Higher-order model checking meets implicit automata (work in progress)

Lê Thành Dũng (Tito) Nguyễn — nltd@nguyentito.eu — Aix-Marseille Univ. (new!) with Abhishek De (Birmingham), Charles Grellois (Sheffield) & Cécilia Pradic (Swansea) Journées du GT Scalp, mardi 19 novembre 2024

Model checking (disclaimer: I don't know anything about it)

A basic setting:

• represent possible system behaviors for time $t \to +\infty$ as a language $L \subseteq \Sigma^{\omega}$ typically, $L = \mathscr{L}(\mathscr{A})$ for some automaton \mathscr{A}

Model checking (disclaimer: I don't know anything about it)

A basic setting:

- represent possible system behaviors for time t → +∞ as a language L ⊆ Σ^ω typically, L = ℒ(𝔄) for some automaton 𝔄
- represent "nice" wanted behaviors as $\mathscr{L}(\varphi)$ for some specification φ typically, φ is a formula in some temporal *logic*
 - "no bad things happen in finite time"
 - "every open is followed by a close"
 - ...

Model checking (disclaimer: I don't know anything about it)

A basic setting:

- represent possible system behaviors for time $t \to +\infty$ as a language $L \subseteq \Sigma^{\omega}$ typically, $L = \mathscr{L}(\mathscr{A})$ for some automaton \mathscr{A}
- represent "nice" wanted behaviors as $\mathscr{L}(\varphi)$ for some specification φ typically, φ is a formula in some temporal *logic*
 - "no bad things happen in finite time"
 - "every open is followed by a close"
 - ...

The model-checking problem: "are all possible behaviors nice?"

$$L \stackrel{?}{\subseteq} \mathscr{L}(\varphi)$$

System behavior = just one <u>infinite tree</u> *T*

Branching may be used for conditionals (if/then/else), uninterpreted function symbols with multiple arguments, ...

System behavior = just one <u>infinite tree</u> *T*

Branching may be used for conditionals (if/then/else), uninterpreted function symbols with multiple arguments, ...

Wanted behaviors = formula φ in monadic second-order (MSO) logic

MSO is a canonical choice from theoretical POV:

- more expressive than temporal logics
- corresponds to regular languages on finite words
- on infinite trees: corresponds to e.g. alternating parity automata

System behavior = just one <u>infinite tree</u> *T*

Branching may be used for conditionals (if/then/else), uninterpreted function symbols with multiple arguments, ...

Wanted behaviors = formula φ in monadic second-order (MSO) logic

MSO is a canonical choice from theoretical POV:

- more expressive than temporal logics
- corresponds to regular languages on finite words
- on infinite trees: corresponds to e.g. alternating parity automata

Decision problem: $T \in \mathscr{L}(\varphi)$? i.e. $T \vDash \varphi$?

System behavior = just one <u>infinite tree</u> *T*

Branching may be used for conditionals (if/then/else), uninterpreted function symbols with multiple arguments, ...

Wanted behaviors = formula φ in monadic second-order (MSO) logic

MSO is a canonical choice from theoretical POV:

- more expressive than temporal logics
- corresponds to regular languages on finite words
- on infinite trees: corresponds to e.g. alternating parity automata

Decision problem: $T \in \mathscr{L}(\varphi)$? i.e. $T \vDash \varphi$?

www system representation = finite description of tree with decidable MSO theory

Which infinite trees have a decidable MSO theory?

- Some open problems, for instance automatic structures
- A big class of trees equivalently described by
 - higher-order pushdown automata (related to Muchnik iteration)
 - tree→graph transductions + graph→tree unfolding: "Caucal hierarchy" (2002)
 - safe higher-order recursion schemes (HORS) [Knapkik, Niwiński & Urzyczyn '02]

Which infinite trees have a decidable MSO theory?

- Some open problems, for instance automatic structures
- A big class of trees equivalently described by
 - higher-order pushdown automata (related to Muchnik iteration)
 - tree→graph transductions + graph→tree unfolding: "Caucal hierarchy" (2002)
 - safe higher-order recursion schemes (HORS) [Knapkik, Niwiński & Urzyczyn '02]
- Natural extension [Ong 2006, reproved many times]: unsafe HORS = *simply typed λ-calculus* with *recursive* definitions
 ~~ verification of functional programs! (tested on subsets of OCaml)

Which infinite trees have a decidable MSO theory?

