Two Decreasing Measures for Simply Typed λ-Terms

Pablo Barenbaum ✉
ICC, Universidad de Buenos Aires, Argentina
Universidad Nacional de Quilmes (CONICET), Buenos Aires, Argentina

Cristian Sottile ✉
ICC, Universidad de Buenos Aires (CONICET), Argentina
Universidad Nacional de Quilmes, Buenos Aires, Argentina

Abstract

This paper defines two decreasing measures for terms of the simply typed λ-calculus, called the $\mathcal{W}$-measure and the $\mathcal{T}^{\mathbf{m}}$-measure. A decreasing measure is a function that maps each typable λ-term to an element of a well-founded ordering, in such a way that contracting any β-redex decreases the value of the function, entailing strong normalization. Both measures are defined constructively, relying on an auxiliary calculus, a non-erasing variant of the λ-calculus. In this system, dubbed the $\lambda^{\mathbf{m}}$-calculus, each β-step creates a "wrapper" containing a copy of the argument that cannot be erased and cannot interact with the context in any other way. Both measures rely crucially on the observation, known to Turing and Prawitz, that contracting a redex cannot create redexes of higher degree, where the degree of a redex is defined as the height of the type of its λ-abstraction. The $\mathcal{W}$-measure maps each λ-term to a natural number, and it is obtained by evaluating the term in the $\lambda^{\mathbf{m}}$-calculus and counting the number of remaining wrappers. The $\mathcal{T}^{\mathbf{m}}$-measure maps each λ-term to a structure of nested multisets, where the nesting depth is proportional to the maximum redex degree.

2012 ACM Subject Classification Theory of computation → Equational logic and rewriting; Theory of computation → Lambda calculus

Keywords and phrases Lambda Calculus, Rewriting, Termination, Strong Normalization, Simple Types

Digital Object Identifier 10.4230/LIPIcs.FSCD.2023.11

Funding This work was partially supported by project grants PICT-2021-I-A-00090, PICT-2021-I-INVI-00602, PIP 11220200100368CO, PICT 2019-1272, PUNQ 2218/22, and PUNQ 2219/22.

Acknowledgements To Giulio Manzonetto for fruitful discussions that led to the development of this work. To Eduardo Bonelli and the anonymous reviewers for feedback on earlier versions of this paper. The second author would like to thank his advisors Alejandro Díaz-Caro and Pablo E. Martínez López.

1 Introduction

In this paper we revisit a fundamental question, that of **strong normalization** of the simply typed λ-calculus (STLC). We begin by recalling that a reduction relation is *weakly normalizing* (WN) if every term can be reduced to normal form in a finite number of steps, whereas it is *strongly normalizing* (SN) if there are no infinite reduction sequences ($a_1 \to a_2 \to a_3 \to \ldots$). Let us review three proof techniques for proving strong normalization of the STLC.

One of the better known ways to prove that the STLC is SN is through arguments based on **reducibility models**. The idea is to interpret each type A as a set $[\![A]\!]$ of strongly normalizing terms, and to prove that each term M of type A is an element of $[\![A]\!]$.

Many variants of these ideas can be found in the literature, including Girard's reducibility candidates [17] and Tait's saturated sets [30]. These techniques provide relatively succint proofs and they generalize well to extensions of the STLC, *e.g.* to dependent type theory [6] or classical calculi [13]. On the other hand, the abstract nature of reducibility arguments does not provide a "tangible" insight on why a β-reduction step brings a term closer to normal form. More specifically, reducibility arguments do not construct explicit **decreasing measures**. By decreasing measure we mean a function "#" mapping each λ-term to a well-founded ordering $(X, >)$ such that $M \to_\beta N$ implies $\#(M) > \#(N)$.

Another way to prove strong normalization is based on **redex degrees**. A *redex* in the STLC is an applied abstraction, *i.e.* a term of the form $(\lambda x.\, M)\, N$. The *degree* of a redex is defined as the *height* of the type of its abstraction. A crucial observation, that can be attributed to an unpublished note of Turing (as reported by Gandy [15]; see also [4]), is that *contracting a redex cannot create a redex of higher or equal degree.* Recall that a redex S is *created* by the contraction of a redex R if S has no *ancestor* before R. Indeed, as shown by Lévy [22], in the λ-calculus, redexes can be created in exactly one of the three ways below:

1 $\quad (\underline{\lambda} x.\, x)\, (\lambda y.\, M)\, N \quad \to_\beta \quad (\underline{\lambda} y.\, M)\, N$

2 $\quad (\underline{\lambda} x.\, \lambda y.\, M)\, N\, P \quad \to_\beta \quad (\underline{\lambda} y.\, M[x := N])\, P$

3 $\quad (\underline{\lambda} x.\, \ldots x\, M \ldots)\, (\lambda y.\, N) \quad \to_\beta \quad \ldots (\underline{\lambda} y.\, N)\, M[x := \lambda y.\, N] \ldots$

where we underline the λ of the contracted redex on the left, and the λ of the created redex on the right. In each of these cases, it can be seen that the degree of the created redex is strictly lower than the degree of the contracted redex. For instance, in creation case **1**, the type of the contracted redex is of the form $(A \to B) \to (A \to B)$, while the type of the created redex is $A \to B$, so the height strictly decreases.

With this fact in mind, for each term M one can define what we call **Turing's measure**, *i.e.* the multiset $\mathcal{T}(M)$ of the degrees of all the redexes of M. One may hope that any reduction step $M \to_\beta N$ decreases the measure, *i.e.* $\mathcal{T}(M) \succ \mathcal{T}(N)$, where "$\succ$" is the usual well-founded multiset ordering induced by the ordering $(\mathbb{N}, >)$ of its elements [12]. Unfortunately, this is not the case: even though contracting a redex can only create redexes of strictly lower degree, it can still make an arbitrary number of *copies* of redexes of arbitrarily large degrees.

In his notes, Turing observed that one can follow a reduction strategy that always selects the *rightmost* redex of highest degree. This strategy ensures that the contracted redex does not copy redexes of higher or equal degree, which makes the $\mathcal{T}(-)$ measure strictly decrease, thus proving that the λ-calculus is WN. An even simpler measure that also decreases, if one follows this strategy, is $\mathcal{T}'(M) = (D, n)$, where D is the maximum degree of the redexes in M and n is the number of redexes of degree D in M. Similar ideas were exploited by Prawitz [28] and Gentzen (as reported by von Plato [27]) to normalize proofs in natural deduction. After WN has been established, an indirect proof of SN can be obtained by translating each typable λ-term M to a typable term M' of the *λI-calculus*; see for instance [29, Section 3.5].

In summary, redex degrees can be used to define concrete measures such as $\mathcal{T}(M)$ and $\mathcal{T}'(M)$, that are computable in linear time and decrease when following a particular reduction strategy. As already mentioned, these measures do not necessarily decrease when contracting arbitrary β-redexes.

A third way to prove SN relies on an interpretation that maps terms to **increasing functionals**. This approach was pioneered by Gandy [16] and refined by de Vrijer [10]. Each type A is mapped to a partially ordered set $[\![A]\!]$. Specifically, base types are mapped to $(\mathbb{N}, \leq)$, and $[\![A \to B]\!]$ is defined as the set of strictly increasing functions $[\![A]\!] \to [\![B]\!]$, partially ordered by the point-wise order. Each term M of type A is interpreted as an element $[M] \in [\![A]\!]$. Moreover, an element $f \in [\![A]\!]$ can be projected to a natural number $f\star \in \mathbb{N}$ in such a way that $M \to_\beta N$ implies $[M]\star > [N]\star$. This indeed provides a decreasing measure. One of the downsides of this measure is that computing $[M]\star$ is essentially as difficult as evaluating M, because $[M]$ is defined as a higher-order functional with a similar structure as the λ-term M itself.

In this work we propose **two decreasing measures for the STLC**, that we dub the $\mathcal{W}$-measure and the $\mathcal{T}^{\mathbf{m}}$-measure, and we prove that they are decreasing. An ideal decreasing measure should fulfill multiple (partly subjective) requirements: **1.** the measure should be easy to calculate, in terms of computational complexity; **2.** its codomain (a well-founded ordering) should be simple, in terms of its ordinal type; **3.** it should give us insight on why β-reduction terminates; **4.** it should be easy to prove that the measure is decreasing. A measure that excels simultaneously at all these requirements is elusive, and perhaps unattainable. The proposed measures have different strengths and weaknesses.

Contributions and structure of this document. The $\mathcal{W}$-measure and the $\mathcal{T}^{\mathbf{m}}$-measure are defined by means of on an auxiliary calculus that we dub the $\lambda^{\mathbf{m}}$-calculus. The remainder of the paper is structured as follows.

In **Section 2** we **define the $\lambda^{\mathbf{m}}$-calculus**. It is an extension of the STLC with terms[1] of the form $t\{s\}$, called *wrappers*. A wrapper $t\{s\}$ should be understood as essentially the term t in which s is a *memorized term*, that is, leftover garbage that can be reduced but cannot interact with the context in any way. The type of $t\{s\}$ is the same as the type of t, disregarding the type of s.

The β-reduction rule is modified so that contracting a redex $(\lambda x.\, t)\, s$, besides substituting the free occurrences of x by s in t, produces a wrapper that contains a copy of the argument s. The reduction rule is $(\lambda x.\, t)\{u_1\}\ldots\{u_n\}\, s \ \to_{\mathbf{m}}\ t[x := s]\{s\}\{u_1\}\ldots\{u_n\}$. Note that we allow the presence of an arbitrary number of memorized terms mediating between the abstraction and the application. This is to avoid memorized terms *blocking* redexes. For example, if $I = \lambda x.\, x$:

$$(\lambda x.\, x(xy))I \to_{\mathbf{m}} (I(Iy))\{I\} \to_{\mathbf{m}} (Iy)\{Iy\}\{I\} \to_{\mathbf{m}} (Iy)\{y\{y\}\}\{I\} \to_{\mathbf{m}} y\{y\}\{y\{y\}\}\{I\}$$

Then we study some syntactic properties of $\lambda^{\mathbf{m}}$. In particular, we define a relation $t \rhd s$ of *forgetful reduction*, meaning that s is obtained from t by erasing one memorized subterm. For example, $x\,\{x\{y\}\}\{y\{z\}\} \rhd x\,\{y\{z\}\}$. Forgetful reduction is used as a technical tool to prove that the measures are decreasing in the following sections.

In **Section 3**, we **propose the $\mathcal{W}$-measure** (Def. 12), and we prove that it is decreasing. To define the $\mathcal{W}$-measure, we resort to an operation $\mathtt{S}_d(t)$ that simultaneously contracts all the redexes of degree d in a term of the $\lambda^{\mathbf{m}}$-calculus, that is, the result of the *complete development* of all the redexes of degree d. The degree of a redex $(\lambda x.\, t)\{u_1\}\ldots\{u_n\}\, s$ is defined similarly as for the STLC, as the height of the type of the abstraction. To calculate the $\mathcal{W}$-measure of a λ-term M, let D be the maximum degree of the redexes in M, and define $\mathcal{W}(M)$ as the number of wrappers in $\mathtt{S}_1(\mathtt{S}_2(\ldots\mathtt{S}_D(M)))$. For example, if $M = (\lambda x.\, x\,(x\,y))\,(\lambda z.\, w)$, it turns out that $\mathtt{S}_1(\mathtt{S}_2(M)) = w\{w\{y\}\}\{\lambda z.\, w\}$ which has three

[1] Note that terms of the $\lambda^{\mathbf{m}}$-calculus are ranged over by $t, s, \ldots$ (rather than $M, N, \ldots$).

wrappers, so $\mathcal{W}(M) = 3$. The $\mathcal{W}$-measure maps each typable λ-term to a natural number. The main result of Section 3 is Thm. 15, stating that $\mathcal{W}$ **is decreasing**, *i.e.* that $M \to_\beta N$ implies $\mathcal{W}(M) > \mathcal{W}(N)$.

In **Section 4** we study **reduction by degrees**, a restricted notion of reduction in the $\lambda^{\mathbf{m}}$-calculus, written $t \xrightarrow{d}_{\mathbf{m}} s$, meaning that t reduces to s by contracting a redex of degree d. This section contains technical commutation, termination, and postponement results.

In **Section 5**, we **propose the $\mathcal{T}^{\mathbf{m}}$-measure**, and we prove that it is decreasing. To define the $\mathcal{T}^{\mathbf{m}}$-measure, we define two auxiliary measures $\mathcal{T}^{\mathbf{m}}_{\leq D}(t)$ and $\mathcal{R}^{\mathbf{m}}_{D}(t)$, indexed by a natural number $D \in \mathbb{N}_0$, mutually recursively:

- $\mathcal{T}^{\mathbf{m}}_{\leq D}(t)$ is the multiset of pairs $(d, \mathcal{R}^{\mathbf{m}}_{d}(t))$, for each redex occurrence of degree $d \leq D$ in t;
- $\mathcal{R}^{\mathbf{m}}_{D}(t)$ is the multiset of elements $\mathcal{T}^{\mathbf{m}}_{\leq D-1}(t')$, for each reduction sequence $t \xrightarrow{D}^{*}_{\mathbf{m}} t'$.

