# Midiendo programas para demostrar terminación

Pablo Barenbaum
UNQ/CONICET & ICC/UBA

Cristian Sottile
ICC/UBA/CONICET & UNQ

V Jornadas de Investigadores en Formación en Ciencia y Tecnología
Departamento de Ciencia y Tecnología
Universidad Nacional de Quilmes

Bernal, Buenos Aires, Argentina

28 de Septiembre de 2023

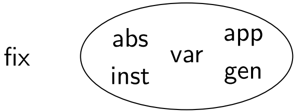
# Outline

- Terminación de programas
- Medidas decrecientes
- La medida de Turing
- El cálculo auxiliar sin borrado $\lambda^m$
- La medida $\mathcal{W}$: basada en operaciones sobre $\lambda^m$
- La medida $\mathcal{T}^m$: generalización de una medida dada por Turing

# Terminación de programas

## La propiedad de Terminación

- Resoluble en lenguajes no Turing completos como $\lambda^{\rightarrow}$, $\lambda^2$, $CoC$, Coq, Agda (TC se recupera con fix que permite loops)




- Terminación en el cálculo $\lambda$ tipado: normalización fuerte (SN)

- ¿Por qué demostrar terminación?
  - Correctitud del lenguaje
  - Correspondencia entre la lógica y la programación (paradigma de "propositions as types")

- Técnicas de demostración
  - Semánticas
  - Sintácticas

# Medidas decrecientes

**Definition (Medida decreciente)**

$$\# : \Lambda \to \mathit{WFO} \qquad\qquad M \to_\beta N \implies \#(M) > \#(N)$$

**Una medida decreciente implica SN**

**Motivación** (vs. otras técnicas)

- Aporta más información
- Permite un análisis más profundo (*e.g.* pasos restantes de computación)

**Nuestro trabajo**

- Dos medidas: $\mathcal{W}$ y $\mathcal{T}^m$
- Contribución a la comprensión de por qué los programas del $\lambda^{\to}$ terminan

# La medida de Turing

## Definición

### Redexes y grados

| redex | altura de un tipo | grado de un redex |
|---|---|---|
| $\dots(\lambda x.t)s\dots$ | *e.g.* $\mathsf{h}((\tau \to \tau) \to (\tau \to \tau)) = 2$ | *e.g.* $\delta((\lambda x.x^{\tau\to\tau})t) = 2$ |

### Ejemplo

$$
\begin{array}{lcccccc}
I = \lambda x^{\tau}.x & \quad & \delta(I\,x) & = & \mathsf{h}(\tau \to \tau) & = & 1 \\
K = \lambda x^{\tau}.\lambda y^{\tau}.x & \quad & \delta(K\,(I\,x)) & = & \mathsf{h}(\tau \to \tau \to \tau) & = & 2
\end{array}
$$

$$\underbrace{K\,(\underbrace{I\,x}_{1})}_{2}\,(\underbrace{I\,x}_{1})$$

### Dos observaciones importantes [Turing, 1940s]

- la contracción de un redex **no** puede **crear redexes de igual o mayor grado**
- la contracción de un redex puede **copiar redexes de cualquier grado**
- eligiendo correctamente el redex a contraer podemos dar una **medida decreciente débil**

**Idea**: multiconjunto de los grados de los redexes de $M$

$$\mathcal{T}(M) = [\ d \mid R \text{ es un redex de grado } d \text{ en } M\ ] \qquad [2,1,1]$$

# El cálculo auxiliar sin borrado $\lambda^m$

**Motivación**: definir una medida decreciente a partir de una creciente (bajo WCR y WN)

**Definición**

$$t \;::=\; x \mid \lambda x.t \mid t\,t \mid t\{t\} \qquad (\lambda x.t)s \to_m t[s/x]\{s\} \qquad t\,\underbrace{\{s\{r\}\}\{u\}}_{L} \implies t\mathsf{L}$$

peso de un término: cantidad de memorias $\qquad \mathsf{w}(x\{y\{z\}\}\{w\}) = 3$

**Lemma**

*1.* $\lambda^m$ *satisface preservación de tipos* $\qquad$ *2.* $\lambda^m$ *es confluente*

