Strong Normalizability as a Finiteness Structure via the Taylor Expansion of λ -terms

Michele Pagani[°], Christine Tasson[°] and <u>Lionel Vaux</u>^{*}

^o IRIF, Université Paris Diderot, France
 * I2M, Aix-Marseille Université, France
 Supported by French ANR Project Coquas (ANR 12 JS02 006 01).

FoSSaCS @ ETAPS 2016 Eindhoven, NL, 4-7 April 2016

Everything is in the title

$\left(\forall M\in\Lambda_{+}\right)\,M\in\mathsf{SN}\iff\mathcal{T}\left(M\right)\in\mathfrak{F}$

We characterize the strong normalizability (SN) of (non-deterministic) λ -terms (Λ_+) as a finiteness structure (\mathfrak{F}) via Taylor expansion (\mathcal{T}).

Everything is in the title

$\left(\forall M\in\Lambda_{+}\right)\,M\in\mathsf{SN}\iff\mathcal{T}\left(M\right)\in\mathfrak{F}$

We characterize the strong normalizability (SN) of (non-deterministic) λ -terms (Λ_+) as a finiteness structure (\mathfrak{F}) via Taylor expansion (\mathcal{T}).

Quantitative semantics

A prime aged idea (Girard, '80s, before LL)

 λ -terms = analytic functions = power series Originally: for the λ -calculus, in an abstract categorical setting (coefficients are sets)

Quantitative semantics

A prime aged idea (Girard, '80s, before LL)

 λ -terms = analytic functions = power series Originally: for the λ -calculus, in an abstract categorical setting (coefficients are sets)

Finitness spaces (Ehrhard, early 2000's)

Reformulate q.s. in a linear logic setting using standard algebra:

- ▶ types \rightsquigarrow particular topological vector spaces: $\llbracket A \rrbracket \subseteq \mathbf{k}^{|A|} +$ some additional structure
- function terms \rightsquigarrow power series

Quantitative semantics

A prime aged idea (Girard, '80s, before LL)

 λ -terms = analytic functions = power series Originally: for the λ -calculus, in an abstract categorical setting (coefficients are sets)

Finitness spaces (Ehrhard, early 2000's)

Reformulate q.s. in a linear logic setting using standard algebra:

▶ types \rightsquigarrow particular topological vector spaces:

 $[\![A]\!]\subseteq \mathbf{k}^{|A|}\,+\,\mathrm{some\ additional\ structure}$

• function terms \rightsquigarrow power series

Differentiation of λ -terms (Ehrhard-Regnier 2003-2004)

So we can *differentiate* λ -terms, and compute their Taylor expansion! And one can mimick that in the syntax:

- ▶ differential λ -calculus
- a finitary fragment: resource λ -calculus = the target of Taylor expansion

Resource λ -calculus

Resource terms

Meaning: $\langle s \rangle [s_1, \ldots, s_n] = (Ds)_0 \cdot (s_1, \ldots, s_n)$

Resource reduction

$$\begin{array}{c} \langle \lambda x.s \rangle \ \bar{t} \rightarrow_{\rho} \partial_{x} s \cdot \bar{t} \quad (\text{anywhere}) \\ \partial_{x} s \cdot \bar{t} = \left\{ \begin{array}{c} \sum_{f \in \mathfrak{S}_{n}} s \left[t_{f(1)}, \ldots, t_{f(n)} / x_{1}, \ldots, x_{n} \right] & \text{if } \deg_{x}(s) = \# \bar{t} = n \\ 0 & \text{otherwise} \end{array} \right. \end{array}$$

Resource λ -calculus

Resource terms

Meaning: $\langle s \rangle [s_1, \ldots, s_n] = (Ds)_0 \cdot (s_1, \ldots, s_n)$