The measure $\mathcal{T}^{\mathbf{m}}_{\leq D}(t)$ is defined for every $D \geq 0$, while $\mathcal{R}^{\mathbf{m}}_{D}(t)$ is defined only for $D \geq 1$. Multisets are ordered according to the usual multiset ordering, and pairs according to the lexicographic ordering. To calculate the $\mathcal{T}^{\mathbf{m}}$-measure of a λ-term M, let D be the maximum degree of the redexes in M, and define $\mathcal{T}^{\mathbf{m}}_{\leq}(M) \overset{\text{def}}{=} \mathcal{T}^{\mathbf{m}}_{\leq D}(M)$. The measure $\mathcal{T}^{\mathbf{m}}_{\leq}(M)$ yields a structure of nested multisets of nesting depth at most $2D$. The main theorem of Section 3 is Thm. 32, stating that $\mathcal{T}^{\mathbf{m}}$ **is decreasing**, *i.e.* that $M \to_\beta N$ implies $\mathcal{T}^{\mathbf{m}}_{\leq}(M) > \mathcal{T}^{\mathbf{m}}_{\leq}(N)$.

Finally, in **Section 6**, we conclude.

2 The $\lambda^{\mathbf{m}}$-calculus

As mentioned in the introduction, the $\lambda^{\mathbf{m}}$-calculus is an extension of the STLC in which the β-reduction rule keeps an extra memorized copy of the argument in a "wrapper" $t\{\!\!\{s\}\!\!\}$, in such a way that contracting a redex like $(\lambda x.\, t)\, s$ does not erase s, even if x does not occur free in t. In this section we define the $\lambda^{\mathbf{m}}$-calculus and we prove some of the properties that are needed in the following sections to prove that the $\mathcal{W}$-measure and the $\mathcal{T}^{\mathbf{m}}$-measure are decreasing. In particular, we discuss *subject reduction* (Prop. 3) and *confluence* (Prop. 4); we define an operation of **simplification** (Def. 5) which turns out to calculate the normal form of a term (Prop. 7); and we define the relation called **forgetful reduction** (Def. 8), which is shown to commute with reduction (Prop. 10).

First we fix the notation and nomenclature. *Types* of the STLC are either base types $(\alpha, \beta, \ldots)$ or arrow types $(A \to B)$. *Terms* are either variables $(x^A, y^A, \ldots)$, abstractions $(\lambda x^A.\, M)$, or applications $(M\, N)$, with the usual typing rules. Terms are defined up to α-renaming of bound variables. We adopt an *à la* Church presentation of the STLC, but we omit most type decorations on variables as long as there is little danger of confusion. The β-reduction rule is $(\lambda x.\, M)\, N \to_\beta M[x := N]$ where $M[x := N]$ is the capture-avoiding substitution of the free occurrences of x in M by N.

The $\lambda^{\mathbf{m}}$-calculus: syntax and reduction. The set of $\lambda^{\mathbf{m}}$*-terms* – or just *terms* – is given by $t, s, \ldots ::= x^A \mid \lambda x^A.\, t \mid t\, s \mid t\{\!\!\{s\}\!\!\}$. The four kinds of terms are respectively called *variables*, *abstractions*, *applications*, and *wrappers*. In a wrapper $t\{\!\!\{s\}\!\!\}$, the subterm t is called the *body* and s is called the *memorized term*. As in the STLC, we usually omit type annotations and terms are regarded up to α-renaming. A *context* is a term $\mathtt{C}$ with a single free occurrence of a distinguished variable $\Box$, and $\mathtt{C}[t]$ is the variable-capturing substitution of the occurrence of $\Box$ in $\mathtt{C}$ by t.

Typing judgments are of the form $\Gamma \vdash t : A$ where Γ is a partial function mapping variables to types. Derivable typing judgments are defined by the following rules:

$$\frac{}{\Gamma, x : A \vdash x^A : A} \qquad \frac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x^A . t : A \to B} \qquad \frac{\Gamma \vdash t : A \to B \quad \Gamma \vdash s : A}{\Gamma \vdash t\, s : B} \qquad \frac{\Gamma \vdash t : A \quad \Gamma \vdash s : B}{\Gamma \vdash t\{s\} : A}$$

A term t is *typable* if $\Gamma \vdash t : A$ holds for some Γ and some A. Unless otherwise specified, when we speak of "terms" we mean "typable terms". It is straightforward to show that a typable term has a unique type. We write $\mathsf{type}(t)$ for the type of t.

A *memory*, written $\mathtt{L}$, is a list of memorized terms, given by the grammar $\mathtt{L} ::= \Box \mid \mathtt{L}\{t\}$. If t is a term and $\mathtt{L}$ is a memory, we write $t\mathtt{L}$ for the term that results from appending all the memorized terms in $\mathtt{L}$ to t, that is, $(t)(\Box\{s_1\} \ldots \{s_n\}) = t\{s_1\} \ldots \{s_n\}$. We write $t[x := s]$ for the operation of capture-avoiding substitution of the free occurrences of x in t by s. The $\lambda^{\mathbf{m}}$-calculus is the rewriting system whose objects are typable $\lambda^{\mathbf{m}}$-terms, endowed with the following notion of reduction, closed by compatibility under arbitrary contexts:

▶ **Definition 1** (Reduction in the $\lambda^{\mathbf{m}}$-calculus). $(\lambda x. t)\mathtt{L}\, s \;\to_{\mathbf{m}}\; t[x := s]\{s\}\mathtt{L}$

Abstractions followed by a memory, *i.e.* terms of the form $(\lambda x. t)\mathtt{L}$, are called $\mathbf{m}$-*abstractions*. Note that all abstractions are also $\mathbf{m}$-abstractions, as $\mathtt{L}$ may be empty. A *redex* is an expression matching the left-hand side of the $\to_{\mathbf{m}}$-reduction rule, which must be an *applied* $\mathbf{m}$-*abstraction*, *i.e.* a term of the form $(\lambda x. t)\mathtt{L}\, s$. The *height* of a type is given by $\mathsf{h}(\alpha) \stackrel{\text{def}}{=} 0$ and $\mathsf{h}(A \to B) \stackrel{\text{def}}{=} 1 + \max(\mathsf{h}(A), \mathsf{h}(B))$. The *degree* of a $\mathbf{m}$-abstraction $(\lambda x. t)\mathtt{L}$ is defined as the height of its type; note that this number is always strictly positive, since the type must be of the form $A \to B$. Moreover, this type is unique, so the operation is well-defined. The *degree* of a redex $(\lambda x. t)\mathtt{L}\, s$ is defined as the degree of the $\mathbf{m}$-abstraction $(\lambda x. t)\mathtt{L}$. The *max-degree* of a term t is written $\mathsf{maxdeg}(t)$ and it is defined as the maximum degree of the redexes in t, or 0 if t has no redexes. The *weight* $\mathsf{w}(t)$ of a $\lambda^{\mathbf{m}}$-term t is the number of wrappers in t.

▶ **Example 2.** Let 0 be a base type and let $t := (\lambda x^{0\to 0}. \lambda y^0. y^0\{x^{0\to 0}\,(x^{0\to 0}\, z^0)\})\, I\, w^0$, where $I := \lambda x^0. x^0$. One possible way to reduce t is:

$$\begin{aligned} (\lambda x. \lambda y. y\{x\,(x\,z)\})\, I\, w \;&\to_{\mathbf{m}}\; (\lambda y. y\{I\,(I\,z)\})\{I\}\, w \;\to_{\mathbf{m}}\; w\{I\,(I\,z)\}\{w\}\{I\} \\ &\to_{\mathbf{m}}\; w\{I\,(z\{z\})\}\,\{w\}\,\{I\} \;\to_{\mathbf{m}}\; w\{z\{z\}\{z\{z\}\}\}\{w\}\{I\} = s \end{aligned}$$

The degrees of the redexes contracted in each step are 2, 1, 1, and 1, in that order. Note that $\mathsf{maxdeg}(t) = 2$ and that the weight of the resulting term is $\mathsf{w}(s) = 6$.

Two basic properties of the $\lambda^{\mathbf{m}}$-calculus are subject reduction and confluence. These are immediate consequences of the fact that the $\lambda^{\mathbf{m}}$-calculus can be understood as an *orthogonal HRS* in the sense of Nipkow [26], *i.e.* a left-linear higher-order rewriting system without critical pairs.

▶ **Proposition 3** (Subject reduction). *Let* $\Gamma \vdash t : A$ *and* $t \to_{\mathbf{m}} s$. *Then* $\Gamma \vdash s : A$.

▶ **Proposition 4** (Confluence). *If* $t_1 \to_{\mathbf{m}}^* t_2$ *and* $t_1 \to_{\mathbf{m}}^* t_3$*, there exists a term* t_4 *such that* $t_2 \to_{\mathbf{m}}^* t_4$ *and* $t_3 \to_{\mathbf{m}}^* t_4$.

Full simplification. Next, we define an operation written $\mathtt{S}_*(t)$ and called *full simplification*.

Let $d \geq 1$ be a natural number. The *simplification of degree d*, written $\mathtt{S}_d(t)$, is the result of simultaneously contracting all the redexes of degree d in t, that is, the result of the *complete development* of all redexes of degree d. Formally, for each $\lambda^{\mathbf{m}}$-term t we define $\mathtt{S}_d(t)$, and, for each memory $\mathtt{L}$, we define $\mathtt{S}_d(\mathtt{L})$ as follows:

▶ **Definition 5** (Simplification).

$$\begin{aligned}
\mathtt{S}_d(x) &\stackrel{\text{def}}{=} x \\
\mathtt{S}_d(\lambda x.\, t) &\stackrel{\text{def}}{=} \lambda x.\, \mathtt{S}_d(t) \\
\mathtt{S}_d(t\, s) &\stackrel{\text{def}}{=} \begin{cases} \mathtt{S}_d(t')[x := \mathtt{S}_d(s)]\{\mathtt{S}_d(s)\}\mathtt{S}_d(\mathtt{L}) & \textit{if } t = (\lambda x.\, t')\mathtt{L} \textit{ and it is of degree } d \\ \mathtt{S}_d(t)\, \mathtt{S}_d(s) & \textit{otherwise} \end{cases} \\
\mathtt{S}_d(t\{s\}) &\stackrel{\text{def}}{=} \mathtt{S}_d(t)\{\mathtt{S}_d(s)\}
\end{aligned}$$

where if $\mathtt{L}$ is a memory, $\mathtt{S}_d(\mathtt{L})$ is defined by $\mathtt{S}_d(\Box) \stackrel{\text{def}}{=} \Box$ and $\mathtt{S}_d(\mathtt{L}\{t\}) \stackrel{\text{def}}{=} \mathtt{S}_d(\mathtt{L})\{\mathtt{S}_d(t)\}$. Furthermore, if t is a $\lambda^{\mathbf{m}}$-term of max-degree D, we define the *full simplification* of t as the term that results from iteratively taking the simplification of degree i from D down to 1. More precisely, $\mathtt{S}_*(t) \stackrel{\text{def}}{=} \mathtt{S}_1(\dots \mathtt{S}_{D-1}(\mathtt{S}_D(t)))$.

▶ **Example 6.** Consider the λ-term $M = (\lambda x^{0\to0}.\, x^{0\to0}(x^{0\to0}\, y^0))(\lambda z^0.\, w^0)$. It can be regarded also as a $\lambda^{\mathbf{m}}$-term, and we have:

$$\begin{aligned}
\mathtt{S}_2(M) &= ((\lambda z^0.\, w^0)\, ((\lambda z^0.\, w^0)\, y^0))\{\lambda z^0.\, w^0\} \\
\mathtt{S}_*(M) = \mathtt{S}_1(\mathtt{S}_2(M)) &= w^0\{w^0\{y^0\}\}\{\lambda z^0.\, w^0\}
\end{aligned}$$

Note that M has only one redex, whose abstraction is of type $(0 \to 0) \to 0$ and hence of degree 2, and that $\mathtt{S}_2(M)$ has two redexes, whose abstractions are of type $0 \to 0$ and hence of degree 1. Moreover, consider the λ-term $N = (\lambda z^0.\, w^0)\, ((\lambda z^0.\, w^0)\, y^0)$. Then $\mathtt{S}_*(N) = \mathtt{S}_1(N) = w\{w\{y\}\}$. Note that N has two redexes whose abstraction is of type $0 \to 0$ and hence of degree 1. As an additional note, in the λ-calculus there is a reduction step $M \to_\beta N$, and we have that $\mathsf{w}(\mathtt{S}_*(M)) = 3 > 2 = \mathsf{w}(\mathtt{S}_*(N))$. So this example illustrates that the $\mathcal{W}$-measure (as defined in Def. 12) is decreasing (as we will show in Thm. 15).

