Derivational Complexity and Context-Sensitive Rewriting

Salvador Lucas[1]

Abstract
Context-sensitive rewriting is a restriction of rewriting where reduction steps are allowed on specific arguments $\mu(f) \subseteq \{1, \ldots, k\}$ of k-ary function symbols f only. Terms which cannot be further rewritten in this way are called μ-normal forms. For left-linear term rewriting systems (TRSs), the so-called *normalization via μ-normalization* procedure provides a systematic way to obtain normal forms by the stepwise computation and combination of intermediate μ-normal forms. In this paper, we show how to obtain bounds on the *derivational complexity* of computations using this procedure by using bounds on the derivational complexity of context-sensitive rewriting. Two main applications are envisaged: Normalization via μ-normalization can be used with non-terminating TRSs where the procedure still terminates; on the other hand, it can be used to improve on bounds of derivational complexity of terminating TRSs as it discards many rewritings.

Keywords Derivational complexity · Term rewriting · Termination.

1 Introduction

In term rewriting [5] and rewriting-based languages, including lazy functional languages like Haskell [19], the normal form t of expressions s is often achieved by first obtaining a *head-normal form* u, i.e., a term which does not rewrite to a redex (i.e., there is no rule $\ell \rightarrow r$ such that $u \rightarrow^* \sigma(\ell)$ for some substitution σ). Therefore, such a term u has a *stable* root symbol f which *cannot* be changed by further rewritings. Thus, we can write $u = f(s_1, \ldots, s_k)$ for some terms $s_1, \ldots, s_k$ and recursively continue the normalization process starting from $s_1, \ldots, s_k$ to obtain normal forms $t_1, \ldots, t_k$ which are then used to finally return $t = f(t_1, \ldots, t_k)$. This procedure, though, has an important drawback: It is *undecidable* whether s *has* (or u *is*) a head-normal form [44, Section 8]. In context-sensitive rewriting (*CSR* [40]), the *normalization via μ-normalization* procedure ([40, Section 9.3], $N\mu N$ in the following) plays a similar role and can be used in programming languages like CafeOBJ [20], Maude [11] or OBJ [24], which permit the use of syntactic replacement

✉ Salvador Lucas
 slucas@dsic.upv.es

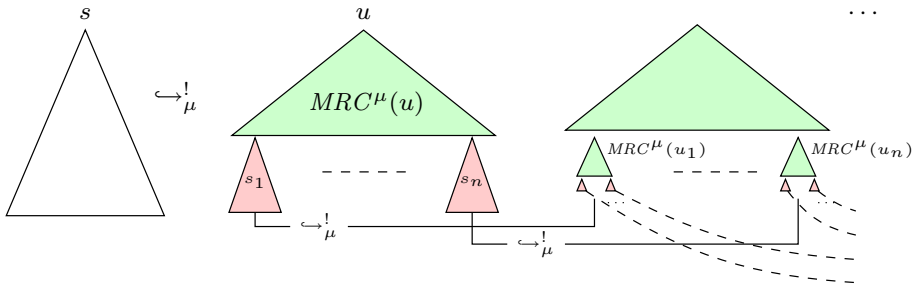1 DSIC & VRAIN, Universitat Politècnica de València, Valencia, Spain

Fig. 1 Normalization via μ-normalization

restrictions *à la* CSR as follows. Each k-ary symbol f of the signature is given replacement restrictions by means of a replacement map μ which defines a subset $\mu(f) \subseteq \{1, \ldots, k\}$ of (*active*) arguments of f which can be rewritten; rewriting is *forbidden* on arguments $i \notin \mu(f)$ (often called *frozen*). Such a restriction is top-down propagated to the structure of terms s so that we distinguish between *active* and frozen *positions* of subterms, with the topmost position always active. In *CSR* only subterms at active positions can be rewritten (denoted $s \hookrightarrow t$).