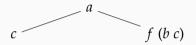
- Some open problems, for instance automatic structures
- A big class of trees equivalently described by
 - higher-order pushdown automata (related to Muchnik iteration)
 - tree→graph transductions + graph→tree unfolding: "Caucal hierarchy" (2002)
 - safe higher-order recursion schemes (HORS) [Knapkik, Niwiński & Urzyczyn '02]
- Natural extension [Ong 2006, reproved many times]: unsafe HORS = *simply typed λ-calculus* with *recursive* definitions
 ~~ verification of functional programs! (tested on subsets of OCaml)

(in practice, higher-order fixpoint logic seems to work better for model-checking functional programs: cf. recent work of Kobayashi et al.)

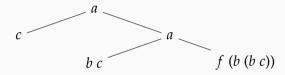
f c

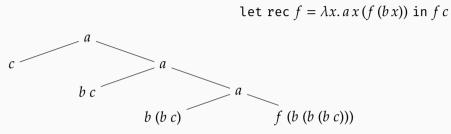
let rec $f = \lambda x. a x (f (b x))$ in f c

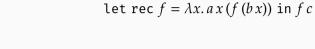
let rec $f = \lambda x. a x (f (b x))$ in f c

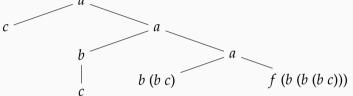


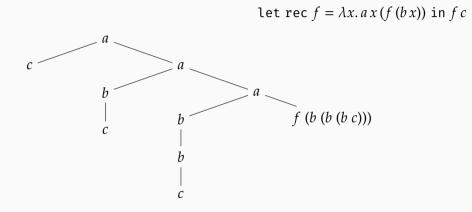
let rec $f = \lambda x. a x (f (b x))$ in f c

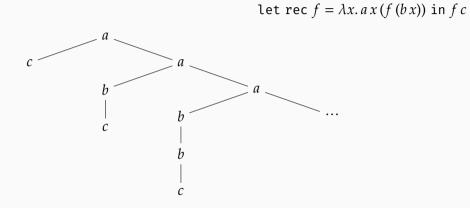












b

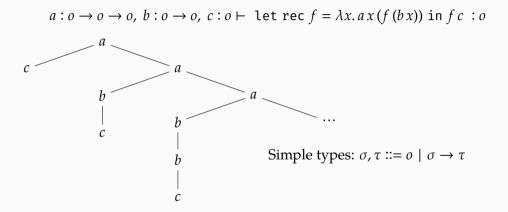
С

$$a: o \to o \to o, b: o \to o, c: o \vdash \text{ let rec } f = \lambda x. ax (f (bx)) \text{ in } f c: o$$

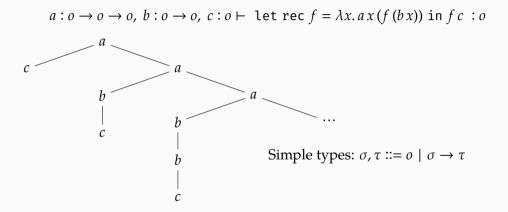
 $c \qquad a \qquad b \qquad b \qquad b \qquad b \qquad b \qquad \dots$

$$a: o \to o \to o, b: o \to o, c: o \vdash \text{ let rec } f = \lambda x. ax (f(bx)) \text{ in } fc: c$$

$$c \qquad a \qquad b \qquad b \qquad a \qquad \dots \\ c \qquad b \qquad b \qquad \dots \\ c \qquad b \qquad b \qquad \dots \\ b \qquad \text{Simple types: } \sigma, \tau ::= o \mid \sigma \to \tau$$



Decidability of MSO logic over such trees: flagship success of denotational semantics



Decidability of MSO logic over such trees: flagship success of *denotational semantics* idea best illustrated by a theorem on finite words, w/o recursion: next slide $\frac{5}{13}$

Church encodings of binary strings [Böhm & Berarducci 1985]

 \simeq fold_right on a list of characters (generalizable to any alphabet; Nat = Str_{{1}}):