As it turns out, **full simplification corresponds to reduction to normal form**. More precisely, we have the following result, which entails in particular that the $\lambda^{\mathbf{m}}$-calculus is weakly normalizing:

▶ **Proposition 7.** *$t \to^*_{\mathbf{m}} \mathtt{S}_*(t)$, and moreover $\mathtt{S}_*(t)$ is a $\to_{\mathbf{m}}$-normal form.*

Proof. To show that $t \to^*_{\mathbf{m}} \mathtt{S}_*(t)$, it suffices to prove a lemma stating that $t \to^*_{\mathbf{m}} \mathtt{S}_d(t)$ for all $d \geq 1$. This implies that $t \to^*_{\mathbf{m}} \mathtt{S}_D(t) \to^*_{\mathbf{m}} \mathtt{S}_{D-1}(\mathtt{S}_D(t)) \dots \to^*_{\mathbf{m}} \mathtt{S}_1(\dots \mathtt{S}_{D-1}(\mathtt{S}_D(t))) = \mathtt{S}_*(t)$, where D is the max-degree of t. The lemma itself is straightforward by induction on t.

To show that $\mathtt{S}_*(t)$ is a $\to_{\mathbf{m}}$-normal form, the key property is that, after performing a simplification of order d, no redexes of order d remain. The reason is that contracting a redex of order d can only create redexes of lower degree. More precisely, we prove a key lemma stating that if $d \geq 1$ and $\mathsf{maxdeg}(t) \leq d$, then $\mathsf{maxdeg}(\mathtt{S}_d(t)) < d$. If we let $\mathsf{maxdeg}(t) \leq D$, we can iterate this lemma, to obtain that $\mathsf{maxdeg}(\mathtt{S}_D(t)) < D$, and $\mathsf{maxdeg}(\mathtt{S}_{D-1}(\mathtt{S}_D(t))) < D - 1$, ..., and finally $\mathsf{maxdeg}(\mathtt{S}_1(\dots \mathtt{S}_{D-1}(\mathtt{S}_D(t)))) < 1$. This means that $\mathtt{S}_*(t) = \mathtt{S}_1(\dots \mathtt{S}_{D-1}(\mathtt{S}_D(t)))$ does not contain redexes, since there are no redexes of degree 0, so $\mathtt{S}_*(t)$ must be a $\to_{\mathbf{m}}$-normal form. ◀

Forgetful reduction. To conclude this section, we introduce the relation of *forgetful reduction* $t \rhd^+ s$, and we prove that it commutes with reduction.

▶ **Definition 8.** *A $\lambda^{\mathbf{m}}$-term t* reduces via a forgetful step *to s, written $t \rhd s$, according to the following axiom, closed by compatibility under arbitrary contexts:*

$$t\{s\} \rhd t$$

We say that t reduces via forgetful reduction *to* s *if and only if* $t \rhd^+ s$*, where* $\rhd^+$ *denotes the transitive closure of* $\rhd$*.*

▶ **Example 9.** $(\lambda x.\, x\{y\{y\}\})\{z\{z\}\} \rhd (\lambda x.\, x\{y\{y\}\})\{z\} \rhd (\lambda x.\, x)\{z\} \rhd \lambda x.\, x.$

▶ **Proposition 10** (Forgetful reduction commutes with reduction)**.** *If* $t \rhd^+ s$ *and* $t \to_{\mathbf{m}}^* t'$*, there exists a term* s' *such that* $t' \rhd^+ s'$ *and* $s \to_{\mathbf{m}}^* s'$*. Furthermore, if* $t \rhd^+ s$ *and* t *is a* $\to_{\mathbf{m}}$*-normal form, then* s *is also a normal form.*

Proof. The result can be reduced to a local commutation result, stating that if $t \rhd s$ and $t \to_{\mathbf{m}} t'$, there exists a term s' such that $t' \rhd^+ s'$ and $s \to_{\mathbf{m}}^= s'$, where $\to_{\mathbf{m}}^=$ is the reflexive closure of $\to_{\mathbf{m}}$. Local commutation can be proved by case analysis. The interesting cases are when a shrinking step $s \rhd s'$ lies inside the argument of a redex, and when a reduction step $r \to_{\mathbf{m}} r'$ is inside erased garbage:

$$\begin{array}{ccc}
(\lambda x.\, t)\mathtt{L}\, s & \rhd & (\lambda x.\, t)\mathtt{L}\, s' \\
\downarrow & & \downarrow \\
t[x := s]\{s\}\mathtt{L} & \rhd^+ & t[x := s']\{s'\}\mathtt{L}
\end{array}
\qquad\qquad
\begin{array}{ccc}
u\{r\} & \rhd & u \\
\downarrow & & \| \\
u\{r'\} & \rhd^+ & u
\end{array}$$

For the last part of the statement, it suffices to show that if $t \rhd s$ in one step and t is a $\to_{\mathbf{m}}$-normal form, then s is also a normal form, which is straightforward by induction on t. ◀

Each step in the STLC has a *corresponding* step in the $\lambda^{\mathbf{m}}$-calculus, that contracts the redex in the same position. For instance the step $(\lambda x.\, x\, y)\, I \to_\beta I\, y$ in the STLC has a corresponding step $(\lambda x.\, x\, y)\, I \to_{\mathbf{m}} (I\, y)\{I\}$ in the $\lambda^{\mathbf{m}}$-calculus. In this example, $(I\, y)\{I\} \rhd I\, y$. The following easy lemma confirms that this is a general fact:

▶ **Lemma 11** (Reduce/forget lemma)**.** *Let* $M \to_\beta N$ *be a* β*-step, and let* $M \to_{\mathbf{m}} s$ *be the corresponding step in* $\lambda^{\mathbf{m}}$*. Then* $s \rhd N$*.*

3 The $\mathcal{W}$-measure

In this section, we **define the $\mathcal{W}$-measure** (Def. 12) and we **prove that it is decreasing** (Thm. 15). Let us try to convey some ideas that led to the definition of the $\mathcal{W}$-measure. Recall that an abstract rewriting system $(A, \to)$ is *weakly Church–Rosser* (WCR) if $\leftarrow\to \subseteq \to^*\leftarrow^*$, *Church–Rosser* (CR) if $\leftarrow^*\to^* \subseteq \to^*\leftarrow^*$, and *increasing* (Inc) if there exists a function $|\cdot| : A \to \mathbb{N}$ such that $a \to b$ implies $|a| < |b|$. Let us also recall *Klop–Nederpelt's lemma* [31, Theorem 1.2.3 (iii)], which states that $\text{Inc} \wedge \text{WCR} \wedge \text{WN} \implies \text{SN} \wedge \text{CR}$.

Let $(A, \to)$ be increasing and WCR. Given a reduction $a \to^* b$, where b is a normal form, we can find a *decreasing* measure for the set of objects reachable from a, that is, the set $\{c \in A \mid a \to^* c\}$. In fact, by Klop–Nederpelt's lemma, we know that for every $c \in A$ such that $a \to^* c$ we have that $c \to^* b$, which implies that $|c| \leq |b|$, and hence we can define $\#(c) := |b| - |c|$. It is easy to see that $\#(-)$ is a decreasing measure, since $c \to c'$ implies that $|c| < |c'|$ so $\#(c) := |b| - |c| > |b| - |c'| = \#(c')$. Furthermore, the value of $\#(c)$ does not depend on the choice of a, by uniqueness of normal forms.

The idea behind the $\mathcal{W}$-measure is that the construction of a *decreasing* measure can be based on an *increasing* measure, according to the previous observation. It is not possible to build an increasing measure directly for the STLC; *e.g.* the following infinite sequence of expansions $t \leftarrow I\, t \leftarrow I\, (I\, t) \leftarrow \ldots$ would induce an infinite decreasing chain of natural numbers $|t| > |I\, t| > |I\, (I\, t)| > \ldots$.

One could try to define an increasing measure in a variant of the STLC such as Endrullis *et al.*'s clocked λ-calculus [14], in which the β-rule becomes $(\lambda x.\, t)\, s \to \tau(t[x := s])$, that is, contracting a β-redex produces a counter "τ" that keeps track of the number of contracted redexes. One could then count the number of τ's: for example, in the reduction sequence $(\lambda x.\, x\,(x\,y))\, I \to \tau(I\,(I\,y)) \to \tau\tau(I\,y) \to \tau\tau\tau y$ the number of counters strictly increases with each step. Unfortunately, this does not define an increasing measure, due to *erasure*. For example, $(\lambda x.\, y)\, t \to \tau y$ erases all the counters in t.

This is the motivation behind the definition of the $\lambda^{\mathbf{m}}$-calculus, which avoids erasure by always keeping an extra copy of the argument in a wrapper. The $\lambda^{\mathbf{m}}$-calculus is indeed increasing: in a step $t \to_{\mathbf{m}} s$ one has that $\mathsf{w}(t) < \mathsf{w}(s)$, where we recall that $\mathsf{w}(t)$ denotes the *weight*, *i.e.* the number of wrappers in t. For example, the step $(\lambda x.\, y)\,(z\{z\}) \to_{\mathbf{m}} y\{z\{z\}\}$ increases the number of wrappers. The decreasing measure $\mathcal{W}(M)$ is defined essentially by reducing M to normal form in the $\lambda^{\mathbf{m}}$-calculus and counting the number of wrappers in the result:

▶ **Definition 12** (The $\mathcal{W}$-measure)**.** *For each typable λ-term M, define $\mathcal{W}(M) \stackrel{\text{def}}{=} \mathsf{w}(\mathtt{S}_*(M))$.*

As we show below, $\mathtt{S}_*(M)$ turns out to be exactly the normal form of M in the $\lambda^{\mathbf{m}}$-calculus. We insist in writing $\mathtt{S}_*(M)$ to emphasize that the *definition* of the $\mathcal{W}$-measure does not require to prove that the $\lambda^{\mathbf{m}}$-calculus is weakly normalizing. Indeed, the simplification $\mathtt{S}_d(t)$ can be defined by structural induction on t, and the full simplification $\mathtt{S}_*(t) = \mathtt{S}_1(\mathtt{S}_2(\ldots \mathtt{S}_D(t)))$ can be calculated in exactly D iterations. On the other hand, the *proof* that the $\mathcal{W}$-measure is decreasing does rely on the fact that $\mathtt{S}_*(M)$ is the normal form of M.

In the remainder of this section, we prove that the $\mathcal{W}$-measure is indeed decreasing. The following lemma states that forgetful reduction decreases weight, and it is straightforward to prove:

▶ **Lemma 13.** *If $t \rhd^+ s$ then $\mathsf{w}(t) > \mathsf{w}(s)$.*

The proof that the $\mathcal{W}$-measure decreases relies on the two following properties that relate full simplification $\mathtt{S}_*(-)$ respectively with reduction ($\to_{\mathbf{m}}$) and forgetful reduction ($\rhd^+$):

▶ **Lemma 14.** 1. *If $t \to_{\mathbf{m}} s$ then $\mathtt{S}_*(t) = \mathtt{S}_*(s)$.* 2. *If $t \rhd^+ s$ then $\mathtt{S}_*(t) \rhd^+ \mathtt{S}_*(s)$.*

Proof. For the first item, note that by Prop. 7, we know that $t \to_{\mathbf{m}}^* \mathtt{S}_*(t)$ and that $t \to_{\mathbf{m}} s \to_{\mathbf{m}}^* \mathtt{S}_*(s)$, where moreover $\mathtt{S}_*(t)$ and $\mathtt{S}_*(s)$ are $\to_{\mathbf{m}}$-normal forms. By confluence (Prop. 4), this means that $\mathtt{S}_*(t) = \mathtt{S}_*(s)$.