Example 1 *(Running example I)* Consider the following variant $\mathcal{R}$ of a TRS in [40, Example 8.20]:

$$a \to h(a) \tag{1}$$
$$h(x) \to b \tag{2}$$
$$f(x) \to g(a) \tag{3}$$
$$f(x) \to g(b) \tag{4}$$

with $\mu(f) = \mu(h) = \varnothing$ and $\mu(g) = \{1\}$. We have $\underline{a} \hookrightarrow \underline{h(a)} \hookrightarrow b$. However, $h(\underline{a}) \not\hookrightarrow h(h(a))$, as subterm a is not active in $h(a)$.

Terms which cannot be further rewritten using *CSR* are called μ-*normal forms*. In general, they can contain reducible terms, i.e., they are *not* normal forms.

Example 2 *(Running example II)* Consider $\mathcal{R} = \{app(app(x, y), z) \to app(x, app(y, z))\}$ [46, Example 1] together with $\mu(app) = \{1\}$. We have

$$s = \underline{app(app(x_1, app(x_2, x_3)), x_4)} \hookrightarrow app(x_1, app(app(x_2, x_3), x_4)) = u \tag{5}$$

where u is a μ-normal form which is *not* a normal form.

In $N\mu N$, μ-normal forms play the role of head-normal forms: In order to obtain the normal form t of a term s, we first obtain a μ-normal form u of s and then *jump* into the maximal (in size) *frozen* subterms of u, which are below the so-called *maximal replacing context* $MRC^\mu(u)$ of u, to continue the normalization. This is sketched in Fig. 1. The good point of μ-normal forms is that they are decidable (see [40, Sect. 6]).

Remark 1 *(Normalization via μ-normalization in* Maude*)* Since replacement maps can be used in Maude (see [40, Section 4.3]), an implementation of $N\mu N$ in Full Maude [11, Chapter 18] wyas developed in [16]. In [41, Sect. 10.2] an implementation using Maude's strategy language [17,54] is described, see

The study of the *length* of rewriting computations (i.e., the number of rewriting steps) can be understood as an *efficiency* measure regarding the 'time' they take. Given a terminating TRS $\mathcal{R}$ and a term s, the *derivational height* of s (written $\mathsf{dh}(s, \rightarrow)$) is defined as the maximum length of rewriting sequences $s = s_1 \rightarrow s_2 \rightarrow \cdots \rightarrow s_m$ starting from s. Then, the *derivational complexity*, $\mathsf{dc}_{\mathcal{R}}(n)$, of $\mathcal{R}$ is the maximum derivational height of terms s of size $|s|$ at most n, i.e., $|s| \leq n$. After Hofbauer and Lautemann's pioneering work [33], the use of *termination proofs* to obtain *bounds* on $\mathsf{dc}_{\mathcal{R}}(n)$ has been investigated by several authors focusing on the use of polynomials [7–9,51], matrices [18,46,48,58], recursive path orderings [10,32,59], dependency pairs [29,45], etc. Also, a number of tools have been developed to automatically (try to) obtain derivational complexity bounds for TRSs, see [23] and the references therein.

1.1 Contributions of the Paper

In this paper, we investigate derivational complexity of *CSR* and derivational complexity of *NμN*, also in connection with derivational complexity of rewriting. It is obvious that *CSR* can be used to avoid reductions. Furthermore, under some assumptions it is able to simulate rewriting computations and compute normal forms (see [40, Section 5.2]). In this case, *CSR* can be used to put bounds on the length of normalizing sequences when *CSR* can be taken as a rewriting strategy.