Resource reduction

$$\begin{array}{c} \langle \lambda x.s \rangle \ \bar{t} \rightarrow_{\rho} \partial_{x} s \cdot \bar{t} \quad (\text{anywhere}) \\ \partial_{x} s \cdot \bar{t} = \left\{ \begin{array}{c} \sum_{f \in \mathfrak{S}_{n}} s \left[t_{f(1)}, \ldots, t_{f(n)} / x_{1}, \ldots, x_{n} \right] & \text{if } \deg_{x}(s) = \# \bar{t} = n \\ 0 & \text{otherwise} \end{array} \right. \end{array}$$

linearity: $\lambda x.0 = 0$, $\langle s \rangle [t_1 + t_2, u] = \langle s \rangle [t_1, u] + \langle s \rangle [t_2, u], \dots$

Resource λ -calculus

Resource terms

Meaning: $\langle s \rangle [s_1, \ldots, s_n] = (Ds)_0 \cdot (s_1, \ldots, s_n)$

Resource reduction

$$\langle \lambda x.s \rangle \ \bar{t} \to_{\rho} \partial_x s \cdot \bar{t} \quad \text{(anywhere)}$$

$$\partial_x s \cdot \bar{t} = \begin{cases} \sum_{f \in \mathfrak{S}_n} s \left[t_{f(1)}, \dots, t_{f(n)} / x_1, \dots, x_n \right] & \text{if } \deg_x(s) = \# \bar{t} = n \\ 0 & \text{otherwise} \end{cases}$$

linearity: $\lambda x.0 = 0, \langle s \rangle [t_1 + t_2, u] = \langle s \rangle [t_1, u] + \langle s \rangle [t_2, u], \dots$

 Resource reduction preserves free variables, is size-decreasing, strongly confluent and normalizing.

Taylor expansion of λ -terms

Semantically, $(M) N = \sum_{n \in \mathbf{N}} \frac{1}{n!} \langle M \rangle N^n$ where $N^n = [N, \dots, N]$.

Taylor expansion of λ -terms

Semantically, $(M) N = \sum_{n \in \mathbb{N}} \frac{1}{n!} \langle M \rangle N^n$ where $N^n = [N, \dots, N]$. Taylor expansion: $\vec{\mathcal{T}}(M) \in \mathbf{Q}^{+\Delta}$

$$\vec{\mathcal{T}}((M) N) = \sum_{n \in \mathbf{N}} \frac{1}{n!} \left\langle \vec{\mathcal{T}}(M) \right\rangle \vec{\mathcal{T}}(N)^{n}$$
$$\vec{\mathcal{T}}(x) = x \quad \vec{\mathcal{T}}(\lambda x.M) = \lambda x.\vec{\mathcal{T}}(M)$$

Taylor expansion of λ -terms

Semantically, $(M) N = \sum_{n \in \mathbb{N}} \frac{1}{n!} \langle M \rangle N^n$ where $N^n = [N, \dots, N]$. Taylor expansion: $\vec{\mathcal{T}}(M) \in \mathbf{Q}^{+\Delta}$

$$\vec{\mathcal{T}}((M) N) = \sum_{n \in \mathbf{N}} \frac{1}{n!} \left\langle \vec{\mathcal{T}}(M) \right\rangle \vec{\mathcal{T}}(N)^{n}$$
$$\vec{\mathcal{T}}(x) = x \quad \vec{\mathcal{T}}(\lambda x.M) = \lambda x.\vec{\mathcal{T}}(M)$$

Theorem (Ehrhard-Regnier, CiE 2006) If $M \in \Lambda$, then $\vec{\mathcal{T}}(M)$ normalizes to $\vec{\mathcal{T}}(\mathsf{BT}(M))$.

Moral

In the ordinary λ -calculus $\mathsf{BT}(M) \simeq \mathsf{NF}(\vec{\mathcal{T}}(M))$.

Normalizing Taylor expansions

But how can $\vec{\mathcal{T}}(M)$ even normalize?

We want to set

$$\mathsf{NF}\left(\vec{\mathcal{T}}\left(M\right)\right) = \sum_{s \in \Delta} \vec{\mathcal{T}}\left(M\right)_{s} .\mathsf{NF}\left(s\right)$$

 \rightsquigarrow infinite sums (and in general we might consider all kinds of coefficients) \rightsquigarrow convergence?

Normalizing Taylor expansions: uniformity to the rescue! But how can $\vec{\mathcal{T}}(M)$ even normalize?

