsilon \doteq \epsilon|p|1\}$$

- $\sigma \rightarrow \psi_0(\varphi(p)) = \varphi(p)$

\rightsquigarrow no model

Uniformly Bounded Example

Example (Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \rightarrow q1, q1 \rightarrow 1p, 1p \rightarrow q0, q0 \rightarrow 0p\}$$

Uniformly Bounded Example

Example (Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{ \mathbf{0}p \longrightarrow q\mathbf{1}, q\mathbf{1} \longrightarrow \mathbf{1}p, \mathbf{1}p \longrightarrow q\mathbf{0}, q\mathbf{0} \longrightarrow \mathbf{0}p \}$$

$\mathbf{0}, p \in \epsilon$

Uniformly Bounded Example

Example (Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \rightarrow q1, q1 \rightarrow 1p, 1p \rightarrow q0, q0 \rightarrow 0p\}$$
$$0|p| \epsilon \xrightarrow{\mathcal{M}} \epsilon|q|1$$

Uniformly Bounded Example

Example (Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \rightarrow q1, q1 \rightarrow 1p, 1p \rightarrow q0, q0 \rightarrow 0p\}$$
$$0|p|\epsilon \xrightarrow{\mathcal{M}} \epsilon|q|1 \xrightarrow{\mathcal{M}} 1|p|\epsilon$$

Uniformly Bounded Example

Example (Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \rightarrow q1, q1 \rightarrow 1p, 1p \rightarrow q0, q0 \rightarrow 0p\}$$
$$0|p|\epsilon \xrightarrow{\mathcal{M}} \epsilon|q|1 \xrightarrow{\mathcal{M}} 1|p|\epsilon \xrightarrow{\mathcal{M}} \epsilon|q|0$$

Uniformly Bounded Example

Example (Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \rightarrow q1, q1 \rightarrow 1p, 1p \rightarrow q0, q0 \rightarrow 0p\}$$
$$0|p| \epsilon \xrightarrow{\mathcal{M}} \epsilon|q|1 \xrightarrow{\mathcal{M}} 1|p| \epsilon \xrightarrow{\mathcal{M}} \epsilon|q|0 \xrightarrow{\mathcal{M}} 0|p| \epsilon$$

Uniformly Bounded Example

Example (Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \rightarrow q1, q1 \rightarrow 1p, 1p \rightarrow q0, q0 \rightarrow 0p\}$$

$$0|p| \epsilon \xrightarrow{\mathcal{M}} \epsilon|q|1 \xrightarrow{\mathcal{M}} 1|p| \epsilon \xrightarrow{\mathcal{M}} \epsilon|q|0 \xrightarrow{\mathcal{M}} 0|p| \epsilon$$

\rightsquigarrow any configuration reaches at most 4 distinct configurations

Uniformly Bounded Example

Example (Uniformly Bounded Stack Machine)

$$\mathcal{M} = \{0p \rightarrow q1, q1 \rightarrow 1p, 1p \rightarrow q0, q0 \rightarrow 0p\}$$
$$0|p|\epsilon \xrightarrow{\mathcal{M}} \epsilon|q|1 \xrightarrow{\mathcal{M}} 1|p|\epsilon \xrightarrow{\mathcal{M}} \epsilon|q|0 \xrightarrow{\mathcal{M}} 0|p|\epsilon$$

\rightsquigarrow any configuration reaches at most 4 distinct configurations

Example (Solvable Constraints)

$$\mathcal{C} = \{0|p|\epsilon \doteq \epsilon|q|1, 1|p|\epsilon \doteq \epsilon|q|1, 1|p|\epsilon \doteq \epsilon|q|0, 0|p|\epsilon \doteq \epsilon|q|0\}$$

- $\sigma \rightarrow \psi_0(\varphi(p)) = \varphi(q)$
- $\sigma \rightarrow \psi_1(\varphi(p)) = \varphi(q)$
- $\psi_1(\varphi(p)) \rightarrow \tau = \varphi(q)$
- $\psi_0(\varphi(p)) \rightarrow \tau = \varphi(q)$

\rightsquigarrow model

$$\varphi(p) = \alpha$$

$$\varphi(q) = \beta \rightarrow \beta$$

$$\psi_0(\alpha) = \psi_1(\alpha) = \beta$$

Reduction Soundness

Definition (Machine Encoding)

Given simple stack machine \mathcal{M} , define

$$\mathcal{C} = \{ \mathbf{a|p|e} \doteq \mathbf{e|q|b} \mid (\mathbf{ap} \longrightarrow \mathbf{qb}) \in \mathcal{M} \text{ or } (\mathbf{qb} \longrightarrow \mathbf{ap}) \in \mathcal{M} \}$$