Example 3 *(Running Example III)* Consider the following TRS $\mathcal{R}$ [4, Example 3.23]:

$$\mathsf{f}(0, y) \rightarrow 0 \tag{6}$$

$$\mathsf{f}(\mathsf{s}(x), y) \rightarrow \mathsf{f}(\mathsf{f}(x, y), y) \tag{7}$$

As discussed in [4], $\mathcal{R}$ cannot be proved terminating by using a simplification ordering [14], in particular polynomial interpretations over the naturals or (variants of) recursive path orderings. Matrix interpretations also fail to prove $\mathcal{R}$ terminating. When using AProVE [22] to obtain derivational complexity bounds for $\mathcal{R}$, a linear lower bound is given, but no upper bound is provided (reported as $\mathsf{BOUNDS}(\mathsf{n}^1, \mathsf{INF})$). However, when considering $\mu(\mathsf{f}) = \mu(\mathsf{s}) = \{1\}$, it is possible to prove that *CSR* suffices to compute any normal form of terms s. Furthermore, we show that the length of context-sensitive computations is bounded by $O(n)$, see Example 19.

The ability of *CSR* to 'reinforce' termination can be used to guarantee that *CSR* or *NμN* terminates even if $\mathcal{R}$ is non-terminating. Thus, we can investigate derivational complexity for such non-terminating TRSs.

Example 4 Although $\mathcal{R}$ in Example 1 is not terminating (due to rule (1)), it is normalizing (i.e., every term has a normal form) and for all ground terms t *CSR* is able to obtain the normal form of t (see Example 10 and Corollary 1). Also, *CSR* terminates for all terms (see Example 14), and we actually obtain a *constant* bound $O(1)$ on the derivational complexity of *CSR*.

In some cases, though, *CSR* is not able to directly obtain normal forms and often stops yielding a μ-normal form (see Example 2). In this case, *NμN* can be used to approximate derivational complexity of terminating TRSs by using derivational complexity of *CSR*, often obtaining 'improved' complexity bounds.

Example 5 *(Running example IV)* Consider the TRS $\mathcal{R}$ in [55, Example 8]:

$$x \wedge (y \vee z) \rightarrow (x \wedge y) \vee (x \wedge z) \tag{8}$$

$$\neg x \vee y \rightarrow x \supset y \tag{9}$$

$$(\neg x) \supset (\neg y) \rightarrow y \supset (y \wedge x) \tag{10}$$

Steinbach proved $\mathcal{R}$ polynomially terminating[1] by using a polynomial interpretation with quadratic polynomials (see Example 11). According to [33], a doubly exponential bound $2^{2^{O(n)}}$ would be given on $\mathsf{dc}_{\mathcal{R}}(n)$. If the replacement map $\mu(\wedge) = \{2\}$ and $\mu(f) = \{1, \ldots, ar(f)\}$ for any other symbol f is used, we can improve it (for $N\mu N$) to $2^{O(n^2)}$ (Example 39).

After some general preliminaries in Sect. 2 and a brief summary of context-sensitive rewriting in Sect. 3, Sect. 4 provides the technical definitions on derivational complexity. Section 5 summarizes the bounds obtained from several termination techniques. Section 6 introduces derivational complexity of *CSR* and provides some results which permit the obtention of bounds on derivational complexity of *CSR*. Section 7 defines normalization via μ-normalization. Section 8 introduces an extension of *CSR* called *layered CSR* (*LCSR*). Then, $N\mu N$ is formulated as normalization using *LCSR*. A sufficient criterion for proving termination of $N\mu N$ for non-terminating TRSs is also given. Section 9 defines computational complexity of $N\mu N$ as the derivational complexity of *LCSR*. Section 10 discusses how to obtain bounds on derivational complexity of $N\mu N$. Section 11 shows how to obtain bounds on derivational complexity of $N\mu N$ from bounds on derivational complexity of *CSR*. Section 12 discusses some related work. Section 13 concludes.

2 Preliminaries

This section collects some definitions and notations about term rewriting. More details and missing notions can be found in [5,50,57]. In the following, $\mathcal{P}(A)$ denotes the power set of a set A. The cardinality of a finite set A is denoted as $|A|$.