For the second item, note that by Prop. 7, we know that $t \to_{\mathbf{m}}^* \mathtt{S}_*(t)$. Since we also know $t \rhd^+ s$ by hypothesis, and since forgetful reduction commutes with reduction (Prop. 10), there exists a term u such that $s \to_{\mathbf{m}}^* u$ and $\mathtt{S}_*(t) \rhd^+ u$. By Prop. 7 we know that $\mathtt{S}_*(t)$ is in normal form, so by Prop. 10 u must also be a normal form. On the other hand, by Prop. 7 we know that $s \to_{\mathbf{m}}^* \mathtt{S}_*(s)$, where $\mathtt{S}_*(s)$ must also be a normal form. In summary, we have that $s \to_{\mathbf{m}}^* u$ and $s \to_{\mathbf{m}}^* \mathtt{S}_*(s)$, where both u and $\mathtt{S}_*(s)$ are normal forms. By confluence (Prop. 4) $u = \mathtt{S}_*(s)$, and from this we obtain that $\mathtt{S}_*(t) \rhd^+ u = \mathtt{S}_*(s)$, as required. ◀

▶ **Theorem 15.** *Let M, N be typable λ-terms such that $M \to_\beta N$. Then $\mathcal{W}(M) > \mathcal{W}(N)$.*

Proof. Given the step $M \to_\beta N$, consider the corresponding step $M \to_{\mathbf{m}} s$, and note that $s \rhd^+ N$ by the reduce/forget lemma (Lem. 11). Since $M \to_{\mathbf{m}} s \rhd^+ N$, by Lem. 14, we have that $\mathtt{S}_*(M) = \mathtt{S}_*(s) \rhd^+ \mathtt{S}_*(N)$. Finally, by Lem. 13, $\mathcal{W}(M) = \mathsf{w}(\mathtt{S}_*(M)) > \mathsf{w}(\mathtt{S}_*(N)) = \mathcal{W}(N)$. ◀

The following is one example that the $\mathcal{W}$-measure decreases (see Ex. 6 for another example):

▶ **Example 16.** Let $M = (\lambda x^0.\, y^{0\to0\to0}\, x^0\, x^0)\, ((\lambda x^{0\to0}.\, x^{0\to0}\, z^0)\, f^{0\to0})$, consider the step $M = (\lambda x.\, y\, x\, x)((\lambda x.\, x\, z)\, f) \to_\beta (\lambda x.\, y\, x\, x)\,(f\, z) = N$, and note that $\mathcal{W}(M) = \mathtt{w}(\mathtt{S}_*(M)) = 4 > 1 = \mathcal{W}(N)$, since:

$$\mathtt{S}_*(M) = (y\,(f\,z)\{\!\!\{f\}\!\!\}\,(f\,z)\{\!\!\{f\}\!\!\})\{\!\!\{(f\,z)\{\!\!\{f\}\!\!\}\}\!\!\} \qquad \mathtt{S}_*(N) = (y\,(f\,z)\,(f\,z))\{\!\!\{f\,z\}\!\!\}$$

4 Reduction by degrees

This section is of purely technical nature. The aim is to develop tools that we use in the following section to reason about the $\mathcal{T}^{\mathbf{m}}$-measure. To do so, we need to introduce witnesses of steps and reduction sequences, treating the $\lambda^{\mathbf{m}}$-calculus as an *abstract rewriting system* in the sense of [31, Def. 8.2.2] or as a *transition system* in the sense of [24, Def. 1]. *Objects* are $\lambda^{\mathbf{m}}$-terms, *steps* are 5-uples $R = (\mathtt{C}, x, t, \mathtt{L}, s)$ witnessing the reduction step $\mathtt{C}[(\lambda x.\, t)\mathtt{L}\, s] \to_{\mathbf{m}} \mathtt{C}[t[x := s]\{\!\!\{s\}\!\!\}\mathtt{L}]$ under a context $\mathtt{C}$, and *reductions* $(\rho, \sigma, \ldots)$ are sequences of composable steps. Similarly, *forgetful steps* are triples $R = (\mathtt{C}, t, s)$ witnessing the forgetful reduction $\mathtt{C}[t\{\!\!\{s\}\!\!\}] \rhd \mathtt{C}[t]$, and *forgetful reductions* (also written $\rho, \sigma, \ldots$) are sequences of composable forgetful steps. We write ρ^{src} and ρ^{tgt} respectively for the source and target terms of ρ.

For each $d \in \mathbb{N}_0$, we define **reduction of degree** d as follows:

▶ **Definition 17.** *$t \xrightarrow{d}_{\mathbf{m}} s$ if and only if $t \to_{\mathbf{m}} s$ by contracting a redex of degree d.*

We write $R : t \xrightarrow{d}_{\mathbf{m}} s$ if R is a step witnessing a reduction step of degree d, and $\rho : t \xrightarrow{d}{}^*_{\mathbf{m}} s$ if ρ is a reduction witnessing a sequence of reduction steps of degree d.

The following results require to explicitly manipulate steps and reductions.

▶ **Proposition 18** (Commutation of reduction by degrees)**.** *For any two reductions $\rho : t_1 \xrightarrow{d}{}^*_{\mathbf{m}} t_2$ and $\sigma : t_1 \xrightarrow{D}{}^*_{\mathbf{m}} t_3$, there exists a term t_4 and one can construct reductions $\sigma/\rho : t_2 \xrightarrow{D}{}^*_{\mathbf{m}} t_4$ and $\rho/\sigma : t_3 \xrightarrow{d}{}^*_{\mathbf{m}} t_4$ such that, furthermore, if $d \neq D$, then* **1.** *ρ/σ contains at least as many steps as ρ; and* **2.** *ρ/σ determines ρ, that is, $\rho_1/\sigma = \rho_2/\sigma$ implies $\rho_1 = \rho_2$.*

Proof. This is reduced to the fact that the $\lambda^{\mathbf{m}}$-calculus can be understood as an orthogonal higher-order rewriting system in the sense of Nipkow [26]. Indeed, ρ/σ and σ/ρ can be taken to be the standard notion of projection based on residuals for orthogonal HRSs. Note that item **1.** holds because the $\lambda^{\mathbf{m}}$-calculus is non-erasing while item **2.** is a consequence of the *unique ancestor* property, *i.e.* each redex *descends* from at most one redex. ◀

▶ **Corollary 19** (Termination of reduction by degrees)**.** *The relation $\xrightarrow{d}_{\mathbf{m}}$ is strongly normalizing.*

Proof. This is a consequence of the fact that HRSs enjoy the Finite Developments property [31, Theorem 11.5.11], observing that reduction of degree d does not create redexes of degree d. Alternatively, it can be easily shown that $t \xrightarrow{d}{}^*_{\mathbf{m}} \mathtt{S}_d(t)$ and $\mathtt{S}_d(t)$ is in $\xrightarrow{d}_{\mathbf{m}}$-normal form, so $\xrightarrow{d}_{\mathbf{m}}$ is WN. Moreover, one can observe that $\xrightarrow{d}_{\mathbf{m}}$ is *uniformly normalizing* [19], given that there is no erasure, which entails that $\xrightarrow{d}_{\mathbf{m}}$ is SN. ◀

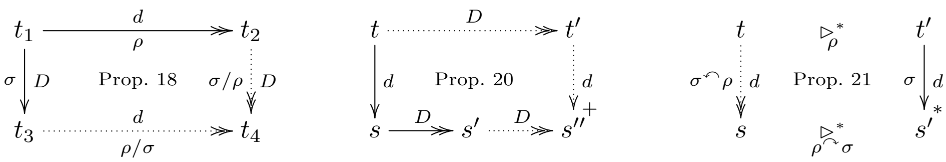
▶ **Proposition 20** (Lifting property for lower steps)**.** *Let $d < D$ and $t \xrightarrow{d}_{\mathbf{m}} s \xrightarrow{D}{}^*_{\mathbf{m}} s'$. Then there exist terms t', s'' such that $t \xrightarrow{D}{}^*_{\mathbf{m}} t'$ and $s' \xrightarrow{D}{}^*_{\mathbf{m}} s''$ and $t' \xrightarrow{d}{}^+_{\mathbf{m}} s''$.*

Proof. Note that $t \xrightarrow{D}_{\mathbf{m}}^{*} \mathtt{S}_D(t)$. By Prop. 18, there exists a term u such that $s \xrightarrow{D}_{\mathbf{m}}^{*} u$ and $\mathtt{S}_D(t) \xrightarrow{d}_{\mathbf{m}}^{+} u$. Again by Prop. 18, there exists s'' such that $u \xrightarrow{D}_{\mathbf{m}}^{*} s''$ and $s' \xrightarrow{D}_{\mathbf{m}}^{*} s''$. Moreover, $\mathtt{S}_D(t)$ is in $\xrightarrow{D}_{\mathbf{m}}$-normal form. Since $\mathtt{S}_D(t) \xrightarrow{d}_{\mathbf{m}}^{*} u$ with $d < D$ and reduction does not create redexes of higher degree, u is also in $\xrightarrow{D}_{\mathbf{m}}$-normal form, so $u = s''$, and we are done. ◀

▶ **Proposition 21** (Postponement of forgetful reduction). *For any two reductions $\rho : t \rhd^* t'$ and $\sigma : t' \xrightarrow{d}_{\mathbf{m}}^{*} s'$, there exists a term s and reductions $\rho^\frown\sigma : s \rhd^* s'$ and $\sigma^\frown\rho : t \xrightarrow{d}_{\mathbf{m}}^{*} s$. Furthermore, $\sigma^\frown\rho$ determines σ, that is, $\sigma_1^\frown\rho = \sigma_2^\frown\rho$ implies $\sigma_1 = \sigma_2$.*

Proof. This can be reduced to an analysis of the critical pairs between the rewriting rules defining $\rhd^{-1}$ and $\to_{\mathbf{m}}$. Critical pairs are of the form $(\lambda x.\, t)\mathtt{L}_1\{s\}\mathtt{L}_2\, u \rhd (\lambda x.\, t)\mathtt{L}_1\mathtt{L}_2\, u \to_{\mathbf{m}} t[x := u]\{u\}\mathtt{L}_1\mathtt{L}_2$ and can be closed by $(\lambda x.\, t)\mathtt{L}_1\{s\}\mathtt{L}_2\, u \to_{\mathbf{m}} t[x := u]\{u\}\mathtt{L}_1\{s\}\mathtt{L}_2 \rhd t[x := u]\{u\}\mathtt{L}_1\mathtt{L}_2$. ◀

The following diagrams depict the statements of the three preceding propositions:



5 The $\mathcal{T}^{\mathbf{m}}$-measure

In this section, we **define the $\mathcal{T}^{\mathbf{m}}$-measure** (Def. 25) and we **prove that it is decreasing** (Thm. 32). We start with some preliminary notions.

A partially ordered set $(X, >)$ is *well-founded* if there are no infinite decreasing chains. $\mathbb{M}(X)$ denotes the set of *finite multisets* over a set X, which are functions $\mathfrak{m} : X \to \mathbb{N}_0$ such that $\mathfrak{m}(x) > 0$ for finitely many values of $x \in X$. We write $\mathfrak{m} + \mathfrak{n}$ for the sum of multisets, and $x \in \mathfrak{m}$ if $\mathfrak{m}(x) > 0$. We write $[x_1, \ldots, x_n]$ for the multiset of elements $x_1, \ldots, x_n$, taking multiplicities into account. If X is a finite set and $f : X \to Y$ is a function, we use the "multiset builder" notation $[f(x) \mid\mid x \in X]$ to denote the multiset $\sum_{x \in X}[f(x)]$. If $(X, >)$ is a partially ordered set, we define a binary relation $\succ^1$ on multisets by declaring that $\mathfrak{m} + [x] \succ^1 \mathfrak{m} + \mathfrak{n}$ if $x > y$ for every $y \in \mathfrak{n}$. The *multiset order* induced by $(X, >)$ is the strict order relation on multisets defined by declaring that $\mathfrak{m} \succ \mathfrak{n}$ if and only if $\mathfrak{m}\ (\succ^1)^+\ \mathfrak{n}$. We recall the following widely known theorem by Dershowitz and Manna [12]:

▶ **Theorem 22.** *If $(X, >)$ is well-founded, then $(\mathbb{M}(X), \succ)$ is well-founded.*

As usual, $\mathfrak{m} \succeq \mathfrak{n}$ stands for $(\mathfrak{m} = \mathfrak{n} \vee \mathfrak{m} \succ \mathfrak{n})$, and $\mathfrak{m} \preceq \mathfrak{n}$ stands for $\mathfrak{n} \succeq \mathfrak{m}$. We define an operation $k \otimes \mathfrak{m}$ by the recursive equations $0 \otimes \mathfrak{m} \stackrel{\text{def}}{=} [\,]$ and $(1 + k) \otimes \mathfrak{m} \stackrel{\text{def}}{=} \mathfrak{m} + k \otimes \mathfrak{m}$. The relation $\mathfrak{m} :\succ: \mathfrak{n}$, called the *pointwise multiset order*, is defined to hold if $\mathfrak{m}$ and $\mathfrak{n}$ can be written as of the forms $\mathfrak{m} = [x_1, \ldots, x_n]$ and $\mathfrak{n} = [y_1, \ldots, y_n]$ in such a way that $x_i > y_i$ for all $i \in 1..n$. Observe that if $\mathfrak{m} :\succ: \mathfrak{n}$ then for all $k \in \mathbb{N}_0$ we have that $\mathfrak{m} \succeq k \otimes \mathfrak{n}$. Another easy-to-check property is that if $\mathfrak{m} :\succ: \mathfrak{n}$ and $\mathfrak{m}$ is non-empty then $\mathfrak{m} \succ \mathfrak{n}$.

A first frustrated attempt. As mentioned in the introduction, Turing's measure, given by $\mathcal{T}(M) \stackrel{\text{def}}{=} [d \mid\mid R$ is a redex occurrence of degree d in $M]$, decreases when contracting the rightmost redex of highest degree. Our goal is to mend the $\mathcal{T}$-measure in such a way that contracting *any* redex decreases the measure. The difficulty is that a redex of degree d may copy redexes of a higher or equal degree $d' \geq d$. So one can wonder: whenever a redex R of degree d makes n copies of a redex S of degree $d' \geq d$, in what sense can the copies of S be considered "smaller" than S? To address this, we generalize the $\mathcal{T}$-measure to a family of measures $\mathcal{T}_D(M) \stackrel{\text{def}}{=} [(d, \mathcal{T}_{d-1}(M)) \mid\mid R \text{ is a redex occurrence of degree } d \leq D \text{ in } M]$ indexed by a degree $D \in \mathbb{N}_0$. Note that $\mathcal{T}_0(M)$ is the empty multiset because there are no redexes of degree 0.