**Simplificación**

- $\mathsf{S}_D(t)$: contracción simultánea de los redexes $D$
- $\mathsf{S}_*(t)$: iteración $\mathsf{S}_i(t)$ $\quad \mathsf{S}_1(\dots\mathsf{S}_D(t)\dots)$ $\quad$ ($D$ max $\delta$)

**Lemma**

3. $t \to_m^* \mathsf{S}_*(t)$ $\quad$ 4. $\mathsf{S}_*(t)$ *forma normal de* $t$

**Relación de olvido**

$t\{s\} \triangleright t \qquad e.g.\ It \to_m t\{t\} \triangleright t$

**Lemma**

5. $\triangleright$ *conmuta con* $\to_m$

6. $M \to_\beta N$ *implica* $M \to_m s \triangleright N$

# Contando memorias

# La medida $\mathcal{W}$

$\lambda^m$ es **creciente**: $\mathsf{w}(t)$

$$(\lambda x.t)\mathsf{L}s \to_m t[s/x]\{s\}\mathsf{L}$$

**Idea**: la forma normal de $M$ tiene más memorias que la de $N$

**Definition**

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

$$\begin{array}{ccc} M & \xrightarrow{\beta} & N \\ \mathsf{S}_*(M) & \triangleright & \mathsf{S}_*(N) \end{array}$$

$$\mathsf{w}(\mathsf{S}_*(M)) > \mathsf{w}(\mathsf{S}_*(N))$$

**Theorem**

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

# Generalización de la medida de Turing

# Medida de Turing: generalización a cualquier redex

- La medida original requiere elegir el redex correcto
- Un redex puede copiar otros redexes de igual o mayor grado

**Por ejemplo**

- $M \underset{R}{\to_\beta} N$
- $R$ with $\delta(R) = 1$ copies a redex $S$ with $\delta(S) = 2$

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

**Nuestra propuesta**: adaptar la medida para que decrezca ante la contracción de *cualquier redex*

# Un enfoque "naive" $\mathcal{T}'$

**Problema**: un redex puede copiar otros redexes de igual o mayor grado

**Idea**

$i)$ generalizar $\mathcal{T}$ a una familia de medidas indexadas por grado

$$\mathcal{T}'_2(M) = [\underset{\mathsf{S}}{2}, \underset{\mathsf{R}}{1}] \qquad \text{y} \qquad \mathcal{T}'_1(M) = [\underset{\mathsf{R}}{1}]$$

$ii)$ en vez de contar redexes de manera aislada,
consideramos además información sobre los redexes restantes

$$\mathcal{T}'_2(M) = [\ (\underset{\mathsf{S}}{2}, \mathcal{T}'_1(M)),\ (\underset{\mathsf{R}}{1}, [])\ ] \qquad \mathcal{T}'_1(M) = [\ (\underset{\mathsf{R}}{1}, [])\ ]$$

**Definition**

- $\mathcal{T}'_D(M) = [(i, \mathcal{T}'_{i-1}(M)) \mid R$ es un redex de grado $i \leq D$ en $M]$
- $\mathcal{T}'(M) = \mathcal{T}'_D(M)$ donde $D$ es el grado máximo de los redexes en $M$

**No funciona** para los casos en que se copian redexes de igual grado

# La medida $\mathcal{T}^m$

**Información a considerar**

- Un **desarrollo de un conjunto de redexes** es una **secuencia de reducción** donde cada paso corresponde a un **residuo** de un redex **en el conjunto**. Notación: $\rho : M \xrightarrow{D}{}^{*}_{m} M'$
- Un **residuo** es una copia de un redex que aparece luego de la contracción de otro redex

**Idea**

*i)* generalizar $\mathcal{T}$ a una familia de medidas indexadas por grado $\mathcal{T}^m_D$

*ii)* en vez de contar redexes de manera aislada, consideramos
- del conjunto de redexes de grado de $D$
- del final $M'$ de cada desarrollo $\quad \rho : M \xrightarrow{D}{}^{*}_{m} M'$
- el multiconjunto de sus medidas para un grado menor $\quad \mathcal{T}^m_{D-1}(M')$