Given a binary relation $\mathsf{R} \subseteq A \times A$ on a set A, we often write $a \mathrel{\mathsf{R}} b$ instead of $(a, b) \in \mathsf{R}$. Given two relations $\mathsf{R}, \mathsf{R}' \subseteq A \times A$, their *composition* is defined by $\mathsf{R} \circ \mathsf{R}' = \{(a, b) \in A \times A \mid (\exists c)\ a \mathrel{\mathsf{R}} c \wedge c \mathrel{\mathsf{R}'} b\}$. Also, for all $n \in \mathbb{N}$, the *n*-fold composition R^n of R is defined by $\mathsf{R}^0 = \{(x, x) \mid x \in A\}$ and $\mathsf{R}^n = \mathsf{R} \circ \mathsf{R}^{n-1}$ if $n > 0$. The *transitive* closure of R is denoted by R^+, and its *reflexive and transitive* closure by R^*. The relation R is *finitely branching* if for all $a \in A$, the set $\{a \mathrel{\mathsf{R}} b \mid b \in A\}$ of direct successors of a is finite. An element $a \in A$ is *irreducible* (or an R-*normal form*), if there exists no b such that $a \mathrel{\mathsf{R}} b$; we say that b is an R-normal form of a (written $a \mathrel{\mathsf{R}^!} b$) if $a \mathrel{\mathsf{R}^*} b$ and b is an R-normal form. We also say that a is R-normalizing, and that a has an R-normal form. Also, R is *normalizing* if every $a \in A$ has an R-normal form. Given $a \in A$, if there is no infinite sequence $a = a_1 \mathrel{\mathsf{R}} a_2 \mathrel{\mathsf{R}} \cdots \mathrel{\mathsf{R}} a_n \mathrel{\mathsf{R}} \cdots$, then a is R-*terminating*; R is *terminating* (or *well-founded*[2]); if a is R-terminating for all $a \in A$. We say that R is *confluent* if, for every $a, b, c \in A$, whenever $a \mathrel{\mathsf{R}^*} b$ and $a \mathrel{\mathsf{R}^*} c$, there exists $d \in A$ such that $b \mathrel{\mathsf{R}^*} d$ and $c \mathrel{\mathsf{R}^*} d$.

Throughout the paper, $\mathcal{X}$ denotes a countable set of variables and $\mathcal{F}$ denotes a signature, i.e., a set of function symbols $f, g \ldots$, each having a fixed arity $ar(f)$. The set of terms built

---

[1] Rule (10) in [55] is $(\neg x) \supset (\neg y) \rightarrow y \supset (x \wedge y)$, but the proof still works (see Example 11).

[2] See [50, Definition 2.1.1] and the paragraph below this definition for a clarifying discussion about the use of 'well-founded' and 'terminating' in mathematics and computer science.

from $\mathcal{F}$ and $\mathcal{X}$ is $\mathcal{T}(\mathcal{F}, \mathcal{X})$. The set of variables in a term t is denoted $Var(t)$. Let $\Box \notin \mathcal{F} \cup \mathcal{X}$ be a special constant symbol. A *context* is a term in $\mathcal{T}(\mathcal{F} \cup \{\Box\}, \mathcal{X})$. If $C[, \ldots,]$ is a context with n occurrences of $\Box$, then $C[t_1, \ldots, t_n]$ is the result of replacing the occurrences of $\Box$ with $t_1, \ldots, t_n$ from left to right. A context containing precisely one occurrence of $\Box$ is denoted by $C[\,]$. A term is said to be *linear* if no variable occurs more than once in t. Terms are viewed as labeled trees in the usual way. *Positions* $p, q, \ldots$ are sequences of positive natural numbers used to address subterms of t. We denote the *empty* sequence by Λ. Given positions p, q, we denote their *concatenation* as $p.q$. Positions are ordered by the standard *prefix ordering* $\leq$, i.e., $p \leq q$ if $q = p.p'$ for some position p'. Given a set of positions P, $minimal_{\leq}(P)$ is the set of *minimal* positions of P w.r.t. $\leq$. If p is a position, and Q is a set of positions, $p.Q = \{p.q \mid q \in Q\}$. We denote the empty chain by Λ. The set of positions of a term t is $\mathcal{P}os(t)$. Positions of non-variable symbols in t are denoted as $\mathcal{P}os_{\mathcal{F}}(t)$, and $\mathcal{P}os_{\mathcal{X}}(t)$ are the positions of variables. The subterm of t at position p is denoted as $t|_p$ and $t[s]_p$ is the term t with the subterm at position p replaced by s. The symbol labeling the root of t is denoted as $root(t)$. Given terms t and s, $\mathcal{P}os_s(t)$ denotes the set of positions of the subterm s in t, i.e., $\mathcal{P}os_s(t) = \{p \in \mathcal{P}os(t) \mid t|_p = s\}$. The *depth* δ_t of a term t is the number of nodes on a branch of the term tree of t of maximal length: $\delta_t = 1$ if t is a variable or a constant; $\delta_t = 1 + \max\{\delta_{t_1}, \ldots, \delta_{t_k}\}$ if $t = f(t_1, \ldots, t_k)$ [57, Definition 2.1.7]. The *size* of t, i.e., the number of symbols occurring in t is denoted $|t|$.

