A syntactic approach to Strong Normalization through decreasing measures

Cristian Sottile

CONICET & Universidad de Buenos Aires

20th Latin American Symposium on Mathematical Logic (XX SLALM)

July 1-5, 2024







Instituto de Ciencias de la Computación





SN proofs: a brief overview

Reducibility [Tait'67, Girard'72]

Interpretation of types into sets of well-behaved terms

$$\llbracket A \to B \rrbracket \hspace{0.2cm} = \hspace{0.2cm} \{ \hspace{0.2cm} t \hspace{0.2cm} | \hspace{0.2cm} \forall s \in \llbracket A \rrbracket. \hspace{0.2cm} t s \in \llbracket B \rrbracket \hspace{0.2cm} \}$$

Reduction of SN to WN [Nederpelt'73, Klop'80]

- Through different flavours of λI ideas
- A WN proof

Why?

Why decreasing measures?

- insight
- intuition
- metrics

The koan #26

- Posed by Gödel
- Submitted by Barendregt
- ▶ Find an "easy" mapping from λ^{\rightarrow} to ordinals

Why "syntactic"

- sort of convention
- soft classification of SN proofs
- maybe abstract vs concrete would be better?
- external vs internal ?
- we stick to the convention

syntactic = "internal" analysis over the structure of terms or the rewriting relation

semantic reducibility (RC)

syntactic

decreasing measures (DM) reduction of SN to WN (NK)

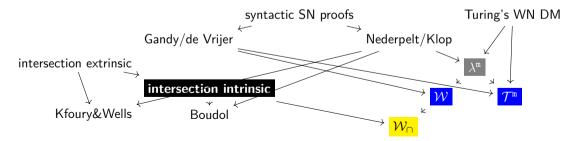
Our work

[Barenbaum & Sottile FSCD'23]

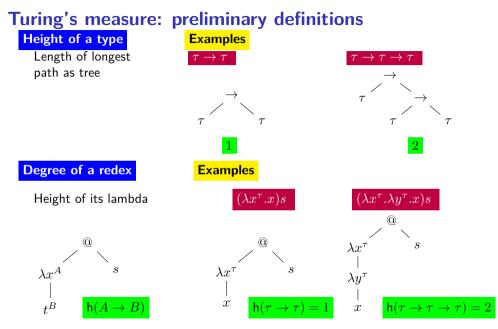
- An auxiliar calculus λ^{m} to manipulate (non-)erasure through memories
- > A simple measure \mathcal{W} based on counting memories
- ► A complex measure \mathcal{T}^{m} generalizing Turing's WN one

[Work in progress with Barenbaum & Ronchi della Rocca]

- A presentation of **idempotent intersection types a la Church**
- ▶ An adaptation of W to idempotent intersection types, $\frac{W_{\cap}}{W_{\cap}}$



The auxiliar non-erasing λ^{m} -calculus



Cristian Sottile

Turing's measure: Weak Normalization

Map terms \mapsto multiset of the redex degrees

Example

$$\mathcal{T}((\lambda x^{\tau}.\lambda y^{\tau}.x)\underbrace{(\lambda x^{\tau}.x)s}_{1}) = [2,1]$$

WN: choosing the redex to contract

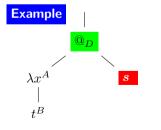
has the greatest degree

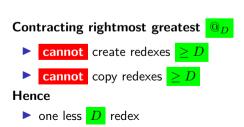
Two crucial observations [Turing, 1940s]

1. a redex cannot create redexes of greater or equal degree

 $\mathcal{T}(M) = [d \mid R \text{ is a redex of degree } d \text{ in } M]$

- 2. a redex can copy redexes of any degree
 - rightmost occurrence of that degree





Cristian Sottile

The auxiliar non-erasing λ^{m} -calculus

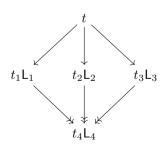
Definition

 $t ::= x \mid \lambda x.t \mid tt \mid t\{t\}$

$$[\lambda x.t)s \to_m t[s/x]$$
 {s

Why not to erase?