We want to set

$$\mathsf{NF}\left(\vec{\mathcal{T}}\left(M\right)\right) = \sum_{s \in \Delta} \vec{\mathcal{T}}\left(M\right)_{s} .\mathsf{NF}\left(s\right)$$

 \rightsquigarrow infinite sums (and in general we might consider all kinds of coefficients) \rightsquigarrow convergence?

Theorem (Ehrhard-Regnier 2004, published in TCS in 2008) Write $\mathcal{T}(M) = |\vec{\mathcal{T}}(M)|$. Then for all $t \in \Delta$, there is at most one $s \in \mathcal{T}(M)$ such that $\mathsf{NF}(s)_t \neq 0$.

Proof.

 λ -terms are uniform: their finitary approximants are pairwise coherent.

Normalizing Taylor expansions: uniformity to the rescue! But how can $\vec{\mathcal{T}}(M)$ even normalize?

We want to set

$$\mathsf{NF}\left(\vec{\mathcal{T}}\left(M\right)\right) = \sum_{s \in \Delta} \vec{\mathcal{T}}\left(M\right)_{s} .\mathsf{NF}\left(s\right)$$

 \rightsquigarrow infinite sums (and in general we might consider all kinds of coefficients) \rightsquigarrow convergence?

Theorem (Ehrhard-Regnier 2004, published in TCS in 2008) Write $\mathcal{T}(M) = |\vec{\mathcal{T}}(M)|$. Then for all $t \in \Delta$, there is at most one $s \in \mathcal{T}(M)$ such that $\mathsf{NF}(s)_t \neq 0$.

Proof.

 λ -terms are uniform: their finitary approximants are pairwise coherent. \Box

This fails in general NF $\left(\sum_{n \in \mathbf{N}} \langle \lambda x. x \rangle^n [y]\right) = ?$ $\langle \lambda x. x \rangle^n [y] = \langle \lambda x. x \rangle [\langle \lambda x. x \rangle [\cdots [y] \cdots]]$

$$\Lambda_+ \ni M, N, \ldots ::= x \mid \lambda x.M \mid (M) N \mid M + N$$

 $(\lambda x.M) N \rightarrow_{\beta} M [N/x]$

$$\begin{split} \Lambda_+ &\ni M, N, \dots ::= x \mid \lambda x.M \mid (M) N \mid M + N \\ & (\lambda x.M) N \rightarrow_\beta M \left[N/x \right] \\ & (M+N) P = (M) P + (N) P \end{split}$$

$$\Lambda_{+} \ni M, N, \dots ::= x \mid \lambda x.M \mid (M) N \mid M + N$$
$$(\lambda x.M) N \to_{\beta} M [N/x]$$
$$(M+N) P = (M) P + (N) P$$

Example Let $\delta_M = \lambda x. (M + (x) x)$ and $\infty_M = (\delta_M) \delta_M: \infty_M \to^*_\beta M + \infty_M.$

$$\Lambda_{+} \ni M, N, \dots ::= x \mid \lambda x.M \mid (M) N \mid M + N$$
$$(\lambda x.M) N \to_{\beta} M [N/x]$$
$$(M+N) P = (M) P + (N) P$$

Example

Let $\delta_M = \lambda x. (M + (x) x)$ and $\infty_M = (\delta_M) \delta_M: \infty_M \to^*_\beta M + \infty_M.$

Taylor expansion in a non uniform setting

$$\vec{\mathcal{T}}\left(M+N\right) = \vec{\mathcal{T}}\left(M\right) + \vec{\mathcal{T}}\left(N\right)$$

$$\Lambda_{+} \ni M, N, \dots ::= x \mid \lambda x.M \mid (M) N \mid M + N$$
$$(\lambda x.M) N \to_{\beta} M [N/x]$$
$$(M+N) P = (M) P + (N) P$$

Example

Then

Let $\delta_M = \lambda x. (M + (x) x)$ and $\infty_M = (\delta_M) \delta_M: \infty_M \to^*_{\beta} M + \infty_M.$

Taylor expansion in a non uniform setting

$$\vec{\mathcal{T}}\left(M+N\right) = \vec{\mathcal{T}}\left(M\right) + \vec{\mathcal{T}}\left(N\right)$$
$$\mathsf{NF}\left(\vec{\mathcal{T}}\left(\infty_{M}\right)\right) = ?$$