**Definition**

$$\mathcal{T}^m_D(t) = [\ (i, \mathcal{V}^m_i(t)) \mid R \text{ es un redex de grado } i \leq D \text{ en } t\ ]$$

$$\mathcal{V}^m_D(t) = [\ \mathcal{T}^m_{D-1}(t') \mid \rho : t \xrightarrow{D}{}^{*}_{m} t'\ ]$$

**Theorem**

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

# Conclusiones y trabajo futuro

**Conclusiones**

- Terminación de programas
- Medidas decrecientes
- Cálculo auxiliar sin borrado $\lambda^m$
- Medida $\mathcal{W}$: basada en el peso (o memoria acumulada) de los términos en $\lambda^m$
- Medida $\mathcal{T}^m$: basada en multiconjuntos anidados de medidas de resultados de desarrollos

**Trabajo futuro**

- Extender las medidas a System F ($\lambda^{\rightarrow}$ con polimorfismo)
- Formalizar las demostraciones en un asistente de pruebas

# The auxiliar $\lambda^m-$calculus

## Motivation

$\beta$ is erasing $\qquad\qquad (\lambda x.y)t \to_\beta y$

### A motivation not to erase

- Klop-Nederpelt lemma $\quad \mathit{INC} \wedge \mathit{WCR} \wedge \mathit{WN} \implies \mathit{SN} \wedge \mathit{CR}$
- We can obtain a decreasing measure from $\mathit{INC} \wedge \mathit{WCR} \wedge \mathit{WN}$
  - 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 $\mathit{inc}(t) < \mathit{inc}(s) < \mathit{inc}(v)$
  - $\mathit{dec}(t) = \mathit{inc}(v) - \mathit{inc}(t)$

$$
\begin{tikzcd}
t \arrow[rr, "\mathit{WN}"] \arrow[dr, "S"'] & & v \\
& s \arrow[ur, "\mathit{WCR}"'] &
\end{tikzcd}
$$

# Turing's measure "failing" example

## Example: copying a redex of greater degree

$$
\begin{aligned}
I_1 &= \lambda x^{\tau}.x & \qquad \delta(I_1\,x) &= \mathsf{h}(\tau \to \tau) &&= 1 \\
I_2 &= \lambda x^{\tau\to\tau}.x & \delta(I_2\,I_1) &= \mathsf{h}((\tau \to \tau) \to (\tau \to \tau)) &&= 2 \\
K &= \lambda x^{\tau}.\lambda y^{\tau}.x & \delta(K\,\_) &= \mathsf{h}(\tau \to \tau \to \tau) &&= 2 \\
S_{K\,I} &= \lambda x^{\tau}.K\,x\,(I_1\,x) & \delta(S_{K\,I}\,\_) &= \mathsf{h}(\tau \to \tau) &&= 1
\end{aligned}
$$

$$
\mathcal{T}(\underset{\text{R1}}{\underline{S_{\underset{\text{S2}}{\underline{K}}\;\underset{\text{T1}}{\underline{I}}}\,(\underset{\text{U2}}{\underline{I_2\,I_1}}\,x)}}) = \{\underset{\text{S}}{2}, \underset{\text{U}}{2}, \underset{\text{R}}{1}, \underset{\text{T}}{1}\}
\qquad
\mathcal{T}(K\,\underset{\text{S2}}{\underline{(\underset{\text{U}'2}{\underline{I_2\,I_1}}\,x)}}\,\underset{\text{T1}}{\underline{(I_1\,(\underset{\text{U}'2}{\underline{I_2\,I_1}}\,x))}}) = \{\underset{\text{S}}{2}, \underset{\text{U}'}{2}, \underset{\text{U}'}{2}, \underset{\text{T}}{1}\}
$$