- Nederpelt-Klop's: INC WCR WN \Rightarrow DFC
- Retain information



Operations

- weight of a term: w(t) = amount of memoriese.g. $w(x\{y\{z\}\}\{w\}) = 3$
- simplification of a term: $S_D(t) =$ "bottom-up" contraction of all D redexes $S_*(t) = S_1(\ldots S_{maxdeg}(t) \ldots)$

Properties

► WCR ► WN ► SR

Properties

- Reduction arrives at simplification $t \rightarrow_m^* S_*(t)$
- $\mathsf{S}_*(t) = \mathtt{nf}(t)$ Simplification is normal form

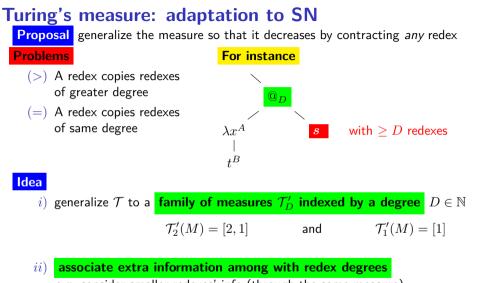
$\mathcal{W}:$ counting memories

Measure ${\cal W}$

Theorem

$$M \to_{\beta} N \implies \mathcal{W}(M) > \mathcal{W}(N)$$

\mathcal{T}^{m} : generalizing Turing's WN measure



e.g. consider smaller redexes' info (through the same measure)

 $\mathcal{T}_{2}'(M) = [(2, \mathcal{T}_{1}'(M)), (1, [])] \qquad \qquad \mathcal{T}_{1}'(M) = [(1, [])]$



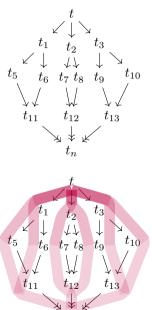
$$\mathcal{T}_2'(M) = [(2, ?), (1, ?)]$$

Idea

Development of degree Dreduction involving only redexes D

All developments of degree D

paths of the complete D-reduction graph from t



A syntactic approach to Strong Normalization through decreasing measures

$\label{eq:W_n} \mathcal{W}_n \text{:} \\ \text{extending } \mathcal{W} \text{ to} \\ \text{Idempotent Intersection Types} \\ \end{array}$

Motivation

Existing decreasing measures

[Kfoury & Wells'95]

- Domain of DM: multiset of numbers
- ► **Methodology:** WN ⇒ SN + DM proving WN (indirect)
- Auxiliary calculus: a la Curry

[Boudol'03]

- Domain of DM: pair of numbers
- Methodology: WN \implies SN + DM proving WN (indirect)
- Auxiliary calculus: a la Church, ad hoc

Our proposal Barenbaum, Ronchi della Rocca & Sottile (WIP)

- Domain of DM: number
- Methodology: DM proving SN (direct)
- Auxiliary calculus: a la Church, correspondent of a la Curry calculus

Idempotent Intersection Types (a la Curry) Key idea

- Variables can have multiple types
- Hence a term can have truly different (non-unifiable) types

Very powerful at charaterizing properties

The typing rule

$$\frac{(\Gamma \vdash N : A_i)_{i \in 1..n} \quad A_i \neq A_j}{\Gamma \Vdash N : \{A_1, \dots, A_n\}} \ e - multi$$

Example

Let

$$A = \tau \to \tau \qquad x : \{ \boldsymbol{A} \to \boldsymbol{A}, \boldsymbol{A} \} \vdash \boldsymbol{x}\boldsymbol{x} : A \qquad \vdash \lambda x.x : \{ \boldsymbol{A} \to \boldsymbol{A}, \boldsymbol{A} \}$$

Then

$$\vdash (\lambda x. xx)(\lambda x. x) : A \qquad (\lambda x. xx)(\lambda x. x) \rightarrow_{\beta} (\lambda x. x)(\lambda x. x)$$

Cristian Sottile

A syntactic approach to Strong Normalization through decreasing measures

e.g. $x : \{\tau, \tau \to \tau\} \vdash x : \tau$

Idempotent Intersection Types a la Church

Key idea

Variables can have multiple types defined a priori

Hence a term modulo erasure can have truly different (non-unifiable) types

Motivation

- \triangleright λ^{m} is a la Church (easier syntactic analysis)
- abscence of standard correspondent Church version of Curry system