Definition (ζ)

$$\zeta(\mathbf{s|p|t}) = \begin{cases} \zeta(\mathbf{s|p|t0}) \rightarrow \zeta(\mathbf{s|p|t1}) & \text{if } \mathbf{s|p|t} \text{ is narrow} \\ \alpha_{[\mathbf{s|p|t}]} & \text{otherwise} \end{cases}$$

where narrowness is decidable and $[\cdot]$ is a total and computable

Lemma (Reduction Soundness)

If \mathcal{M} is uniformly bounded, then $(\varphi, \psi_0, \psi_1) \models \mathcal{C}$ where

$$\varphi(\mathbf{p}) = \zeta(\mathbf{e|p|e}) \quad \psi_0(\alpha_{\mathbf{s|p|t}}) = \zeta(\mathbf{0s|p|t}) \quad \psi_1(\alpha_{\mathbf{s|p|t}}) = \zeta(\mathbf{1s|p|t})$$

Reduction Completeness

Definition (Machine Encoding)

Given simple stack machine \mathcal{M} , define

$$\mathcal{C} = \{ \mathbf{a} | \mathbf{p} | \epsilon \doteq \epsilon | \mathbf{q} | \mathbf{b} \mid (\mathbf{ap} \longrightarrow \mathbf{qb}) \in \mathcal{M} \text{ or } (\mathbf{qb} \longrightarrow \mathbf{ap}) \in \mathcal{M} \}$$

Lemma

$\mathbf{X} \longrightarrow_{\mathcal{M}}^* \mathbf{Y}$ and $(\varphi, \psi_0, \psi_1) \models \mathcal{C}$ implies $(\varphi, \psi_0, \psi_1) \models \mathbf{X} \doteq \mathbf{Y}$

Remark

Size of the syntax tree of $\varphi(\mathbf{p})$ uniformly bounds reachable configuration space from state \mathbf{p} .

Lemma (Reduction Completeness)

If $(\varphi, \psi_0, \psi_1) \models \mathcal{C}$, then \mathcal{M} is uniformly bounded.

Contribution

- Intuitionistic (in sense of Brouwer) Turing reduction from **Turing machine immortality** to **simple stack machine uniform boundedness**
- Fully constructive many-one reduction from **simple stack machine uniform boundedness** to **semi-unification**
 - ▶ simple and direct via ζ
 - ▶ mechanized (axiom-free Coq)
(specification 100 loc, argument 1400 loc)
<https://github.com/uds-psl/2020-fscd-semi-unification>

Ongoing Work

Mechanized reduction from
the **Turing machine halting problem**
to **semi-unification**

- comprehensive
(current mechanization starts with boundedness)
- many-one
(current proof requires Turing reductions)
- axiom-free
(current proof requires fan theorem)
- part of the Coq library of Undecidability Proofs
<https://github.com/uds-psl/coq-library-undecidability>

Ongoing Work

Mechanized reduction from
the **Turing machine halting problem**
to **semi-unification**

- comprehensive
(current mechanization starts with boundedness)
- many-one
(current proof requires Turing reductions)
- axiom-free
(current proof requires fan theorem)
- part of the Coq library of Undecidability Proofs
<https://github.com/uds-psl/coq-library-undecidability>

Thank You

Bibliography I

- Dörre, Jochen and William C. Rounds (1990). “On Subsumption and Semiunification in Feature Algebras.” In: *Proceedings of the Fifth Annual Symposium on Logic in Computer Science (LICS '90), Philadelphia, Pennsylvania, USA, June 4-7, 1990*. IEEE Computer Society, pp. 300–310. DOI: 10.1109/LICS.1990.113756. URL: <https://doi.org/10.1109/LICS.1990.113756>.
- Fähndrich, Manuel, Jakob Rehof, and Manuvir Das (2000). “Scalable context-sensitive flow analysis using instantiation constraints.” In: *Proceedings of the 2000 ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI), Vancouver, British Columbia, Canada, June 18-21, 2000*. Ed. by Monica S. Lam. ACM, pp. 253–263. DOI: 10.1145/349299.349332. URL: <https://doi.org/10.1145/349299.349332>.

Bibliography II

- Henglein, Fritz (1993). "Type Inference with Polymorphic Recursion." In: *ACM Trans. Program. Lang. Syst.* 15.2, pp. 253–289. DOI: 10.1145/169701.169692. URL: <https://doi.org/10.1145/169701.169692>.
- Hooper, Philip K. (1966). "The Undecidability of the Turing Machine Immortality Problem." In: *J. Symb. Log.* 31.2, pp. 219–234. DOI: 10.2307/2269811. URL: <https://doi.org/10.2307/2269811>.
- Kfoury, Assaf J., Jerzy Tiuryn, and Pawel Urzyczyn (1993a). "The Undecidability of the Semi-unification Problem." In: *Inf. Comput.* 102.1, pp. 83–101. DOI: 10.1006/inco.1993.1003. URL: <https://doi.org/10.1006/inco.1993.1003>.
- (1993b). "Type Reconstruction in the Presence of Polymorphic Recursion." In: *ACM Trans. Program. Lang. Syst.* 15.2, pp. 290–311. DOI: 10.1145/169701.169687. URL: <https://doi.org/10.1145/169701.169687>.