A *rewrite rule* is an ordered pair (ℓ, r), written $\ell \to r$, with $\ell, r \in \mathcal{T}(\mathcal{F}, \mathcal{X})$, $\ell \notin \mathcal{X}$ and $Var(r) \subseteq Var(\ell)$. The *left-hand side* (*lhs*) of the rule is ℓ and r is the *right-hand side* (*rhs*). A *term rewriting system* (TRS) is a pair $\mathcal{R} = (\mathcal{F}, R)$ where R is a set of rewrite rules. The set of *lhs*'s of $\mathcal{R}$ is denoted as $L(\mathcal{R})$. An instance $\sigma(\ell)$ of $\ell \in L(\mathcal{R})$ by a substitution σ is a *redex*. The set of redex positions in t is $\mathcal{P}os_{\mathcal{R}}(t)$. A TRS $\mathcal{R}$ is *left-linear* if for all $\ell \in L(\mathcal{R})$, ℓ is a linear term. A rule $\ell \to r$ is *collapsing* if $r \in \mathcal{X}$; it is *duplicating* if $|\mathcal{P}os_x(\ell)| < |\mathcal{P}os_x(r)|$ for some variable x. A TRS $\mathcal{R}$ is collapsing (resp. duplicating) if it contains a collapsing (resp. duplicating) rule. Given $\mathcal{R} = (\mathcal{F}, R)$, we consider $\mathcal{F}$ as the disjoint union $\mathcal{F} = \mathcal{C} \uplus \mathcal{D}$ of symbols $c \in \mathcal{C}$, called *constructors* and symbols $f \in \mathcal{D}$, called *defined functions*, where $\mathcal{D} = \{root(\ell) \mid \ell \to r \in R\}$ and $\mathcal{C} = \mathcal{F} - \mathcal{D}$. Then, $\mathcal{T}(\mathcal{C}, \mathcal{X})$ (resp. $\mathcal{T}(\mathcal{C})$) is the set of (ground) *constructor terms*.

Given a TRS $\mathcal{R}$, a term $s \in \mathcal{T}(\mathcal{F}, \mathcal{X})$ rewrites to t at position p, written $s \xrightarrow{p}_{\mathcal{R}} t$ (or just $s \to t$), if $s|_p = \sigma(\ell)$ and $t = s[\sigma(r)]_p$, for some rule $\ell \to r$ in $\mathcal{R}$, $p \in \mathcal{P}os(s)$ and substitution σ. A TRS $\mathcal{R}$ is *terminating* if $\to_{\mathcal{R}}$ is terminating. A term s is a *head-normal form* (or root-stable) if there is no redex t such that $s \to_{\mathcal{R}}^* t$. A term is said to be *root-normalizing* if it has a root-stable reduct. The $\to_{\mathcal{R}}$-normal forms of $\mathcal{R}$ are simply called *normal* forms; $NF_{\mathcal{R}}$ is the set of normal forms of $\mathcal{R}$. A term is said to be normalizing if it is $\to_{\mathcal{R}}$-normalizing. A TRS is normalizing if all terms are.