Finiteness structures to the rescue

The main artifact of Ehrhard's finiteness spaces: Definition

- If $a, a' \subseteq A$, write $a \perp a'$ iff $a \cap a'$ is finite.
- If $\mathfrak{S} \subseteq \mathfrak{P}(A)$, let $\mathfrak{S}^{\perp} := \{a' \subseteq A; \forall a \in \mathfrak{S}, a \perp a'\}.$
- A finiteness structure is any $\mathfrak{F} = \mathfrak{S}^{\perp}$.

When is $\vec{\mathcal{T}}(M)$ normalizable?

- Write $s \ge t$ if $s \to_{\rho}^{*} t + \cdots$.
- Let $\uparrow t = \{s \in \Delta; s \ge t\}.$
- ▶ $\vec{\mathcal{T}}(M)$ is normalizable iff for all normal $t \in \Delta, \mathcal{T}(M) \perp \uparrow t$.
- ► { $\uparrow t$; t normal $\in \Delta$ }[⊥] is the finiteness structure of (supports of) normalizable vectors.

Typed terms have a finitary Taylor expansion

Let system F_+ be system F plus $\frac{\Gamma \vdash M}{\Gamma}$

$$\frac{\Gamma \vdash M : A \quad \Gamma \vdash N : A}{\Gamma \vdash M + N : A}$$

Theorem (Ehrhard, LICS 2010) If $M \in \Lambda_+$ is typable in system F_+ , then $\mathcal{T}(M) \in \{\uparrow t ; t \in \Delta\}^{\perp}$.

Proof.

Manage sets of resource terms as if they were λ -terms, and follow the usual reducibility technique, associating a finiteness structure $\mathfrak{Fin}(A) \subseteq \{\uparrow t \; ; \; t \in \Delta\}^{\perp}$ with each type A.

- \blacktriangleright Typability in F can be relaxed to strong normalizability.
- ▶ Then the implication

$$M \in \mathsf{SN} \Rightarrow \mathcal{T}(M) \in \{\uparrow t \; ; \; t \in \Delta\}^{\perp}$$

can be reversed...

▶ provided the finiteness $\{\uparrow t ; t \in \Delta\}^{\perp}$ is refined to a tighter one.

 $M \in \mathsf{SN} \Rightarrow \mathcal{T}(M) \in \{\uparrow t \; ; \; t \in \Delta\}^{\perp}$

In the ordinary λ -calculus:

▶ SN = typability in system D (simple types + ∩)

• "any" proof by reducibility for simple types is valid for DSo we:

- ▶ introduce a system D₊ of intersection types for non uniform terms (this needs some care)
- prove that $M \in \mathsf{SN}$ implies $\Gamma \vdash M : A$ in D_+
- adapt Ehrhard's proof to D_+

 $\mathcal{T}(M) \in \{\uparrow t \; ; \; t \in \Delta\}^{\perp} \Rightarrow M \in \mathsf{SN}$

Finiteness prevents loops...

Consider $\delta_n = \lambda x. \langle x \rangle [x^n]$; then for all $n \in \mathbf{N}$, $\mathcal{T}(\Omega) \ni \langle \delta_n \rangle [\delta_0, \delta_0, \delta_1 \dots, \delta_{n-1}] \ge \langle \delta_0 \rangle [] \to_{\rho} 0$. Hence $\mathcal{T}(\Omega) \notin \{\uparrow t ; t \in \Delta\}^{\perp}$.

$\mathcal{T}(M) \in \{\uparrow t \; ; \; t \in \Delta\}^{\perp} \not\Rightarrow M \in \mathsf{SN}$

Finiteness prevents loops...

Consider $\delta_n = \lambda x. \langle x \rangle [x^n]$; then for all $n \in \mathbf{N}$, $\mathcal{T}(\Omega) \ni \langle \delta_n \rangle [\delta_0, \delta_0, \delta_1 \dots, \delta_{n-1}] \ge \langle \delta_0 \rangle [] \to_{\rho} 0$. Hence $\mathcal{T}(\Omega) \notin \{\uparrow t ; t \in \Delta\}^{\perp}$.