Let us try to argue that if $d \leq D$ and $M \xrightarrow{d}_\beta N$ then $\mathcal{T}_D(M) \succ \mathcal{T}_D(N)$. Here $M \xrightarrow{d}_\beta N$ means that $M \to_\beta N$ by contracting a redex of degree d. Suppose that the contraction of the redex $R : M \xrightarrow{d}_\beta N$ copies a redex S of degree d', where we assume that $d < d' \leq D$, producing n copies $S_1, \ldots, S_n$. Note that the contribution of S to the multiset is $(d', \mathcal{T}_{d'-1}(M))$, and the contribution of each S_i is $(d', \mathcal{T}_{d'-1}(N))$. By induction on D, we could inductively argue that $\mathcal{T}_{d'-1}(M) \succ \mathcal{T}_{d'-1}(N)$, since $d' - 1 < d' \leq D$. So far the property would seem to hold.

The problem with this proposal is that a redex R of degree d may still make copies of redexes of degree *exactly* d, whose contribution does not necessarily decrease[2].

A second frustrated attempt. The difficulty is to deal with the situation in which a redex R of degree d makes n copies of a redex S of the same degree d. A key observation is that a reduction sequence $M \xrightarrow{d}{}^*_\beta N$ must be a *development*[3] of the set of redexes of degree d. This is because contracting a redex of degree d can only create redexes of degree strictly less than d, so any redex of degree d that remains after one $\xrightarrow{d}_\beta$-step must be a *residual* of a preexisting redex. This motivates our second attempt to define a measure, consisting of two families of measures $\mathcal{T}^\beta_{\leq D}(-)$ and $\mathcal{R}^\beta_D(-)$, indexed by $D \in \mathbb{N}_0$ and defined mutually recursively:

$$\mathcal{T}^\beta_{\leq D}(M) \stackrel{\text{def}}{=} [(d, \mathcal{R}^\beta_d(M)) \mid\mid R \text{ is a } \beta\text{-redex occurrence of degree } d \leq D \text{ in } M]$$

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

Note that there are no redexes of degree 0, so $\mathcal{T}^\beta_{\leq D}(M)$ may not depend on $\mathcal{R}^\beta_0(M)$. In fact, $\mathcal{R}^\beta_D(M)$ is defined only for $D \geq 1$. The recursive definition is well-founded because $\mathcal{T}^\beta_{\leq D}(M)$ may depend on $\mathcal{R}^\beta_1(M), \ldots, \mathcal{R}^\beta_D(M)$ which in turn may only depend on $\mathcal{T}^\beta_{\leq d}(M')$ for $d < D$. The multiplicity of $\mathcal{T}^\beta_{\leq D-1}(M')$ in the multiset $\mathcal{R}^\beta_D(M)$ is given by the number of reduction sequences that contract only redexes of degree D, that is, the number of different paths $M \xrightarrow{D}{}^*_{\mathbf{m}} M'$. One important point is that, for the measure $\mathcal{R}^\beta_D(t)$ to be well defined, one needs to argue that the number of paths $M \xrightarrow{D}{}^*_{\mathbf{m}} M'$ is finite. Since $M \xrightarrow{D}{}^*_{\mathbf{m}} M'$ is a development, this is a consequence of the *finite developments* (FD) property for orthogonal HRSs [31, Theorem 11.5.11].[4]

[2] For example, in $M = (\lambda x^0.\, y^{0\to0\to0}\, x^0\, x^0)\, ((\lambda z^0.\, z^0)\, w^0) \xrightarrow{1}_\beta y^{0\to0\to0}\, ((\lambda z^0.\, z^0)\, w^0)\, ((\lambda z^0.\, z^0)\, w^0) = N$ the measure does not decrease, as $\mathcal{T}_1(M) = [(1, []), (1, [])] = \mathcal{T}_1(N)$.

[3] Recall that a development of a set of redexes X is a reduction sequence $M \to^*_\beta N$ in which each step contracts a *residual* of a redex in X. The residuals of a redex $S : t \to_\beta s$ after the contraction of a redex $R : t \to_\beta t'$ are, informally speaking, the "copies" left of S in t'. For formal definitions see [3, Section 11.2].

[4] Note that FD only ensures that developments are finite. To see that the set $\{\rho \mid M \xrightarrow{D}{}^*_{\mathbf{m}} M'\}$ is finite, one should resort to König's lemma, together with the fact that the STLC is finitely branching. For a constructive proof, one can use a computable decreasing measure, such as in de Vrijer's proof of FD [9].

Let us try to argue that if $d \leq D$ and $M \xrightarrow{d}_\beta N$ then $\mathcal{T}^\beta_{\leq D}(M) \succ \mathcal{T}^\beta_{\leq D}(N)$. On the first hand, if a redex $R : M \xrightarrow{d}_\beta N$ of degree d copies a redex S of *exactly* the same degree d making n copies $S_1, \ldots, S_n$, the contribution of S to the multiset is $(d, \mathcal{R}^\beta_d(M))$, whereas each S_i contributes $(d, \mathcal{R}^\beta_d(N))$, and we can argue that $\mathcal{R}^\beta_d(M) \succ \mathcal{R}^\beta_d(N)$, because we can injectively map each reduction sequence $\rho : N \xrightarrow{d}^*_\beta N'$ to the reduction sequence $R\rho : M \xrightarrow{d}_\beta N \xrightarrow{d}^*_\beta N'$, where $R\rho$ denotes the composition of R and ρ. Furthermore, there is an empty reduction sequence $M \xrightarrow{d}^*_\beta M$ contributing an element $\mathcal{T}^\beta_{\leq d-1}(M)$ to $\mathcal{R}^\beta_d(M)$ but not to $\mathcal{R}^\beta_d(N)$.

On the other hand, if the contraction of a redex $R : M \xrightarrow{d}_\beta N$ of degree d copies a redex S of *strictly* greater degree $d' > d$ making n copies $S_1, \ldots, S_n$, the weight of S is $(d', \mathcal{R}^\beta_{d'}(M))$ and the weight of each S_i is $(d', \mathcal{R}^\beta_{d'}(N))$, and we would need to show that $\mathcal{R}^\beta_{d'}(M) \succ \mathcal{R}^\beta_{d'}(N)$. One way to do so would be to map each reduction sequence $\rho : N \xrightarrow{d}^*_\beta N'$ to a reduction sequence $\sigma : M \xrightarrow{d}^*_\beta M'$ such that $\mathcal{T}^\beta_{\leq d'-1}(M') \succ \mathcal{T}^\beta_{\leq d'-1}(N')$. However, there does not seem to be a way to rule out the possibility that σ might erase R and that $M' = N'$, which would yield $\mathcal{T}^\beta_{\leq d'-1}(M') = \mathcal{T}^\beta_{\leq d'-1}(N')$, rather than a strict inequality. The root of the problem seems again to be *erasure*.

Definition of the $\mathcal{T}^{\mathbf{m}}$-measure. The $\mathcal{T}^{\mathbf{m}}$-measure is based on the ideas described above, but considering reduction in the $\lambda^{\mathbf{m}}$-calculus rather than in the STLC, to ensure that there is no erasure. Informally, the $\mathcal{T}^{\mathbf{m}}$-measure is defined by means of the two following equations. These equations are exactly as the ones defining $\mathcal{T}^\beta_{\leq D}(-)$ and $\mathcal{R}^\beta_D(-)$ above, with the only difference that they deal with $\lambda^{\mathbf{m}}$-terms and $\to_{\mathbf{m}}$-reduction rather than with pure λ-terms and $\to_\beta$-reduction:

$$\mathcal{T}^{\mathbf{m}}_{\leq D}(t) \stackrel{\text{def}}{=} [(d, \mathcal{R}^{\mathbf{m}}_d(t)) \mid\mid R \text{ is a } \mathbf{m}\text{-redex occurrence of degree } d \leq D \text{ in } t]$$

$$\mathcal{R}^{\mathbf{m}}_D(t) \stackrel{\text{def}}{=} [\mathcal{T}^{\mathbf{m}}_{\leq D-1}(t') \mid\mid \rho : t \xrightarrow{D}^*_{\mathbf{m}} t']$$

To be able to reason about these measures inductively, it will be convenient to define an auxiliary measure $\mathcal{T}^{\mathbf{m}}_d(t_0, t)$ as the multiset of elements of the form $(d, \mathcal{R}^{\mathbf{m}}_d(t_0))$ for each $\mathbf{m}$-redex occurrence of degree *exactly* d in t. This auxiliary measure takes two arguments t_0 and t, and it is defined by structural recursion on the second argument (t), while the first argument (t_0) is used to keep track of the original term. Note that, with this auxiliary definition, we can write $\mathcal{T}^{\mathbf{m}}_{\leq D}(t)$ as a sum, namely $\mathcal{T}^{\mathbf{m}}_{\leq D}(t) = \mathcal{T}^{\mathbf{m}}_1(t, t) + \ldots + \mathcal{T}^{\mathbf{m}}_D(t, t)$.

To define the measure formally, we start by precisely defining its codomain.

▶ **Definition 23** (Codomain of the $\mathcal{T}^{\mathbf{m}}$-measure). *For each $d \geq 0$, we define a set $\mathbb{T}_d$, and for $d \geq 1$ we define a set $\mathbb{R}_d$, mutually recursively:*

$$\mathbb{T}_d \stackrel{\text{def}}{=} \mathbb{M}(\{(i, b) \mid 1 \leq i \leq d,\ b \in \mathbb{R}_i\}) \qquad \mathbb{R}_d \stackrel{\text{def}}{=} \mathbb{M}(\mathbb{T}_{d-1})$$

The sets $\mathbb{T}_d$ and $\mathbb{R}_d$ are partially ordered by the induced multiset ordering on their elements. Tuples (i, b) are ordered with the lexicographic order, that is, $(i, b) > (i', b')$ if and only if $i > i' \vee (i = i' \wedge b \succ b')$. Note that $\mathbb{T}_0 = \{[\,]\}$ and that if $d \leq d'$ then $\mathbb{T}_d \subseteq \mathbb{T}_{d'}$ and $\mathbb{R}_d \subseteq \mathbb{R}_{d'}$. Moreover, $(\mathbb{T}_d, \succ)$ and $(\mathbb{R}_d, \succ)$ are well-founded partial orders by Thm. 22.

Given typable $\lambda^{\mathbf{m}}$-terms t_0, t, and $d \in \mathbb{N}_0$, we define $\mathcal{T}^{\mathbf{m}}_d(t_0, t) \in \mathbb{T}_d$ and $\mathcal{T}^{\mathbf{m}}_{\leq d}(t) \in \mathbb{T}_d$, and if $d > 0$ we define $\mathcal{R}^{\mathbf{m}}_d(t) \in \mathbb{R}_d$, by induction on d as follows. Note that $\mathcal{T}^{\mathbf{m}}_d(t_0, t)$ is defined by a nested induction on t, and it is also defined on memories ($\mathcal{T}^{\mathbf{m}}_d(t_0, \mathtt{L})$):

▶ **Definition 24** (The measures $\mathcal{T}^{\mathbf{m}}_{d}(-,-)$, $\mathcal{T}^{\mathbf{m}}_{\leq d}(-)$, and $\mathcal{R}^{\mathbf{m}}_{d}(-)$)**.**

$$
\begin{aligned}
\mathcal{T}^{\mathbf{m}}_{d}(t_0, x) &\overset{\text{def}}{=} [\,] \\
\mathcal{T}^{\mathbf{m}}_{d}(t_0, \lambda x.\, s) &\overset{\text{def}}{=} \mathcal{T}^{\mathbf{m}}_{d}(t_0, s) \\
\mathcal{T}^{\mathbf{m}}_{d}(t_0, s\, u) &\overset{\text{def}}{=} \begin{cases} \mathcal{T}^{\mathbf{m}}_{d}(t_0, s') + \mathcal{T}^{\mathbf{m}}_{d}(t_0, \mathtt{L}) + \mathcal{T}^{\mathbf{m}}_{d}(t_0, u) + [(d, \mathcal{R}^{\mathbf{m}}_{d}(t_0))] \\ \qquad\qquad\qquad \textit{if } s = (\lambda x.\, s')\mathtt{L} \textit{ and it is of degree } d \\ \mathcal{T}^{\mathbf{m}}_{d}(t_0, s) + \mathcal{T}^{\mathbf{m}}_{d}(t_0, u) \qquad\qquad \textit{otherwise} \end{cases} \\
\mathcal{T}^{\mathbf{m}}_{d}(t_0, s\{u\}) &\overset{\text{def}}{=} \mathcal{T}^{\mathbf{m}}_{d}(t_0, s) + \mathcal{T}^{\mathbf{m}}_{d}(t_0, u)
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{T}^{\mathbf{m}}_{d}(t_0, \Box) &\overset{\text{def}}{=} [\,] \\
\mathcal{T}^{\mathbf{m}}_{d}(t_0, \mathtt{L}\{t\}) &\overset{\text{def}}{=} \mathcal{T}^{\mathbf{m}}_{d}(t_0, \mathtt{L}) + \mathcal{T}^{\mathbf{m}}_{d}(t_0, t)
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{T}^{\mathbf{m}}_{\leq d}(t) &\overset{\text{def}}{=} \textstyle\sum_{i=1}^{d} \mathcal{T}^{\mathbf{m}}_{i}(t, t) \\
\mathcal{R}^{\mathbf{m}}_{d}(t) &\overset{\text{def}}{=} [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(t') \mid\mid \rho : t \xrightarrow{d}{}^{*}_{\mathbf{m}} t']
\end{aligned}
$$

Moreover, the $\mathcal{T}^{\mathbf{m}}$**-measure** itself is defined for λ-terms as follows:

▶ **Definition 25.** *If M is a typable λ-term, $\mathcal{T}^{\mathbf{m}}(M) \overset{\text{def}}{=} \mathcal{T}^{\mathbf{m}}_{\leq D}(M)$ where $D := \mathsf{maxdeg}(M)$.*

When we write $\mathcal{T}^{\mathbf{m}}_{\leq D}(M)$, we implicitly regard M as a $\lambda^{\mathbf{m}}$-term without any memorized terms.