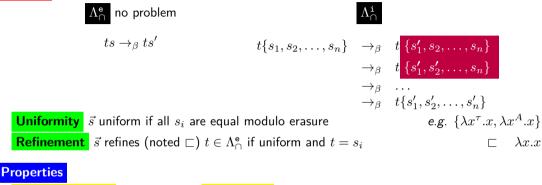
e.g. $x: \{\tau, \tau \to \tau\} \vdash x^{\tau}: \tau$

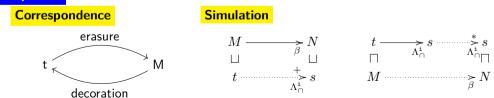
Idempotent Intersection Types a la Church Key changes Type unicity

 \blacktriangleright Λ^{e}_{\Box} assigns **multiple** types to each term \blacktriangleright Λ^{i}_{\Box} assigns **one** type to each term $\implies \qquad \frac{(\Gamma \vdash s_i : A_i)_{i \in 1..n} \quad A_i \neq A_j}{\Gamma \Vdash \{s_1, \dots, s_n\} : \{A_1, \dots, A_n\}}$ $\frac{(\Gamma \vdash N : A_i)_{i \in 1..n} \quad A_i \neq A_j}{\Gamma \Vdash N : \{A_1, \dots, A_n\}} \ e$ **Reduction refinement** \blacktriangleright Λ^{e}_{\cap} agnostic substitution \blacktriangleright Λ_{\Box}^{i} depending (on types) substitution Recall Λ^{e}_{\cap} $\vdash \lambda x.x: \{ A \rightarrow A, A \}$ $(\lambda x. xx)(\lambda x. x) \rightarrow_{\beta} (\lambda x. x)(\lambda x. x)$ $x: \{ \boldsymbol{A} \to \boldsymbol{A}, \boldsymbol{A} \} \vdash x^{\boldsymbol{A} \to \boldsymbol{A}} x^{\boldsymbol{A}} : \boldsymbol{A} \qquad \vdash \lambda x^{\boldsymbol{A}} x: \boldsymbol{A} \to \boldsymbol{A} \qquad \vdash \lambda x^{\boldsymbol{\tau}} x: \boldsymbol{A}$ Now Then $(\lambda x^{\{A \to A, A\}} . x^{A \to A} x^{A}) \{\lambda x^{A} . x, \lambda x^{\tau} . x\} \to_{\beta} (\lambda x^{A} . x) (\lambda x^{\tau} . x)$ $(\lambda x.t)s \to_{\beta} t[s/x] \qquad \Longrightarrow \qquad (\lambda x^{\vec{A}}.t)\vec{s} \to_{\beta} t[s_1/x^{A_1}]\dots[s_n/x^{A_n}]$ So

Idempotent Intersection Types a la Church Correspondence

Problem Reducing the argument of an application





Introducing memories in Λ_{\cap}^{i}

Extension to λ_{\cap}^{m}

- Addition of memories to the terms in Λ^i_{\cap}
- Adaptation of definitions, properties and proofs of λ^{m} to multi-terms and multi-types

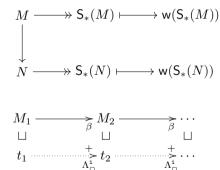
Definition

Measure \mathcal{W}_{\cap}

$$\mathcal{W}(M) = \mathsf{w}(\mathsf{S}_*(M))$$

Strong Normalization of $\Lambda_{\cap}^{\rm e}$

- $\blacktriangleright\,$ SN of $\Lambda^{\rm i}_\cap$
- Correspondence
- Simulation



Conclusions and future work

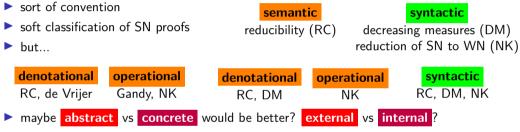
Conclusions

- Overview of techniques for proving Strong Normalization
- Decreasing measures
- Auxiliar non-erasing λ^m calculus, which allowed us to:
 - define \mathcal{W} : DM based on counting accumulated memories in λ^{m}
 - extend \mathcal{W} to Λ_{\cap} , obtaining a simpler measure than existing ones
 - ▶ generalize Turing's WN measure to SN by adding smaller measures of *D*-reachable terms