Bibliography III

- Leiß, Hans (1989). “Polymorphic recursion and semi-unification.” In: *International Workshop on Computer Science Logic*. Springer, pp. 211–224.
- Mycroft, Alan and Richard A. O’Keefe (1984). “A Polymorphic Type System for Prolog.” In: *Artif. Intell.* 23.3, pp. 295–307. DOI: 10.1016/0004-3702(84)90017-1. URL: [https://doi.org/10.1016/0004-3702\(84\)90017-1](https://doi.org/10.1016/0004-3702(84)90017-1).
- Purdom, Paul Walton (1987). “Detecting Looping Simplifications.” In: *Rewriting Techniques and Applications, 2nd International Conference, RTA-87, Bordeaux, France, May 25-27, 1987, Proceedings*. Ed. by Pierre Lescanne. Vol. 256. Lecture Notes in Computer Science. Springer, pp. 54–61. DOI: 10.1007/3-540-17220-3_5. URL: https://doi.org/10.1007/3-540-17220-3_5.

Bibliography IV

Wells, J. B. (1999). "Typability and Type Checking in System F are Equivalent and Undecidable." In: *Ann. Pure Appl. Log.* 98.1-3, pp. 111–156. DOI: [10.1016/S0168-0072\(98\)00047-5](https://doi.org/10.1016/S0168-0072(98)00047-5). URL: [https://doi.org/10.1016/S0168-0072\(98\)00047-5](https://doi.org/10.1016/S0168-0072(98)00047-5).

Backup Slides

Constraint-based Semi-unification

Definition (Substitution Composition)

For substitutions $\psi_0, \psi_1 : \mathbb{V} \rightarrow \mathbb{T}$ and word $\mathbf{v} \in \{0, 1\}^*$ define

$$\psi_\epsilon(\sigma) = \sigma \quad \psi_{\mathbf{v}0}(\sigma) = \psi_{\mathbf{v}}(\psi_0(\sigma)) \quad \psi_{\mathbf{v}1}(\sigma) = \psi_{\mathbf{v}}(\psi_1(\sigma))$$

Definition (Path Function)

For $\mathbf{w} \in \{0, 1\}^*$ define

$$\pi_\epsilon(\sigma) = \sigma \quad \pi_{0\mathbf{w}}(\sigma \rightarrow \tau) = \pi_{\mathbf{w}}(\sigma) \quad \pi_{1\mathbf{w}}(\sigma \rightarrow \tau) = \pi_{\mathbf{w}}(\tau)$$

Definition (Constraint)

Constraint: $\mathbf{s}|\alpha|\mathbf{t} \doteq \mathbf{v}|\beta|\mathbf{w}$

where $\alpha, \beta \in \mathbb{V}$ and $\mathbf{s}, \mathbf{t}, \mathbf{v}, \mathbf{w} \in \{0, 1\}^*$

Model: $(\varphi, \psi_0, \psi_1) \models (\mathbf{s}|\alpha|\mathbf{t} \doteq \mathbf{v}|\beta|\mathbf{w})$ if

$$\pi_{\mathbf{t}}(\psi_{\mathbf{s}}(\varphi(\alpha))) = \pi_{\mathbf{w}}(\psi_{\mathbf{v}}(\varphi(\beta)))$$

Narrow Configuration, Representative

Definition (Joinable Configurations)

Configurations \mathbf{X} , \mathbf{Y} are *joinable* in \mathcal{M} ,
if $\mathbf{X} \xrightarrow{*}_{\mathcal{M}} \mathbf{Z} \xleftarrow{*}_{\mathcal{M}} \mathbf{Y}$ for some configuration \mathbf{Z} .

Definition (Narrow Configuration)

A configuration \mathbf{X} is *narrow* in \mathcal{M} ,
if \mathbf{X} and $\mathbf{s|p|e}$ are joinable in \mathcal{M} for some state p and a word $s \in \mathbb{B}^*$.

Definition (Representative $[\mathbf{X}]_{\mathcal{M}}$)

The *representative* of \mathbf{X} in \mathcal{M} is the lexicographically smallest configuration \mathbf{Y} such that \mathbf{X} and \mathbf{Y} are joinable in \mathcal{M} .

- Joinability is decidable
- Narrowness is decidable
- Representative is computable