From a higher-level perspective, the $\mathcal{T}^{\mathbf{m}}_{d}(t_0, t)$ measure defined above is the multiset of pairs of the form $(d, \mathcal{R}^{\mathbf{m}}_{d}(t_0))$ for each redex of degree d in t. Similarly, $\mathcal{T}^{\mathbf{m}}_{\leq D}(t)$ is the multiset of pairs of the form $(d, \mathcal{R}^{\mathbf{m}}_{d}(t))$ for each redex of degree $d \leq D$ in t. In particular, $\mathcal{T}^{\mathbf{m}}_{0}(t_0, t)$ and $\mathcal{T}^{\mathbf{m}}_{\leq 0}(t)$ are empty multisets, because there are no redexes of degree 0. Two easy remarks are that $D \leq D'$ implies $\mathcal{T}^{\mathbf{m}}_{\leq D}(t) \preceq \mathcal{T}^{\mathbf{m}}_{\leq D'}(t)$, and that $\mathcal{T}^{\mathbf{m}}_{d}(t_0, t\mathtt{L}) = \mathcal{T}^{\mathbf{m}}_{d}(t_0, t) + \mathcal{T}^{\mathbf{m}}_{d}(t_0, \mathtt{L})$.

▶ Remark 26. As mentioned in the preceding discussion, one important point is that for $\mathcal{R}^{\mathbf{m}}_{d}(-)$ to be well-defined we need to argue that the set $\{\rho \mid \exists t'.\ \rho : t \xrightarrow{d}{}^{*}_{\mathbf{m}} t'\}$ is finite. This is a consequence of Coro. 19.

▶ **Example 27.** Let $\Delta := \lambda x^{0\to 0}.\, x^{0\to 0}(x^{0\to 0} z^0)$ and $W := \lambda y^0.\, w^0$ and consider the diagram:

$$
t_0 = \Delta\, W \xrightarrow{2} t_1 = (W(Wz))\{W\} \qquad t_1 \xrightarrow{1} t_2 = w\{Wz\}\{W\} \xrightarrow{1} t_4 = w\{w\{z\}\}\{W\} \qquad t_1 \xrightarrow{1} t_3 = (W(w\{z\}))\{W\} \xrightarrow{1} t_4
$$

Then $\mathcal{T}^{\mathbf{m}}_{\leq 0}(t_1) = \mathcal{T}^{\mathbf{m}}_{\leq 0}(t_2) = \mathcal{T}^{\mathbf{m}}_{\leq 0}(t_3) = \mathcal{T}^{\mathbf{m}}_{\leq 0}(t_4) = \mathcal{T}^{\mathbf{m}}_{\leq 1}(t_4) = \mathcal{T}^{\mathbf{m}}_{\leq 2}(t_4) = [\,]$, and:

$$
\begin{array}{ll}
\mathcal{T}^{\mathbf{m}}_{\leq 2}(t_0) = [(2, \mathcal{R}^{\mathbf{m}}_{2}(t_0))] & \mathcal{R}^{\mathbf{m}}_{2}(t_0) = [\mathcal{T}^{\mathbf{m}}_{\leq 1}(t_0), \mathcal{T}^{\mathbf{m}}_{\leq 1}(t_1)] \\
\mathcal{T}^{\mathbf{m}}_{\leq 2}(t_1) = \mathcal{T}^{\mathbf{m}}_{\leq 1}(t_1) = [(1, \mathcal{R}^{\mathbf{m}}_{1}(t_1)), (1, \mathcal{R}^{\mathbf{m}}_{1}(t_1))] & \mathcal{R}^{\mathbf{m}}_{1}(t_1) = [\mathcal{T}^{\mathbf{m}}_{\leq 0}(t_1), \mathcal{T}^{\mathbf{m}}_{\leq 0}(t_2), \mathcal{T}^{\mathbf{m}}_{\leq 0}(t_3), \mathcal{T}^{\mathbf{m}}_{\leq 0}(t_4)] \\
\mathcal{T}^{\mathbf{m}}_{\leq 2}(t_2) = \mathcal{T}^{\mathbf{m}}_{\leq 1}(t_2) = [(1, \mathcal{R}^{\mathbf{m}}_{1}(t_2))] & \mathcal{R}^{\mathbf{m}}_{1}(t_2) = [\mathcal{T}^{\mathbf{m}}_{\leq 0}(t_2), \mathcal{T}^{\mathbf{m}}_{\leq 0}(t_4)] \\
\mathcal{T}^{\mathbf{m}}_{\leq 2}(t_3) = \mathcal{T}^{\mathbf{m}}_{\leq 1}(t_3) = [(1, \mathcal{R}^{\mathbf{m}}_{1}(t_3))] & \mathcal{R}^{\mathbf{m}}_{1}(t_3) = [\mathcal{T}^{\mathbf{m}}_{\leq 0}(t_3), \mathcal{T}^{\mathbf{m}}_{\leq 0}(t_4)]
\end{array}
$$

In particular, $\mathcal{T}^{\mathbf{m}}_{\leq 2}(t_0) \succ \mathcal{T}^{\mathbf{m}}_{\leq 2}(t_1) \succ \mathcal{T}^{\mathbf{m}}_{\leq 2}(t_2) \succ \mathcal{T}^{\mathbf{m}}_{\leq 2}(t_4)$ and $\mathcal{T}^{\mathbf{m}}_{\leq 2}(t_1) \succ \mathcal{T}^{\mathbf{m}}_{\leq 2}(t_3) \succ \mathcal{T}^{\mathbf{m}}_{\leq 2}(t_4)$.

The $\mathcal{T}^{\mathbf{m}}$-measure is decreasing. Lastly, we show the main theorem of this section, stating that if $M \to_\beta N$ then $\mathcal{T}^{\mathbf{m}}(M) \succ \mathcal{T}^{\mathbf{m}}(N)$. This theorem is based on three technical results, that we call *high/increase*, *low/decrease*, and *forget/decrease*:

1. **High/increase** (Prop. 29) establishes – perhaps confusingly – that $\mathcal{T}^{\mathbf{m}}_{\leq d}(-)$ (non-strictly) **increases** if one contracts a redex of higher degree $D > d$. More precisely, if $0 \leq d < D$ and $t \xrightarrow{D}_{\mathbf{m}} t'$ then $\mathcal{T}^{\mathbf{m}}_{\leq d}(t) \preceq \mathcal{T}^{\mathbf{m}}_{\leq d}(t')$. Note that $\mathcal{T}^{\mathbf{m}}_{\leq d}(t)$ only looks at redexes of degree $i \leq d$, and contracting a redex of degree $D > d$ *cannot erase* a redex of any degree $i \leq d$, because the $\lambda^{\mathbf{m}}$-calculus is non-erasing. Contracting a redex of degree D can, at most, replicate redexes of degree i. This property is needed for a technical reason to prove the low/decrease property, and it relies crucially on the commutation result of the previous section (Prop. 18).
2. **Low/decrease** (Prop. 30) establishes that $\mathcal{T}^{\mathbf{m}}_{\leq D}(-)$ **strictly** decreases if one contracts a redex of lower degree $d < D$. More precisely, if $1 \leq d \leq D$ and $t \xrightarrow{d}_{\mathbf{m}} t'$ then $\mathcal{T}^{\mathbf{m}}_{\leq D}(t) \succ \mathcal{T}^{\mathbf{m}}_{\leq D}(t')$. This is the core of the argument, and the most technically difficult part to prove. It relies crucially on the lifting property of the previous section (Prop. 20).
3. **Forget/decrease** (Prop. 31) establishes that forgetful reduction (non-strictly) decreases the measure. More precisely, if $t \rhd t'$ then $\mathcal{T}^{\mathbf{m}}_{\leq d}(t) \succeq \mathcal{T}^{\mathbf{m}}_{\leq d}(t')$. This property is used as a final step in the main theorem, and it relies crucially on postponement of forgetful reduction, as studied in the previous section (Prop. 21).

Below we sketch the proofs of these three properties. Let us first mention a straightforward lemma.

▶ **Lemma 28** (Measure of a substitution). **1.** $\mathcal{T}^{\mathbf{m}}_{d}(t_0, t) \preceq \mathcal{T}^{\mathbf{m}}_{d}(t_0, t[x := s])$. **2.** *If s is not a $\mathbf{m}$-abstraction of degree d, then $\mathcal{T}^{\mathbf{m}}_{d}(t_0, t[x := s]) = \mathcal{T}^{\mathbf{m}}_{d}(t_0, t) + k \otimes \mathcal{T}^{\mathbf{m}}_{d}(t_0, s)$ for some $k \in \mathbb{N}_0$.*

Proof. By induction on t. ◀

▶ **Proposition 29** (High/increase). *Let $D \in \mathbb{N}_0$. Then the following hold:*
1. *If $1 \leq d < D$ and $t \xrightarrow{D}_{\mathbf{m}} t'$ then $\mathcal{R}^{\mathbf{m}}_{d}(t) \preceq \mathcal{R}^{\mathbf{m}}_{d}(t')$.*
2. *If $0 \leq d < D$ and $t_0 \xrightarrow{D}_{\mathbf{m}} t'_0$ then $\mathcal{T}^{\mathbf{m}}_{d}(t_0, t) \preceq \mathcal{T}^{\mathbf{m}}_{d}(t'_0, t)$.*
3. *If $0 \leq d < D$ and $t_0 \xrightarrow{D}_{\mathbf{m}} t'_0$ and $t \xrightarrow{D}_{\mathbf{m}} t'$ then $\mathcal{T}^{\mathbf{m}}_{d}(t_0, t) \preceq \mathcal{T}^{\mathbf{m}}_{d}(t'_0, t')$.*
4. *If $0 \leq d < D$ and $t \xrightarrow{D}_{\mathbf{m}} t'$ then $\mathcal{T}^{\mathbf{m}}_{\leq d}(t) \preceq \mathcal{T}^{\mathbf{m}}_{\leq d}(t')$.*

Proof. The four items are proved simultaneously by induction on d, where item **1** resorts to the IH, and the following items may resort to the previous items without decreasing d. Items **2** and **3** proceed by a nested induction on t. Most cases are straightforward.

One interesting situation occurs in item **3** when $t = (\lambda x.\, s)\mathtt{L}\, u$ is the redex of degree D contracted by the step $t \xrightarrow{D}_{\mathbf{m}} t'$. Then we resort to the first part of Lem. 28.