3 Context-Sensitive Rewriting

In this section, we provide some background on *CSR*, following [40]. Given a signature $\mathcal{F}$, a *replacement map* is a mapping $\mu : \mathcal{F} \to \wp(\mathbb{N})$ satisfying that, for all symbols f in $\mathcal{F}$, $\mu(f) \subseteq \{1, \ldots, ar(f)\}$. The set of replacement maps for the signature $\mathcal{F}$ is $M_{\mathcal{F}}$ (or $M_{\mathcal{R}}$ for a TRS $\mathcal{R} = (\mathcal{F}, R)$). Replacement maps are compared as follows: $\mu \sqsubseteq \mu'$ if for all $f \in \mathcal{F}$, $\mu(f) \subseteq \mu'(f)$; we often say that μ is *more restrictive* than μ'. Extreme cases are $\mu_{\bot}$, which disallows replacements in all arguments: $\mu_{\bot}(f) = \varnothing$ for all $f \in \mathcal{F}$, and $\mu_{\top}$, which restricts no replacement: $\mu_{\top}(f) = \{1, \ldots, k\}$ for all k-ary symbols $f \in \mathcal{F}$. We
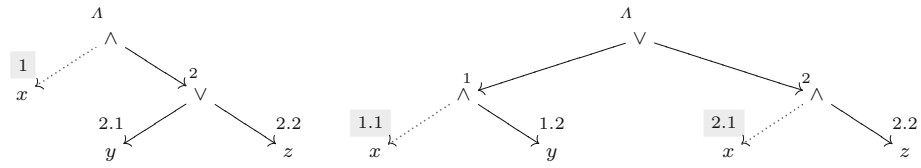
**Fig. 2** Active and frozen (in gray) positions of $\ell_{(8)}$ and $r_{(8)}$

say that a binary relation R on terms is μ-*monotonic* if for all k-ary symbols f, $i \in \mu(f)$, and terms $s_1, \ldots, s_k, t_i$, if s_i R t_i, then $f(s_1, \ldots, s_i, \ldots, s_k)$ R $f(s_1, \ldots, t_i, \ldots, s_k)$. The set $\mathcal{P}os^\mu(t)$ of μ-*replacing* (*or active*) *positions* of t is:

$$\mathcal{P}os^\mu(t) = \begin{cases} \{\Lambda\} & \text{if } t \in \mathcal{X} \\ \{\Lambda\} \cup \bigcup_{i \in \mu(f)} i.\mathcal{P}os^\mu(t_i) & \text{if } t = f(t_1, \ldots, t_k) \end{cases}$$

and $\overline{\mathcal{P}os^\mu}(t) = \mathcal{P}os(t) - \mathcal{P}os^\mu(t)$ is the set of *non-μ-replacing* (or *frozen*) positions.

Example 6 For $\mathcal{R}$ and μ as in Example 5, Fig. 2 depicts the active and frozen (in gray) positions of $\ell_{(8)}$ and $r_{(8)}$.