Future work

- Build a decreasing measure to System F
- Formalize them in a proof assistant
- Adapt \mathcal{W} to idempotent intersection types characterizing head normal forms
- Further compare our measures with those by Gandy and de Vrijer

Why "syntactic"



we stick to the soft convention

syntactic = "internal" analysis over the **structure of terms** or the **rewriting relation**

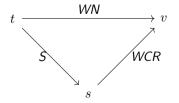
The auxiliar λ^m -calculus

Motivation

$$\beta$$
 is erasing $(\lambda x.y)t \rightarrow_{\beta} y$

A motivation not to erase

- $\blacktriangleright \text{ Klop-Nederpelt lemma } INC \land WCR \land WN \implies SN \land CR$
- ▶ We can obtain a decreasing measure from $INC \land WCR \land WN$
 - \blacktriangleright by WN there is a normal form v for any t
 - by WCR it is the same for every reduct s of t
 - ▶ by INC inc(t) < inc(s) < inc(v)
 - $\textbf{b} \quad \textit{dec}(t) = \textit{inc}(v) \textit{inc}(t)$



Intuitive definition of $\ensuremath{\mathcal{W}}$

Turing's measure "failing" example

Example: copying a redex of greater degree

$$\begin{array}{cccc} I_1 = \lambda x^{\tau}.x & \delta(I_1 x) &= \mathsf{h}(\tau \to \tau) &= 1 \\ I_2 = \lambda x^{\tau \to \tau}.x & \delta(I_2 I_1) &= \mathsf{h}((\tau \to \tau) \to (\tau \to \tau)) = 2 \\ K = \lambda x^{\tau}.\lambda y^{\tau}.x & \delta(K_{_}) &= \mathsf{h}(\tau \to \tau \to \tau) &= 2 \\ S_{KI} = \lambda x^{\tau}.K x \left(I_1 x\right) & \delta(S_{KI_}) = \mathsf{h}(\tau \to \tau) &= 1 \\ \end{array} \\ \begin{array}{c} \mathcal{T}(S_{\frac{K}{52} \frac{I}{11}} \left(\frac{I_2 I_1}{\mathsf{U}^2} x\right)) = \{ \stackrel{2}{2}, \stackrel{2}{2}, \stackrel{1}{\mathsf{U}}, \stackrel{1}{\mathsf{T}} \} & \frac{\mathcal{T}(K \left(\frac{I_2 I_1}{\mathsf{U}^2} x\right) \left(I_1 \left(\frac{I_2 I_1}{\mathsf{U}} x\right)\right)) = \{ \stackrel{2}{2}, \stackrel{2}{\mathsf{U}}, \stackrel{2}{\mathsf{U}}, \stackrel{1}{\mathsf{T}} \} \\ \hline \end{array}$$

A first attempt: \mathcal{T}' measure

Problems

- (>) A redex copies redexes of greater degree
- (=) A redex copies redexes of same degree

Idea

$$\mathcal{T}(M) = [2, 1] \longrightarrow \mathcal{T}(N) = [2, 2]$$
$$\mathcal{T}(M) = [1, 1] \longrightarrow \mathcal{T}(N) = [1, 1]$$

i) generalize $\mathcal T$ to a family of measures $\mathcal T'_D$ indexed by a degree $D\in\mathbb N,$ so e.g.

$$\mathcal{T}_2'(M) = [\begin{smallmatrix} 2 \\ \mathsf{S} \\ \mathsf{R} \end{smallmatrix} \qquad \text{and} \qquad \mathcal{T}_1'(M) = [\begin{smallmatrix} 1 \\ \mathsf{R} \\ \mathsf{R} \end{smallmatrix}$$

ii) instead of counting redex degrees in an isolated way, consider also the information about remaining smaller redexes, so *e.g.*

$$\mathcal{T}_2'(M) = \left[\begin{array}{c} (2 \atop \mathsf{S}, \mathcal{T}_1'(M)), \ (\frac{1}{\mathsf{R}}, []) \end{array} \right] \qquad \qquad \mathcal{T}_1'(M) = \left[\begin{array}{c} (1 \atop \mathsf{R}, []) \end{array} \right]$$