Another interesting part of the proof is item **1**. Let $1 \leq d < D$ and $t \xrightarrow{D}_{\mathbf{m}} t'$ and let us show that $\mathcal{R}^{\mathbf{m}}_{d}(t) \preceq \mathcal{R}^{\mathbf{m}}_{d}(t')$. Indeed, let $X := \{\rho \mid (\exists s)\ \rho : t \xrightarrow{d}{}^{*}_{\mathbf{m}} s\}$ and $Y := \{\sigma \mid (\exists s')\ \sigma : t' \xrightarrow{d}{}^{*}_{\mathbf{m}} s'\}$, and let $R : t \xrightarrow{D}_{\mathbf{m}} t'$. Using Prop. 18, we can define an injective function $\varphi : X \to Y$ by $\varphi(\rho) := \rho/R$. Note that $\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\rho^{\mathsf{tgt}}) \preceq \mathcal{T}^{\mathbf{m}}_{\leq d-1}(\varphi(\rho)^{\mathsf{tgt}})$ holds for every $\rho \in X$ using item **4** of the IH (noting that $1 \leq d-1 < D$ holds because $1 \leq d < D$), resorting to the IH as many times as the length of the reduction $s \xrightarrow{D}{}^{*}_{\mathbf{m}} s'_{\rho}$. To conclude the proof, let $Z = Y \setminus \varphi(X)$. Then:

$$\begin{aligned}\mathcal{R}^{\mathbf{m}}_{d}(t) &= [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in X] \preceq^{(\star)} [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\varphi(\rho)^{\mathsf{tgt}}) \mid\mid \rho \in X] =^{(\star\star)} [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\sigma^{\mathsf{tgt}}) \mid\mid \sigma \in \varphi(X)] \\ &\preceq [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\sigma^{\mathsf{tgt}}) \mid\mid \sigma \in \varphi(X)] + [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\sigma^{\mathsf{tgt}}) \mid\mid \sigma \in Z] = [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\sigma^{\mathsf{tgt}}) \mid\mid \sigma \in Y] = \mathcal{R}^{\mathbf{m}}_{d}(t')\end{aligned}$$

To justify the step marked with ($\star$), note that $[\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in X] = \sum_{\rho \in X}[\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\rho^{\mathsf{tgt}})] \preceq \sum_{\rho \in X}[\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\varphi(\rho)^{\mathsf{tgt}})] = [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\varphi(\rho)^{\mathsf{tgt}}) \mid\mid \rho \in X]$ because $\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\rho^{\mathsf{tgt}}) \preceq \mathcal{T}^{\mathbf{m}}_{\leq d-1}(\varphi(\rho)^{\mathsf{tgt}})$, as we have already claimed. To justify the step marked with ($\star\star$), note that φ is injective. ◀

▶ **Proposition 30** (Low/decrease). *Let $D \in \mathbb{N}_0$. Then the following hold:*

1. *If $1 \leq d \leq j \leq D$ and $t \xrightarrow{d}_{\mathbf{m}} t'$ then $\mathcal{R}^{\mathbf{m}}_j(t) \succ \mathcal{R}^{\mathbf{m}}_j(t')$.*
2. *If $1 \leq d \leq j \leq D$ and $t_0 \xrightarrow{d}_{\mathbf{m}} t'_0$ then $\mathcal{T}^{\mathbf{m}}_j(t_0, t) :\succ: \mathcal{T}^{\mathbf{m}}_j(t'_0, t)$.*
3. *If $1 \leq d \leq D$ and $t_0 \xrightarrow{d}_{\mathbf{m}} t'_0$ and $t \xrightarrow{d}_{\mathbf{m}} t'$, then for all $\mathfrak{m} \in \mathbb{T}_{d-1}$ we have $\mathcal{T}^{\mathbf{m}}_d(t_0, t) \succ \mathcal{T}^{\mathbf{m}}_d(t'_0, t') + \mathfrak{m}$.*
4. *If $1 \leq d < j \leq D$ and $t_0 \xrightarrow{d}_{\mathbf{m}} t'_0$ and $t \xrightarrow{d}_{\mathbf{m}} t'$ then $\mathcal{T}^{\mathbf{m}}_j(t_0, t) \succeq \mathcal{T}^{\mathbf{m}}_j(t'_0, t')$.*
5. *If $1 \leq d \leq D$ and $t \xrightarrow{d}_{\mathbf{m}} t'$ then $\mathcal{T}^{\mathbf{m}}_{\leq D}(t) \succ \mathcal{T}^{\mathbf{m}}_{\leq D}(t')$.*

Proof. The five items are proved simultaneously by induction on D, where item **1** resorts to the IH, and the following items may resort to the previous items without decreasing d. Items **2**–**4** proceed by a nested induction on t. We mention some of the interesting parts of the proof.

For item **1**, let $1 \leq d \leq j \leq D$ and $t \xrightarrow{d}_{\mathbf{m}} t'$ and let us show that $\mathcal{R}^{\mathbf{m}}_j(t) \succ \mathcal{R}^{\mathbf{m}}_j(t')$. Let $X := \{\rho \mid (\exists s)\ \rho : t \xrightarrow{j}{}^{*}_{\mathbf{m}} s\}$ and $Y := \{\sigma \mid (\exists s')\ \sigma : t' \xrightarrow{j}{}^{*}_{\mathbf{m}} s'\}$, and consider two subcases:

- If $d = j$, let $R : t \xrightarrow{d}_{\mathbf{m}} t'$, define an injective function $\varphi : Y \to X$ by $\varphi(\sigma) = R\,\sigma$, let $Z = X \setminus \varphi(Y)$, and note that:

$$\begin{aligned}\mathcal{R}^{\mathbf{m}}_j(t) &= [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in \varphi(Y)] + [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in Z] \\ &= [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(R\sigma^{\mathsf{tgt}}) \mid\mid \sigma \in Y] + [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in Z] \text{ since } \varphi \text{ is injective} \\ &= [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\sigma^{\mathsf{tgt}}) \mid\mid \sigma \in Y] + [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in Z] = \mathcal{R}^{\mathbf{m}}_j(t') + [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in Z]\end{aligned}$$

  To conclude that $\mathcal{R}^{\mathbf{m}}_j(t) \succ \mathcal{R}^{\mathbf{m}}_j(t')$, note that Z is non-empty because it contains the empty reduction $\epsilon : t \xrightarrow{d}{}^{*}_{\mathbf{m}} t$.
- If $d < j$, we construct a function $\varphi : Y \to X$ as follows. By Prop. 20, for each reduction $\sigma : t' \xrightarrow{j}{}^{*}_{\mathbf{m}} s'$ there exist s_σ, u_σ, and reductions $\varphi(\sigma) : t \xrightarrow{j}{}^{*}_{\mathbf{m}} s_\sigma$ and $s' \xrightarrow{j}{}^{*}_{\mathbf{m}} u_\sigma$ and $s_\sigma \xrightarrow{d}{}^{+}_{\mathbf{m}} u_\sigma$. Note that for every $\sigma \in Y$ we have $\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\varphi(\sigma)^{\mathsf{tgt}}) = \mathcal{T}^{\mathbf{m}}_{\leq j-1}(s_\sigma) \succ^{\dagger} \mathcal{T}^{\mathbf{m}}_{\leq j-1}(u_\sigma) \succeq^{\ddagger} \mathcal{T}^{\mathbf{m}}_{\leq j-1}(s') = \mathcal{T}^{\mathbf{m}}_{\leq j-1}(\sigma^{\mathsf{tgt}})$ where † holds by item 5 of the IH observing that $1 \leq d \leq j-1 < D$ because $d < j \leq D$, and ‡ holds by high/increase (Prop. 29) observing that $0 \leq j-1 < j$. To conclude the proof, let $Z = X \setminus \varphi(Y)$, and note that:

$$\begin{aligned}\mathcal{R}^{\mathbf{m}}_j(t) &= [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in \varphi(Y)] + [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in Z] \\ &= [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\varphi(\sigma)^{\mathsf{tgt}}) \mid\mid \sigma \in Y] + [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in Z] \\ &\succeq [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\varphi(\sigma)^{\mathsf{tgt}}) \mid\mid \sigma \in Y] \succ^{(\star)} [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\sigma^{\mathsf{tgt}}) \mid\mid \sigma \in Y] = \mathcal{R}^{\mathbf{m}}_j(t')\end{aligned}$$

  For the step marked with ($\star$), note that $[\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\varphi(\sigma)^{\mathsf{tgt}}) \mid\mid \sigma \in Y] :\succ: [\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\sigma^{\mathsf{tgt}}) \mid\mid \sigma \in Y]$ because $\mathcal{T}^{\mathbf{m}}_{\leq j-1}(\varphi(\sigma)^{\mathsf{tgt}}) \succ \mathcal{T}^{\mathbf{m}}_{\leq j-1}(\sigma^{\mathsf{tgt}})$ holds by the claim above where, moreover, Y is non-empty because it contains the empty reduction $\epsilon : t' \xrightarrow{j}{}^{*}_{\mathbf{m}} t'$.

Another interesting situation occurs in item **3**, when $t = (\lambda x.\, s)\mathtt{L}\, u$ is the redex of degree d contracted by the step $t \xrightarrow{d}_{\mathbf{m}} t'$. The step is of the form $t = (\lambda x.\, s)\mathtt{L}\, u \xrightarrow{d}_{\mathbf{m}} s[x := u]\{u\}\mathtt{L} = t'$. Note that u is not an abstraction of degree d, because it is the argument of an abstraction of degree d. So by Lem. 28 there exists $k \in \mathbb{N}_0$ such that $\mathcal{T}^{\mathbf{m}}_d(t'_0, s[x := u]) = \mathcal{T}^{\mathbf{m}}_d(t'_0, s) + k \otimes \mathcal{T}^{\mathbf{m}}_d(t'_0, u)$. The crucial observation is that $\mathcal{T}^{\mathbf{m}}_d(t_0, u) \succeq (1 + k) \otimes \mathcal{T}^{\mathbf{m}}_d(t'_0, u)$, which is because by item 2 we have that $\mathcal{T}^{\mathbf{m}}_d(t_0, u) :\succ: \mathcal{T}^{\mathbf{m}}_d(t'_0, u)$.

Finally, for item **5**, let $1 \leq d \leq D$ and $t \xrightarrow{d}_{\mathbf{m}} t'$ and let us show that $\mathcal{T}^{\mathbf{m}}_{\leq D}(t) \succ \mathcal{T}^{\mathbf{m}}_{\leq D}(t')$. Indeed:

$$\begin{aligned}
\mathcal{T}^{\mathbf{m}}_{\leq D}(t) &= \sum_{i=1}^{D} \mathcal{T}^{\mathbf{m}}_{i}(t,t) \succeq \mathcal{T}^{\mathbf{m}}_{d}(t,t) + \sum_{j=d+1}^{D} \mathcal{T}^{\mathbf{m}}_{j}(t,t) \\
&\succ \mathcal{T}^{\mathbf{m}}_{\leq d-1}(t') + \mathcal{T}^{\mathbf{m}}_{d}(t',t') + \sum_{j=d+1}^{D} \mathcal{T}^{\mathbf{m}}_{j}(t,t) \text{ by item } \mathbf{3}\text{, taking } \mathfrak{m} := \mathcal{T}^{\mathbf{m}}_{\leq d-1}(t') \\
&\succeq \mathcal{T}^{\mathbf{m}}_{\leq d-1}(t') + \mathcal{T}^{\mathbf{m}}_{d}(t',t') + \sum_{j=d+1}^{D} \mathcal{T}^{\mathbf{m}}_{j}(t',t') = \mathcal{T}^{\mathbf{m}}_{\leq D}(t') \text{ by item } \mathbf{4}.
\end{aligned}$$ ◀

▶ **Proposition 31** (Forget/decrease). *Let $d \in \mathbb{N}_0$. Then the following hold:*

1. *If $t \rhd t'$ then $\mathcal{R}^{\mathbf{m}}_{d}(t) \succeq \mathcal{R}^{\mathbf{m}}_{d}(t')$.*
2. *If $t_0 \rhd t'_0$ then $\mathcal{T}^{\mathbf{m}}_{d}(t_0,t) \succeq \mathcal{T}^{\mathbf{m}}_{d}(t'_0,t)$.*
3. *If $t_0 \rhd t'_0$ and $t \rhd t'$ then $\mathcal{T}^{\mathbf{m}}_{d}(t_0,t) \succeq \mathcal{T}^{\mathbf{m}}_{d}(t'_0,t')$.*
4. *If $t \rhd t'$ then $\mathcal{T}^{\mathbf{m}}_{\leq d}(t) \succeq \mathcal{T}^{\mathbf{m}}_{\leq d}(t')$.*

Proof. The four items are proved simultaneously by induction on D, where item **1** resorts to the IH, and the following items may resort to the previous items without decreasing d. Items **2** and **3** proceed by a nested induction on t.