... but not divergence

Let $\Delta_3 := \lambda x. (x) x x$ and $\Omega_3 := (\Delta_3) \Delta_3$, then $\mathcal{T}(\Omega_3) \perp \uparrow s$ for all s.

$\mathcal{T}(M) \in \{\uparrow t \; ; \; t \in \Delta\}^{\perp} \not\Rightarrow M \in \mathsf{SN}$

Finiteness prevents loops...

Consider $\delta_n = \lambda x. \langle x \rangle [x^n]$; then for all $n \in \mathbf{N}$, $\mathcal{T}(\Omega) \ni \langle \delta_n \rangle [\delta_0, \delta_0, \delta_1 \dots, \delta_{n-1}] \ge \langle \delta_0 \rangle [] \to_{\rho} 0$. Hence $\mathcal{T}(\Omega) \notin \{\uparrow t ; t \in \Delta\}^{\perp}$.

... but not divergence

Let $\Delta_3 := \lambda x. (x) x x$ and $\Omega_3 := (\Delta_3) \Delta_3$, then $\mathcal{T}(\Omega_3) \perp \uparrow s$ for all s. However, writing $\ell(M)$ for the only linear term $\in \mathcal{T}(M)$:

$$\mathcal{T}(\Omega_3) \not\perp \bigcup_{\Omega_3 \to_{\beta}^* M} \uparrow \ell(M)$$

$\mathcal{T}(M) \in \{\uparrow t \; ; \; t \in \Delta\}^{\perp} \not\Rightarrow M \in \mathsf{SN}$

Finiteness prevents loops...

Consider $\delta_n = \lambda x. \langle x \rangle [x^n]$; then for all $n \in \mathbf{N}$, $\mathcal{T}(\Omega) \ni \langle \delta_n \rangle [\delta_0, \delta_0, \delta_1 \dots, \delta_{n-1}] \ge \langle \delta_0 \rangle [] \to_{\rho} 0$. Hence $\mathcal{T}(\Omega) \notin \{\uparrow t ; t \in \Delta\}^{\perp}$.

... but not divergence

Let $\Delta_3 := \lambda x. (x) x x$ and $\Omega_3 := (\Delta_3) \Delta_3$, then $\mathcal{T}(\Omega_3) \perp \uparrow s$ for all s. However, writing $\ell(M)$ for the only linear term $\in \mathcal{T}(M)$:

$$\mathcal{T}(\Omega_3) \not\perp \bigcup_{\Omega_3 \to_{\beta}^* M} \uparrow \ell(M)$$

Fix: add more tests

- Consider a structure $\mathfrak{S} \subseteq \mathfrak{P}(\Delta)$ and let $\mathfrak{F}_{\mathfrak{S}} = \{\uparrow a ; a \in \mathfrak{S}\}^{\perp}$ with $\uparrow a = \bigcup_{s \in a} \uparrow s$.
- ▶ Idea: \mathfrak{S} is a set of *tests*: *M* passes the test $a \in \mathfrak{S}$ if $\mathcal{T}(M) \perp \uparrow a$.
- ▶ Ehrhard's finiteness is $\mathfrak{F}_{\mathfrak{P}_f(\Delta)}$: we need to consider infinite tests.
- Of course, not all \mathfrak{S} are acceptable, otherwise we reject too many terms (consider $\mathfrak{S} = \mathfrak{P}(\Delta)$).

- We can adapt the reducibility proof and show that $M \in SN \Rightarrow \mathcal{T}(M) \in \mathfrak{F}_{\mathfrak{S}}$ provided \mathfrak{S} satisfies:
 - for all $n \in \mathbf{N}$, for all $a \in \mathfrak{S}$, $\{s \in a; \text{ height}(s) = n\}$ is finite.
 - ▶ + some additional, purely technical conditions.