 $\overline{\texttt{011}} = \lambda f_0. \ \lambda f_1. \ \lambda x. \ f_0 \ (f_1 \ (f_1 \ x)): \mathsf{Str}_{\{\texttt{0},\texttt{1}\}}[\tau] = (\tau \to \tau) \to (\tau \to \tau) \to \tau \to \tau$

Church encodings of binary strings [Böhm & Berarducci 1985]

 \simeq fold_right on a list of characters (generalizable to any alphabet; Nat = Str_{{1}}):

 $\overline{\texttt{011}} = \lambda f_0. \ \lambda f_1. \ \lambda x. \ f_0 \ (f_1 \ (f_1 \ x)) : \mathsf{Str}_{\texttt{(0,1)}}[\tau] = (\tau \to \tau) \to (\tau \to \tau) \to \tau \to \tau$

Simply typed λ -terms $t : Str_{\{0,1\}}[\tau] \to Bool define$ **languages** $<math>L \subseteq \{0, 1\}^*$

Church encodings of binary strings [Böhm & Berarducci 1985]

 \simeq fold_right on a list of characters (generalizable to any alphabet; Nat = Str_{{1}}):

 $\overline{\texttt{011}} = \lambda f_0. \ \lambda f_1. \ \lambda x. \ f_0 \ (f_1 \ (f_1 \ x)) : \mathsf{Str}_{\texttt{(0,1)}}[\tau] = (\tau \to \tau) \to (\tau \to \tau) \to \tau \to \tau$

Simply typed λ -terms $t : Str_{\{0,1\}}[\tau] \to Bool define$ **languages** $<math>L \subseteq \{0,1\}^*$ Example: $t = \lambda s. s$ id not true : $Str_{\{0,1\}}[Bool] \to Bool$ (even number of 1s)

$$t \overline{\text{Oll}} \longrightarrow_{\beta} \overline{\text{Oll}} \text{ id not true} \longrightarrow_{\beta} \text{ id (not (not true))} \longrightarrow_{\beta} \text{true}$$

Church encodings of binary strings [Böhm & Berarducci 1985]

 \simeq fold_right on a list of characters (generalizable to any alphabet; Nat = Str_{1}):

 $\overline{\texttt{011}} = \lambda f_0. \ \lambda f_1. \ \lambda x. \ f_0 \ (f_1 \ (f_1 \ x)) : \mathsf{Str}_{\texttt{(0,1)}}[\tau] = (\tau \to \tau) \to (\tau \to \tau) \to \tau \to \tau$

Simply typed λ -terms t: Str_{{0,1}[τ] \rightarrow Bool define **languages** $L \subseteq \{0, 1\}^*$ Example: $t = \lambda s. s$ id not true : Str_{0,1}[Bool] \rightarrow Bool (even number of 1s)

 $t \overline{\operatorname{Oll}} \longrightarrow_{\beta} \overline{\operatorname{Oll}}$ id not true \longrightarrow_{β} id (not (not true)) \longrightarrow_{β} true

Theorem (Hillebrand & Kanellakis 1996)

All regular (i.e. MSO-definable) languages, and only those, can be defined this way.

Proof of STLC-definable \implies regular, by semantic evaluation

Theorem (Hillebrand & Kanellakis, LICS'96)

For any type A and any simply typed λ -term $t : \operatorname{Str}_{\Sigma}[\tau] \to \operatorname{Bool}$, the language $\mathscr{L}(t) = \{w \in \Sigma^* \mid t \overline{w} \to_{\beta}^* \operatorname{true}\}$ is regular.

Part 1 of proof.

Fix a type τ . Any *denotational semantics* [-] quotients words:

$$w \in \Sigma^* \rightsquigarrow \overline{w} : \mathsf{Str}[\tau] \rightsquigarrow \llbracket \overline{w} \rrbracket_{\mathsf{Str}_{\Sigma}[\tau]} \in \llbracket \mathsf{Str}_{\Sigma}[\tau] \rrbracket$$

 $\llbracket \overline{w} \rrbracket_{\operatorname{Str}_{\Sigma}[\tau]}$ determines behavior of w w.r.t. all $\operatorname{Str}_{\Sigma}[\tau] \to \operatorname{Bool}$ terms:

$$w \in \mathscr{L}(t) \iff t \,\overline{w} \to_{\beta}^{*} \operatorname{true} \underset{\text{assuming } \llbracket \operatorname{true} \rrbracket \neq \llbracket \operatorname{false} \rrbracket}{\longleftarrow} [\![t \,\overline{w}]\!] = [\![t]\!] (\llbracket \overline{w}]\!]) = [\![\operatorname{true}]\!]$$

Goal: to decide $\mathscr{L}(t)$, compute $w \mapsto \llbracket \overline{w} \rrbracket$ in some denotational model.