Definition

- $\blacktriangleright \ {\mathcal T}'_D(M) = [(i,{\mathcal T}'_{i-1}(M)) \ | \ R \text{ is a redex of degree } i \leq D \text{ in } M]$
- $\blacktriangleright \ {\cal T}'(M) = {\cal T}'_D(M)$ where D is the maximum degree of M

A first attempt: \mathcal{T}' measure

A working? example (>)

Definition

▶
$$\mathcal{T}'_D(M) = [(d, \mathcal{T}'_{d-1}(M)) \mid R \text{ is a redex of degree } d \leq D \text{ in } M]$$

▶ $\mathcal{T}'(M) = \mathcal{T}'_D(M)$ where D is the maximum degree of M

Example

$$\begin{array}{cccc} M & = & \underbrace{S_{\underline{K}} \underset{\underline{S_2} \, \mathrm{T}_1}{I} (\underline{I_2} \, \underline{I_1} \, x)}_{\mathbf{R}1} & \longrightarrow_{\beta} & & \underbrace{K \, (\underline{I_2} \, \underline{I_1} \, x) \, (I_1 \, (\underline{I_2} \, \underline{I_1} \, x))}_{\mathbf{S}2} & = & N \\ & & \underbrace{U'^2}_{\mathbf{S}2} & \underbrace{U''^2}_{\mathbf{T}1} \end{array}$$

$$\begin{aligned} \mathcal{T}_{2}'(M) &= \left[\begin{array}{c} (2, \mathcal{T}_{1}'(M)), \ (2, \mathcal{T}_{1}'(M)), \ (1, \|]), \ (1, \|]) \end{array} \right] & \mathcal{T}_{1}'(M) = \left[\begin{array}{c} (1, \|]), \ (1, \|]) \end{array} \right] \\ \mathcal{T}_{2}'(N) &= \left[\begin{array}{c} (2, \mathcal{T}_{1}'(M)), \ (2, \mathcal{T}_{1}'(M)), \ (2, \mathcal{T}_{1}'(M)), \ (1, \|]) \end{array} \right] & \mathcal{T}_{1}'(N) = \left[\begin{array}{c} (1, \|]) \end{array} \right] \end{aligned}$$

(2, [(1, []), (1, [])]) > (2, [(1, [])])

A first attempt: \mathcal{T}' measure

A failing example (=)

Definition

▶
$$\mathcal{T}'_D(M) = [(d, \mathcal{T}'_{d-1}(M)) \mid R \text{ is a redex of degree } d \leq D \text{ in } M]$$

▶ $\mathcal{T}'(M) = \mathcal{T}'_D(M)$ where D is the maximum degree of M

Example Example

$$M = \underbrace{S_{\underline{K}}}_{\mathbf{S}_{2}} \underbrace{I}_{\mathbf{1}} \underbrace{(I_{1} x)}_{\mathbf{R}_{1}} \qquad \longrightarrow_{\beta} \qquad \underbrace{K(\underline{I}_{1} x)}_{\underline{U'1}} \underbrace{((I_{1} x))}_{\underline{U''1}} = N$$

$$\mathcal{T}_{2}'(M) = [\ (\underset{\mathsf{S}}{2}, \mathcal{T}_{1}'(M)), \ (\underset{\mathsf{R}}{1}, []), \ (\underset{\mathsf{I}}{1}, []), \ (\underset{\mathsf{U}}{1}, []), \] \qquad \qquad \mathcal{T}_{1}'(M) = [\ (\underset{\mathsf{R}}{1}, []), \ (\underset{\mathsf{U}}{1}, []), \ (\underset{\mathsf{U}}{1}, []), \]$$

 $\mathcal{T}_{2}'(N) = [\ (\underset{\mathsf{S}}{2}, \mathcal{T}_{1}'(M)), \ (\underset{\mathsf{T}}{1}, []), \ (\underset{\mathsf{U}'}{1}, []), \ (\underset{\mathsf{U}''}{1}, []) \] \qquad \qquad \mathcal{T}_{1}'(N) = [\ (\underset{\mathsf{T}}{1}, []), \ (\underset{\mathsf{U}'}{1}, []), \ (\underset{\mathsf{U}''}{1}, []) \]$

(2, [(1, []), (1, []), (1, [])]) = (2, [(1, []), (1, []), (1, [])])