# A first attempt: $\mathcal{T}'$ measure

*A working? example* $(>)$

**Definition**

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

**Example**

$$M \quad = \quad \underset{\color{red}{\mathsf{R1}}}{\underline{S_{\underset{\mathsf{S2}}{\underline{K}}}\,\underset{\mathsf{T1}}{\underline{I}}\,(\underset{\mathsf{U2}}{\underline{I_2\,I_1}}\,x)}} \quad \longrightarrow_\beta \quad \underset{\mathsf{S2}}{\underline{K\,(\underset{\mathsf{U'2}}{\underline{I_2\,I_1}}\,x)}}\,\underset{\mathsf{T1}}{\underline{(I_1\,(\underset{\mathsf{U''2}}{\underline{I_2\,I_1}}\,x))}} \quad = \quad N$$

$$\mathcal{T}'_2(M) = [\ (\underset{\mathsf{S}}{2}, \mathcal{T}'_1(M)),\ (\underset{\mathsf{U}}{2}, \mathcal{T}'_1(M)),\ (\underset{\mathsf{R}}{1}, []),\ (\underset{\mathsf{T}}{1}, [])\ ] \qquad \mathcal{T}'_1(M) = [\ (\underset{\mathsf{R}}{1}, []),\ (\underset{\mathsf{T}}{1}, [])\ ]$$

$$\mathcal{T}'_2(N) = [\ (\underset{\mathsf{S}}{2}, \mathcal{T}'_1(M)),\ (\underset{\mathsf{U'}}{2}, \mathcal{T}'_1(M)),\ (\underset{\mathsf{U''}}{2}, \mathcal{T}'_1(M)),\ (\underset{\mathsf{T}}{1}, [])\ ] \qquad \mathcal{T}'_1(N) = [\ (\underset{\mathsf{T}}{1}, [])\ ]$$

$$(2, [\ (1, []),\ \ (1, [])\ ]) \qquad > \qquad (2, [\ (1, [])\ ])$$

# A first attempt: $\mathcal{T}'$ measure

*A failing example* $(=)$

**Definition**

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

**Example Example**

$$M \quad = \quad \underset{\mathsf{R1}}{\underline{S_{\underset{\mathsf{S2}}{\underline{K}}\,\underset{\mathsf{T1}}{\underline{I}}}\;\underset{\mathsf{U1}}{\underline{(I_1\,x)}}}} \quad \longrightarrow_\beta \quad \underset{\mathsf{S2}}{\underline{K\,\underset{\mathsf{U'1}}{\underline{(I_1\,x)}}}}\;\underset{\mathsf{T1}}{\underline{(\underset{\mathsf{U''1}}{\underline{(I_1\,x)}})}} \quad = \quad N$$

$$\mathcal{T}'_2(M) = [\ (\underset{\mathsf{S}}{2}, \mathcal{T}'_1(M)),\ (\underset{\mathsf{R}}{1}, []),\ (\underset{\mathsf{T}}{1}, []),\ (\underset{\mathsf{U}}{1}, []),\ ] \qquad \mathcal{T}'_1(M) = [\ (\underset{\mathsf{R}}{1}, []),\ (\underset{\mathsf{T}}{1}, []),\ (\underset{\mathsf{U}}{1}, []),\ ]$$

$$\mathcal{T}'_2(N) = [\ (\underset{\mathsf{S}}{2}, \mathcal{T}'_1(M)),\ (\underset{\mathsf{T}}{1}, []),\ (\underset{\mathsf{U'}}{1}, []),\ (\underset{\mathsf{U''}}{1}, [])\ ] \qquad \mathcal{T}'_1(N) = [\ (\underset{\mathsf{T}}{1}, []),\ (\underset{\mathsf{U'}}{1}, []),\ (\underset{\mathsf{U''}}{1}, [])\ ]$$

$$(2, [\ (1, []),\ (1, []),\ (1, [])\ ]) \qquad = \qquad (2, [\ (1, []),\ (1, []),\ (1, [])\ ])$$

# A second attempt: $\mathcal{T}^\beta$ measure

**Definition**

$$\mathcal{T}^\beta_D(M) = [\ (i, \mathcal{V}^\beta_i(M)) \mid R \text{ is a redex of degree } i \leq D \text{ in } M\ ]$$

$$\mathcal{V}^\beta_D(M) = [\ \mathcal{T}^\beta_{D-1}(M') \mid \rho : M \xrightarrow{D}{}^{*}_{\beta} M'\ ]$$

**Reasoning about the auxiliar measure $\mathcal{V}^\beta_D$**

Consider

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

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

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

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

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