Proof of STLC-definable \implies regular, by semantic evaluation

Theorem (Hillebrand & Kanellakis, LICS'96)

For any type A and any simply typed λ -term $t : \operatorname{Str}_{\Sigma}[\tau] \to \operatorname{Bool}$, the language $\mathscr{L}(t) = \{ w \in \Sigma^* \mid t \overline{w} \to_{\beta}^* \operatorname{true} \}$ is regular.

Part 2 of proof.

We use [-]: STLC \rightarrow FinSet to build a DFA with states $Q = [[Str_{\Sigma}[\tau]]]$, acceptation as [[t]](-) = [[true]].

$$w \in \mathscr{L}(t) \iff \llbracket t \rrbracket \left(\llbracket \overline{w} \rrbracket_{\operatorname{Str}_{\Sigma}[\tau]} \right) = \llbracket \operatorname{true} \rrbracket \iff w \text{ accepted.}$$

Proof of STLC-definable \implies regular, by semantic evaluation

Theorem (Hillebrand & Kanellakis, LICS'96)

For any type A and any simply typed λ -term $t : \operatorname{Str}_{\Sigma}[\tau] \to \operatorname{Bool}$, the language $\mathscr{L}(t) = \{ w \in \Sigma^* \mid t \overline{w} \to_{\beta}^* \operatorname{true} \}$ is regular.

Part 2 of proof.

We use $\llbracket - \rrbracket$: STLC \rightarrow FinSet to build a DFA with states $Q = \llbracket \operatorname{Str}_{\Sigma}[\tau] \rrbracket$, acceptation as $\llbracket t \rrbracket(-) = \llbracket \operatorname{true} \rrbracket$. $(|Q| < \infty, \operatorname{e.g.} 2^{2^{134}} \operatorname{when} \tau = \operatorname{Bool} \& | \llbracket v \rrbracket | = 2 = |\Sigma|)$

$$w \in \mathscr{L}(t) \iff \llbracket t \rrbracket \left(\llbracket \overline{w} \rrbracket_{\operatorname{Str}_{\Sigma}[\tau]} \right) = \llbracket \operatorname{true} \rrbracket \iff w \text{ accepted.}$$

Takeaways from the Hillebrand–Kanellakis theorem

From the theorem statement

analogous to implicit complexity (e.g. "Light Linear Logic captures P")
 → implicit automata research programme (N. & Pradic), for instance:
 star-free languages in a non-commutative linear¹ λ-calculus (ICALP'20)

¹Affine in the paper, strictly linear in my PhD dissertation.

From the theorem statement

 analogous to implicit complexity (e.g. "Light Linear Logic captures P")
 → implicit automata research programme (N. & Pradic), for instance: star-free languages in a non-commutative linear¹ λ-calculus (ICALP'20)

From the proof: connection with finite automata via finitary semantics

¹Affine in the paper, strictly linear in my PhD dissertation.

From the theorem statement

 analogous to implicit complexity (e.g. "Light Linear Logic captures P")
 → implicit automata research programme (N. & Pradic), for instance: star-free languages in a non-commutative linear¹ λ-calculus (ICALP'20)

From the proof: connection with finite automata via finitary semantics

✓ for MSO on ∞ trees: alternating parity automata ↔ bespoke semantics
 → decidability proofs for trees generated by STλC+recursion
 [Kobayashi & Ong, Salvati & Walukiewicz, Grellois & Melliès, ...]

Combining the two: *implicit automata over infinite words/trees*

¹Affine in the paper, strictly linear in my PhD dissertation.