The interesting part is item **1**, so let $t \rhd t'$ and let us show that $\mathcal{R}^{\mathbf{m}}_{d}(t) \succeq \mathcal{R}^{\mathbf{m}}_{d}(t')$. Let $X := \{\rho \mid (\exists s)\ \rho : t \xrightarrow{d}^{*}_{\mathbf{m}} s\}$ and $Y := \{\sigma \mid (\exists s')\ \sigma : t' \xrightarrow{d}^{*}_{\mathbf{m}} s'\}$. Define an injective function $\varphi : Y \to X$ by $\varphi(\sigma) := \sigma \frown R$, resorting to Prop. 21, where $\sigma \frown R : t \xrightarrow{d}^{*}_{\mathbf{m}} s_\sigma$. and $s_\sigma \rhd^{*} s'$. Note that for every $\sigma \in Y$ we have $\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\varphi(\sigma)^{\mathsf{tgt}}) = \mathcal{T}^{\mathbf{m}}_{\leq d-1}(s_\sigma) \succeq^{\dagger} \mathcal{T}^{\mathbf{m}}_{\leq d-1}(s') = \mathcal{T}^{\mathbf{m}}_{\leq d-1}(\sigma^{\mathsf{tgt}})$, where † holds by item **4** of the IH, observing that $d-1 < d$. To conclude the proof, let $Z = X \setminus \varphi(Y)$, and note that:

$$\begin{aligned}
\mathcal{R}^{\mathbf{m}}_{d}(t) &= [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in \varphi(Y)] + [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in Z] \succeq [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\rho^{\mathsf{tgt}}) \mid\mid \rho \in \varphi(Y)] \\
&=^{(\star)} [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\varphi(\sigma)^{\mathsf{tgt}}) \mid\mid \sigma \in Y] \succeq^{(\star\star)} [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\sigma^{\mathsf{tgt}}) \mid\mid \sigma \in Y] = \mathcal{R}^{\mathbf{m}}_{d}(t')
\end{aligned}$$

For the step marked with $(\star)$, note that φ is injective. For the step marked with $(\star\star)$, note that $[\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\varphi(\sigma)^{\mathsf{tgt}}) \mid\mid \sigma \in Y] = \sum_{\sigma \in Y} [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\varphi(\sigma)^{\mathsf{tgt}})] \succeq \sum_{\sigma \in Y} [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\sigma^{\mathsf{tgt}})] = [\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\sigma^{\mathsf{tgt}}) \mid\mid \sigma \in Y]$ because $\mathcal{T}^{\mathbf{m}}_{\leq d-1}(\varphi(\sigma)^{\mathsf{tgt}}) \succeq \mathcal{T}^{\mathbf{m}}_{\leq d-1}(\sigma^{\mathsf{tgt}})$, as we have already justified. ◀

Finally, we prove the main theorem in this section:

▶ **Theorem 32.** *Let M, N be typable λ-terms such that $M \to_\beta N$. Then $\mathcal{T}^{\mathbf{m}}(M) > \mathcal{T}^{\mathbf{m}}(N)$.*

Proof. Let $D = \mathsf{maxdeg}(M)$ and $D' = \mathsf{maxdeg}(N)$. Let $M \to_{\mathbf{m}} s$ be the step corresponding to $M \to_\beta N$. By Lem. 11 note that $s \rhd N$. Then:

$$\mathcal{T}^{\mathbf{m}}(M) = \mathcal{T}^{\mathbf{m}}_{\leq D}(M) \succ^{\text{Prop. 30}} \mathcal{T}^{\mathbf{m}}_{\leq D}(s) \succeq^{\text{Prop. 31}} \mathcal{T}^{\mathbf{m}}_{\leq D}(N) \succeq \mathcal{T}^{\mathbf{m}}_{\leq D'}(N) = \mathcal{T}^{\mathbf{m}}(N)$$

The last inequality holds because $D \geq D'$ since, as is well-known, contraction of a β-redex in the simply typed λ-calculus cannot create a redex of higher degree. ◀

6 Conclusion

We have defined two decreasing measures for the STLC, the $\mathcal{W}$-measure (Def. 12) and the $\mathcal{T}^{\mathbf{m}}$-measure (Def. 25). These measures are decreasing (Thm. 15 and Thm. 32 respectively) and, to the best of our knowledge, they provide two new proofs of strong normalization for the STLC. Both measures are defined constructively and by purely syntactic methods, using the $\lambda^{\mathbf{m}}$-calculus as an auxiliary tool.

The problem of finding a "straightforward" decreasing measure for β-reduction in the simply typed λ-calculus is posed as Problem #26 in the TLCA list of open problems [5], and as Problem #19 in the RTA list of open problems [11].

One strength of the $\mathcal{W}$-measure is that its codomain is simple: each term is mapped to a natural number. One weakness is that the definition of the $\mathcal{W}$-measure relies on reduction in the $\lambda^{\mathbf{m}}$-calculus, and computing the $\mathcal{W}$-measure is at least as costly as evaluating the λ-term itself. Measures based on Gandy's [16, 10] have similar characteristics. One question is whether the values of the $\mathcal{W}$-measure and measures based on Gandy's can be related. It is not immediate to establish a precise correspondence.

On the other hand, one strength of the $\mathcal{T}^{\mathbf{m}}$-measure is that it shows how to extend Turing's measure $\mathcal{T}(-)$ so that it decreases when contracting *any* redex. The proof is based on a delicate analysis of how contracting a redex of degree d may create and copy redexes of degree d', depending on whether $d < d'$, or $d = d'$, or $d > d'$. We hope that this may provide novel insights on why the STLC is SN. The codomain of the $\mathcal{T}^{\mathbf{m}}$-measure is not so simple, as the $\mathcal{T}^{\mathbf{m}}$-measure maps each term to a structure of nested multisets. Yet, it is "reasonably simple": the fact that the partial orders $\mathbb{T}_d$ and $\mathbb{R}_d$ are well-founded only relies on the ordinary multiset and lexicographic orderings. The $\mathcal{T}^{\mathbf{m}}$-measure is costly to compute; in particular $\mathcal{R}^{\mathbf{m}}_d(t)$ is defined as a sum over all reductions $\rho : t \xrightarrow{d}^{*}_{\mathbf{m}} t'$, which may produce a combinatorial explosion. Another weakness is that our proofs make use of relatively heavy rewriting machinery, as we have to keep explicit track of witnesses (*e.g.* in Section 4).

Besides the techniques mentioned in the introduction, other proofs of SN of the STLC can be found in the literature. For example, David [7] gives a purely syntactic proof of SN relying on the standardization theorem; Loader [23], as well as Joachimski and Matthes [18], give combinatorial proofs of SN based on inductive predicates characterizing strongly normalizing terms. As far as we know, the only proofs that explicitly construct decreasing measures are those based on Gandy's.

The idea of keeping "leftover garbage" can be traced back to at least the works of Nederpelt [21] and Klop [20], who studied non-erasing variants of (possibly) erasing rewriting systems, in order to relate weak and strong normalization. Many variations of these ideas have been explored in the past, such as in de Groote's notion of β_S reduction [8] or Neergaard and Sørensen calculus with memory [25]. Instead of using the $\lambda^{\mathbf{m}}$-calculus, it is possible that other non-erasing systems may be used. For instance, Gandy [16] translates λ-terms to the terms of λI-calculus to avoid erasing arguments.

The definition of reduction in the $\lambda^{\mathbf{m}}$-calculus, which allows arbitrary memory in between the abstraction and the application, is inspired by Accattoli and Kesner's work on calculi with explicit substitutions "at a distance" [1]. This mechanism can be traced back, again, to at least the work of Nederpelt [21].

The definition of the $\lambda^{\mathbf{m}}$-calculus as a means to obtain an *increasing* measure was inspired by the fact that, in explicit substitution calculi without erasure, labeled reduction (in the sense of Lévy labels [22]) increases the sum of the sizes of all the labels in the term [2].

References

1 Beniamino Accattoli and Delia Kesner. The structural *lambda*-calculus. In *Computer Science Logic, 24th International Workshop, CSL 2010, 19th Annual Conference of the EACSL, Brno, Czech Republic, August 23-27, 2010. Proceedings*, pages 381–395, 2010.

2 Pablo Barenbaum and Eduardo Bonelli. Optimality and the linear substitution calculus. In *2nd International Conference on Formal Structures for Computation and Deduction, FSCD 2017, September 3-9, 2017, Oxford, UK*, pages 9:1–9:16, 2017. doi:10.4230/LIPIcs.FSCD.2017.9.

3 Henk Barendregt. *The Lambda Calculus: Its Syntax and Semantics*, volume 103. Elsevier, 1984.

4 Henk P. Barendregt and Giulio Manzonetto. Turing's contributions to lambda calculus. In B. Cooper and J. van Leeuwen, editors, *Alan Turing - His Work and Impact*, pages 139–143. Elsevier, 2013.

5 TCLA Editorial Board. TLCA list of open problems. http://tlca.di.unito.it/opltlca/, 2006.

6 Thierry Coquand. Canonicity and normalization for dependent type theory. *Theor. Comput. Sci.*, 777:184–191, 2019. doi:10.1016/j.tcs.2019.01.015.

7 René David. Normalization without reducibility. *Ann. Pure Appl. Log.*, 107(1-3):121–130, 2001. doi:10.1016/S0168-0072(00)00030-0.

8 Philippe de Groote. The conservation theorem revisited. In Marc Bezem and Jan Friso Groote, editors, *Typed Lambda Calculi and Applications, International Conference on Typed Lambda Calculi and Applications, TLCA '93, Utrecht, The Netherlands, March 16-18, 1993, Proceedings*, volume 664 of *Lecture Notes in Computer Science*, pages 163–178. Springer, 1993. doi:10.1007/BFb0037105.

9 Roel de Vrijer. A direct proof of the finite developments theorem. *The Journal of symbolic logic*, 50(2):339–343, 1985.

10 Roel de Vrijer. Exactly estimating functionals and strong normalization. In *Indagationes Mathematicae (Proceedings)*, volume 90, pages 479–493. North-Holland, 1987.

11 Nachum Dershowitz, Jean-Pierre Jouannaud, and Jan Willem Klop. Open problems in rewriting. In *International Conference on Rewriting Techniques and Applications*, pages 445–456. Springer, 1991.

12 Nachum Dershowitz and Zohar Manna. Proving termination with multiset orderings. *Communications of the ACM*, 22(8):465–476, 1979.

13 Paul Downen, Philip Johnson-Freyd, and Zena M. Ariola. Abstracting models of strong normalization for classical calculi. *J. Log. Algebraic Methods Program.*, 111:100512, 2020. doi:10.1016/j.jlamp.2019.100512.

14 Jörg Endrullis, Dimitri Hendriks, Jan Willem Klop, and Andrew Polonsky. Clocked lambda calculus. *Math. Struct. Comput. Sci.*, 27(5):782–806, 2017. doi:10.1017/S0960129515000389.

15 Robin O. Gandy. An early proof of normalization by A.M. Turing. In J.P. Seldin and J.R. Hindley, editors, *To H.B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, pages 453–455. Academic Press, 1980.

16 Robin O. Gandy. Proofs of strong normalization. In J.P. Seldin and J.R. Hindley, editors, *To H.B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, pages 457–477. Academic Press, 1980.

17 Jean-Yves Girard. *Interprétation fonctionnelle et élimination des coupures de l'arithmétique d'ordre supérieur.* PhD thesis, Université Paris 7, 1972.

18 Felix Joachimski and Ralph Matthes. Short proofs of normalization for the simply- typed lambda-calculus, permutative conversions and Gödel's T. *Arch. Math. Log.*, 42(1):59–87, 2003. doi:10.1007/s00153-002-0156-9.

19 Zurab Khasidashvili, Mizuhito Ogawa, and Vincent van Oostrom. Uniform normalisation beyond orthogonality. In Aart Middeldorp, editor, *Rewriting Techniques and Applications, 12th International Conference, RTA 2001, Utrecht, The Netherlands, May 22-24, 2001, Proceedings*, volume 2051 of *Lecture Notes in Computer Science*, pages 122–136. Springer, 2001. doi:10.1007/3-540-45127-7_11.

20 Jan Willem Klop. *Combinatory Reduction Systems.* PhD thesis, Utrecht University, 1980.

21 Robert Pieter Nederpelt Lazarom. *Strong normalization in a typed lambda calculus with lambda structured types.* PhD thesis, TU Eindhoven, 1973.

22 Jean-Jacques Lévy. *Réductions correctes et optimales dans le lambda-calcul.* PhD thesis, Université de Paris 7, 1978.

23 Ralph Loader. Notes on simply typed lambda calculus. Technical Report ECS-LFCS-98-381, University of Edinburgh, 1998.

24 Paul-André Melliès. Axiomatic rewriting theory I: A diagrammatic standardization theorem. In Aart Middeldorp, Vincent van Oostrom, Femke van Raamsdonk, and Roel C. de Vrijer, editors, *Processes, Terms and Cycles: Steps on the Road to Infinity, Essays Dedicated to Jan Willem Klop, on the Occasion of His 60th Birthday*, volume 3838 of *Lecture Notes in Computer Science*, pages 554–638. Springer, 2005. `doi:10.1007/11601548_23`.

25 Peter Møller Neergaard and Morten Heine Sørensen. Conservation and uniform normalization in lambda calculi with erasing reductions. *Inf. Comput.*, 178(1):149–179, 2002. `doi:10.1006/inco.2002.3153`.

26 Tobias Nipkow. Higher-order critical pairs. In *Proceedings 1991 Sixth Annual IEEE Symposium on Logic in Computer Science*, pages 342–343. IEEE Computer Society, 1991.

27 Jan von Plato. Gentzen's proof of normalization for natural deduction. *Bulletin of Symbolic Logic*, 14(2):240–257, 2008.

28 Dag Prawitz. *Natural deduction: a proof-theoretical study.* PhD thesis, Almqvist & Wiksell, 1965.

29 Morten Heine Sørensen and Pawel Urzyczyn. *Lectures on the Curry-Howard isomorphism*, volume 149. Elsevier, 2006.

30 William W. Tait. A realizability interpretation of the theory of species. In Rohit Parikh, editor, *Logic Colloquium*, pages 240–251, Berlin, Heidelberg, 1975. Springer Berlin Heidelberg.

31 Terese. *Term Rewriting Systems*, volume 55 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2003.