- We can adapt the reducibility proof and show that $M \in SN \Rightarrow \mathcal{T}(M) \in \mathfrak{F}_{\mathfrak{S}}$ provided \mathfrak{S} satisfies:
 - for all $n \in \mathbf{N}$, for all $a \in \mathfrak{S}$, $\{s \in a; \text{ height}(s) = n\}$ is finite.
 - ▶ + some additional, purely technical conditions.
- ▶ $\mathcal{T}(M) \in \mathfrak{F}_{\mathfrak{S}} \Rightarrow M \in \mathsf{SN}$ as soon as \mathfrak{S} contains all sets of linear resource terms

- We can adapt the reducibility proof and show that $M \in SN \Rightarrow \mathcal{T}(M) \in \mathfrak{F}_{\mathfrak{S}}$ provided \mathfrak{S} satisfies:
 - for all $n \in \mathbf{N}$, for all $a \in \mathfrak{S}$, $\{s \in a; \text{ height}(s) = n\}$ is finite.
 - ▶ + some additional, purely technical conditions.
- ▶ $\mathcal{T}(M) \in \mathfrak{F}_{\mathfrak{S}} \Rightarrow M \in \mathsf{SN}$ as soon as \mathfrak{S} contains all sets of linear resource terms
- ▶ We want $\{s\} \in \mathfrak{S}$ for all $s \in \Delta$ (or at least those in normal form) so that $\mathcal{T}(M) \in \mathfrak{F}_{\mathfrak{S}} \Rightarrow \mathsf{NF}\left(\vec{\mathcal{T}}(M)\right)$ is defined.

- We can adapt the reducibility proof and show that $M \in SN \Rightarrow \mathcal{T}(M) \in \mathfrak{F}_{\mathfrak{S}}$ provided \mathfrak{S} satisfies:
 - for all $n \in \mathbf{N}$, for all $a \in \mathfrak{S}$, $\{s \in a; \text{ height}(s) = n\}$ is finite.
 - \blacktriangleright + some additional, purely technical conditions.
- ▶ $\mathcal{T}(M) \in \mathfrak{F}_{\mathfrak{S}} \Rightarrow M \in \mathsf{SN}$ as soon as \mathfrak{S} contains all sets of linear resource terms
- We want $\{s\} \in \mathfrak{S}$ for all $s \in \Delta$ (or at least those in normal form) so that $\mathcal{T}(M) \in \mathfrak{F}_{\mathfrak{S}} \Rightarrow \mathsf{NF}\left(\vec{\mathcal{T}}(M)\right)$ is defined.

Example

 $\mathfrak{B} = \{a \subseteq \Delta; \ \#(a) \text{ is bounded}\} \text{ where } \#(a) = \{\#(s); \ s \in a\} \text{ and } \#(s) \text{ is the maximum size of a bag of arguments in } s.$

Theorem (Pagani-Tasson-V.)

The following are equivalent:

- ► $M \in SN;$
- $\blacktriangleright \mathcal{T}(M) \in \mathfrak{F}_{\mathfrak{B}}.$

Moreover, in this case, $\vec{\mathcal{T}}(M)$ is normalizable.

Conclusion

We are happy.

We have established a nice and novel characterization of SN .

Conclusion No, wait!

We are happy. We have established a nice and novel characterization of $\mathsf{SN}.$

Are we? The really useful bit is that:

 $the \ Taylor \ expansion \ of \ a \ strongly \ normalizable \ term \ is \ normalizable$

which is a bit frustrating (why "strongly"?).

Conclusion No, wait! Conclusion

We are happy.

We have established a nice and novel characterization of SN.

Are we? The really useful bit is that:

the Taylor expansion of a strongly normalizable term is normalizable

which is a bit frustrating (why "strongly"?).

Yes we are: plenty of future work!

- Our machinery is modular enough that it can be adapted to weak- and head-normalizability (WIP with Pagani and Tasson).
- ► That $\vec{\mathcal{T}}(\mathsf{NF}(M)) = \mathsf{NF}(\vec{\mathcal{T}}(M))$ follows from the fact that we can track β -reduction through Taylor expansion (*WIP*).
- ► Towards a semantically founded notion of Böhm trees for various non uniform settings (quantitative non-determinism, probabilistic stuff, *etc.*).

The end

Thanks for your attention.

Thanks for your attention. Questions?