Conjecture

If $f: \Sigma^{\omega} \to \Gamma^{\omega}$ is computed by some infinitary simply typed λ -term $t: Str_{\Sigma}[\tau] \to Str_{\Gamma}$, then for every ω -regular $L \subseteq \Gamma^{\omega}$, the preimage $f^{-1}(L)$ is also ω -regular.

Should also hold on infinite trees...

Conjecture

If $f: \Sigma^{\omega} \to \Gamma^{\omega}$ is computed by some infinitary simply typed λ -term $t: Str_{\Sigma}[\tau] \to Str_{\Gamma}$, then for every ω -regular $L \subseteq \Gamma^{\omega}$, the preimage $f^{-1}(L)$ is also ω -regular.

Should also hold on infinite trees...

Corollary (assuming the conjecture is true)

L is ω -regular $\iff L = f^{-1}(\text{parity language})$ for such an *f*.

 \rightarrow because deterministic parity automata recognize all ω -regular languages (fails for infinite trees!)

Conjecture

If $f: \Sigma^{\omega} \to \Gamma^{\omega}$ is computed by some infinitary simply typed λ -term $t: Str_{\Sigma}[\tau] \to Str_{\Gamma}$, then for every ω -regular $L \subseteq \Gamma^{\omega}$, the preimage $f^{-1}(L)$ is also ω -regular.

Should also hold on infinite trees...

Corollary (assuming the conjecture is true)

L is ω -regular $\iff L = f^{-1}(\text{parity language})$ for such an *f*.

 \rightarrow because deterministic parity automata recognize all ω -regular languages (fails for infinite trees!)

Proof method for conjecture?

Conjecture

If $f: \Sigma^{\omega} \to \Gamma^{\omega}$ is computed by some infinitary simply typed λ -term $t: Str_{\Sigma}[\tau] \to Str_{\Gamma}$, then for every ω -regular $L \subseteq \Gamma^{\omega}$, the preimage $f^{-1}(L)$ is also ω -regular.

Should also hold on infinite trees...

Corollary (assuming the conjecture is true)

L is ω -regular $\iff L = f^{-1}(\text{parity language})$ for such an *f*.

 \rightarrow because deterministic parity automata recognize all ω -regular languages (fails for infinite trees!)

Proof method for conjecture?

Finitary semantics of infinitary ST λ C with parity conditions \rightsquigarrow colors

Cartesian closed categories (CCCs): semantics of simply typed λ -calculus

Coloring modality for *k* **colors in a CCC** \mathscr{C} [Melliès 2017; Walukiewicz 2019] $\Box A = \underbrace{A \times \cdots \times A}_{k+1 \text{ times}} + \text{ comultiplication } \Box A \rightarrow \Box \Box A \text{ taking max of indices}$

 \Box is a linear exponential comonad $\implies \mathscr{C}_{\Box}$ is also a CCC

- (Scott model of linear logic)_□ with clever interpretation of recursion
 = Grellois and Melliès's (2015) colored semantics for deciding MSO
- interpretation in Scott_{\square} = int. in Scott \circ ("coloring translation" i.e. int. in Λ_{\square}) over ST λ C (w/o recursion), where Λ = syntactic (or initial) CCC

The coloring translation, purely syntactically

On types:
$$\hat{o} = o$$
 and $\widehat{\sigma \to \tau} = \underbrace{\widehat{\sigma} \to \dots \to \widehat{\sigma}}_{k+1 \text{ times}} \to \widehat{\tau}$.

$$\overline{\Gamma, x : \tau \vdash x : \tau} \rightsquigarrow \overline{\widehat{\Gamma}, x^{(0)} : \widehat{\tau}, \dots, x^{(k)} : \widehat{\tau} \vdash x^{(0)} : \widehat{\tau}}$$

$$\frac{\Gamma, x : \sigma \vdash t : \tau}{\Gamma \vdash \lambda x. t : \sigma \to \tau} \rightsquigarrow \frac{\widehat{\Gamma}, x^{(0)} : \widehat{\sigma}, \dots, x^{(k)} : \widehat{\sigma} \vdash \widehat{t} : \widehat{\tau}}{\widehat{\Gamma} \vdash \lambda x. t : \sigma \to \tau}$$