A second attempt: \mathcal{T}^{β} measure

Definition (development of a set of redexes)

reduction sequence where each step corresponds to a residual of a redex in the set

- ▶ a **residual** is a copy of a redex left after contracting another
- notation: $\rho: m \xrightarrow{D^*}_{\beta} m'$

Idea

- i) generalize $\mathcal T$ to a family of measures $\mathcal T_D^\beta$ indexed by a degree $D\in\mathbb N$
- ii) instead of isolatedly counting redexes degrees, consider:
 - from set of redexes of degree D
 - ► target M' from every development $\rho: M \xrightarrow{D}_{\beta}^{*} M'$
 - multiset of those $\mathcal{T}^{\beta}_{D-1}(M')$

Definition

$$\begin{split} \mathcal{T}_D^\beta(M) &= [\ (i,\mathcal{V}_i^\beta(M)) \mid R \text{ is a redex of degree } i \leq D \text{ in } M \] \\ \mathcal{V}_D^\beta(M) &= [\ \mathcal{T}_{D-1}^\beta(M') \mid \rho : M \xrightarrow{D}_{\beta}^* M' \] \end{split}$$

Problem: our technique to prove it decreases does not work because of erasing

A second attempt: \mathcal{T}^{β} measure

Definition

$$\begin{split} \mathcal{T}_D^\beta(M) &= [\ (i,\mathcal{V}_i^\beta(M)) \mid R \text{ is a redex of degree } i \leq D \text{ in } M \\ \mathcal{V}_D^\beta(M) &= [\ \mathcal{T}_{D-1}^\beta(M') \mid \rho : M \xrightarrow{D}_{\beta}^* M' \] \end{split}$$

Reasoning about the auxiliar measure \mathcal{V}^{β}_D

Consider

$$M \underset{R}{\to_{\beta}} N \qquad \mathcal{T}_{D}^{\beta}(M) > \mathcal{T}_{D}^{\beta}(N) \qquad \mathcal{V}_{D}^{\beta}(M) > \mathcal{V}_{D}^{\beta}(N)$$

- 1. Copying a redex of same degree (=)
 - ▶ injective mapping from devs of $\mathcal{V}_D^m(N)$ to devs of $\mathcal{V}_D^m(M)$ $R\rho: M \to_\beta N \to_\beta^* N'$

$$\mathcal{V}_D^\beta(M) > \mathcal{V}_D^\beta(N) \qquad \qquad \mathcal{T}_D^\beta(M) > \mathcal{T}_D^\beta(N)$$

- 2. Copying a redex of higher degree (>)
 - \blacktriangleright not clear the same can be done: a ho may erase R

$$\mathcal{V}_D^\beta(M') = \mathcal{V}_D^\beta(N') \qquad \qquad \mathcal{T}_D^\beta(M') = \mathcal{T}_D^\beta(N')$$

\mathcal{T}^m measure

Idea

- i) generalize $\mathcal T$ to a family of measures $\mathcal T_D^m$ indexed by a degree $D\in\mathbb N$
- ii) instead of isolatedly counting redexes degrees,

consider the multiset of the measures \mathcal{T}_{D-1}^m of every target of a development of degree D

Definition

$$\begin{split} \mathcal{T}_D^m(t) &= \left[\ (i,\mathcal{V}_i^m(t)) \ | \ R \text{ is a redex of degree } i \leq D \text{ in } t \ \right] \\ \mathcal{V}_D^m(t) &= \left[\ \mathcal{T}_{D-1}^m(t') \ | \ \rho: t \xrightarrow{D}_m^* t' \ \right] \end{split}$$

Lemmas

- **Forget/decrease**: forgetful reduction \triangleright decreases \mathcal{T}^m
- ► **High/increase**: contracting a redex of degree D > i increases (non-strictly) \mathcal{T}_i^m only $\leq i$, no D, in \mathcal{T}_i^m no erasing of any $\leq i$ maybe copies of $\leq i$
- ► Low/decrease: contracting a redex of degree i < D decreases (strictly) \mathcal{T}_D^m injective mappings from devs of $\mathcal{V}_D^m(N)$ to devs of $\mathcal{V}_D^m(M)$

Theorem

$$M \to_{\beta} N \implies \mathcal{T}^m(M) > \mathcal{T}^m(N)$$