$$\frac{\Gamma \vdash t : \sigma \to \tau \quad \Gamma \vdash u : \sigma}{\Gamma \vdash t u : \tau} \rightsquigarrow \frac{\widehat{\Gamma} \vdash \widehat{t} : \widehat{\sigma \to \tau} \quad \widehat{\Gamma} \vdash \uparrow_0 \widehat{u} : \widehat{\sigma}}{\widehat{\Gamma} \vdash \widehat{t} u = \widehat{t} (\uparrow_0 \widehat{u}) \dots (\uparrow_k \widehat{u}) : \widehat{\tau}}$$

where $\uparrow_c \hat{u} = \hat{u}[x^{(i)} := x^{(\max(c,i))} \mid x \in \operatorname{dom}(\Gamma), i = 0, \dots, k].$

Important remark

The syntactic presentation of the coloring translation extends effortlessly to *infinitary* simply typed λ -terms, by reading the rules coinductively.

colored Scott semantics only exists for $ST\lambda C$ +recursion; we would like to *define*: infinitary colored Scott := infinitary Scott \circ coloring translation

Important remark

The syntactic presentation of the coloring translation extends effortlessly to *infinitary* simply typed λ -terms, by reading the rules coinductively.

Important remark

The syntactic presentation of the coloring translation extends effortlessly to *infinitary* simply typed λ -terms, by reading the rules coinductively.

colored Scott semantics only exists for STλC+recursion; we would like to *define*:
infinitary colored Scott := infinitary Scott ∘ coloring translation
but this requires encoding the parity acceptance conditions
→ a subset of ∞ branches in the output of the translation are "accepting"
→ a *boundary* in the sense of [Melliès 2017]: introduces and studies a well-behaved Scott semantics for infinitary STλC with boundaries

Important remark

The syntactic presentation of the coloring translation extends effortlessly to *infinitary* simply typed λ -terms, by reading the rules coinductively.

colored Scott semantics only exists for $ST\lambda C$ +recursion; we would like to *define*: infinitary colored Scott := infinitary Scott \circ coloring translation

but this requires encoding the parity acceptance conditions \rightsquigarrow a subset of ∞ branches in the output of the translation are "accepting"

→ a *boundary* in the sense of [Melliès 2017]: introduces and studies a well-behaved Scott semantics for infinitary ST λ C with boundaries

Conjecture (needed to get invariance for infinitary colored Scott)

The infinitary colored translation with boundary is compatible with $\rightarrow_{\beta}^{\infty}$ *.*

Conclusion

Conjecture

The infinitary colored translation with boundary is compatible with $\rightarrow_{\beta}^{\infty}$ *.*

- would give us a well-behaved colored Scott semantics of infinitary $ST\lambda C$
- as a consequence we could prove our "implicit ω -automata" conjecture
- for more details (e.g. intersection types): cf. our ITRS'24 abstract

Conclusion

Conjecture

The infinitary colored translation with boundary is compatible with $\rightarrow_{\beta}^{\infty}$ *.*

- would give us a well-behaved colored Scott semantics of infinitary $ST\lambda C$
- as a consequence we could prove our "implicit ω -automata" conjecture
- for more details (e.g. intersection types): cf. our ITRS'24 abstract

Some points to remember

- Which infinite structures have decidable MSO theory? (for verification)
 → simply typed λ-calculus + recursion provides a large class of trees
 (*higher-order model checking* also reprovable from above conjecture!)
- regular languages in STλC [Hillebrand & Kanellakis 1996] → implicit automata
 + its proof: evaluation in a finitary semantics

Conclusion

Conjecture

The infinitary colored translation with boundary is compatible with $\rightarrow_{\beta}^{\infty}$ *.*

- would give us a well-behaved colored Scott semantics of infinitary $ST\lambda C$
- as a consequence we could prove our "implicit ω -automata" conjecture
- for more details (e.g. intersection types): cf. our ITRS'24 abstract

Some points to remember

- Which infinite structures have decidable MSO theory? (for verification)
 → simply typed λ-calculus + recursion provides a large class of trees
 (*higher-order model checking* also reprovable from above conjecture!)
- regular languages in STλC [Hillebrand & Kanellakis 1996] → implicit automata
 + its proof: evaluation in a finitary semantics