The frozen positions of t have a *frontier* set $\mathcal{F}r^\mu(t) = minimal_{\leq}(\overline{\mathcal{P}os^\mu}(t))$ with the active positions. The *maximal replacing context* $MRC^\mu(t) = t[\Box]_{\mathcal{F}r^\mu(t)}$ of t is the maximal prefix of t whose symbols occur at active positions only, i.e., $t = C[t_1, \ldots, t_n]$ for $C[, \ldots,] = MRC^\mu(t)$ and appropriate terms $t_1, \ldots, t_n$.
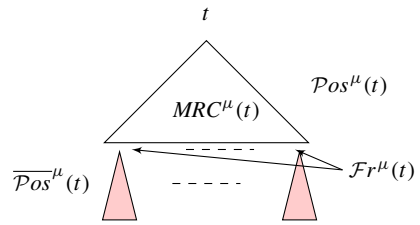
Given a term t, $\mathcal{V}ar^\mu(t)$ (resp. $\mathcal{V}ar^{/\!\!/}(t)$) is the set of variables occurring at active (resp. frozen) positions in t: $\mathcal{V}ar^\mu(t) = \{x \in \mathcal{V}ar(t) \mid \exists p \in \mathcal{P}os^\mu(t), x = t|_p\}$ and $\mathcal{V}ar^{/\!\!/}(t) = \{x \in \mathcal{V}ar(t) \mid \exists p \in \overline{\mathcal{P}os^\mu}(t), x = t|_p\}$. Note that variables in $\mathcal{V}ar^{/\!\!/}(t)$ could also be in $\mathcal{V}ar^\mu(t)$ and vice versa. For instance, $\mathcal{V}ar^\mu(r_{(10)}) = \{x, y\}$ and $\mathcal{V}ar^{/\!\!/}(t)(r_{(10)}) = \{y\}$. Also, we write $s \rhd_{/\!\!/} t$ if t is a frozen subterm of s.

Context-sensitive rewriting (*CSR*) is the restriction of rewriting obtained when a replacement map μ is used to specify the redex positions that can be contracted.

Definition 1 (Context-sensitive rewriting) Let $\mathcal{R}$ be a TRS, $\mu \in M_\mathcal{R}$ and s and t be terms. Then, s μ-*rewrites* to t, written $s \overset{p}{\hookrightarrow}_{\mathcal{R},\mu} t$ (or $s \hookrightarrow_{\mathcal{R},\mu} t$, $s \hookrightarrow_\mu t$, or even $s \hookrightarrow t$), if $s \overset{p}{\to}_\mathcal{R} t$ and p is active in s (i.e., $p \in \mathcal{P}os^\mu(s)$). We often use $\to_{/\!\!/}$ to denote rewriting at *frozen* positions: $\to_{/\!\!/} = \to - \hookrightarrow_\mu$.

If $\hookrightarrow_{\mathcal{R},\mu}$ is confluent (resp. terminating), we say that $\mathcal{R}$ is μ-*confluent* (resp. μ-*terminating*). The $\hookrightarrow_{\mathcal{R},\mu}$-normal forms are called μ-*normal forms* and $\mathsf{NF}^\mu_\mathcal{R}$ is the set of μ-normal forms of $\mathcal{R}$. The *canonical replacement map* $\mu^{can}_\mathcal{R}$ of a TRS $\mathcal{R}$ is *the most restrictive replacement map μ ensuring that the non-variable subterms of the left-hand sides ℓ of the rules $\ell \to r$ of $\mathcal{R}$ are active*, i.e., $\mathcal{P}os_\mathcal{F}(\ell) \subseteq \mathcal{P}os^\mu(\ell)$. Given a TRS $\mathcal{R}$, we let $CM_\mathcal{R} = \{\mu \in M_\mathcal{R} \mid \mu^{can}_\mathcal{R} \sqsubseteq \mu\}$ be the set of replacement maps that are equal or *less restrictive* than the canonical replacement map. If $\mu \in CM_\mathcal{R}$, we also say that μ is *a* canonical replacement map for $\mathcal{R}$; if μ is exactly $\mu^{can}_\mathcal{R}$, we will speak about *the* canonical replacement map of $\mathcal{R}$.

Example 7 The canonical replacement map $\mu_{\mathcal{R}}^{can}$ for $\mathcal{R}$ in Example 1 is $\mu_{\perp}$. Thus, μ in Example 1 is *not* the canonical replacement map (as $\mu(\mathsf{c}) = \{1\}$), but $\mu \in CM_{\mathcal{R}}$, i.e., it is *a canonical replacement map*.

For TRSs $\mathcal{R}$ and *canonical* replacement maps $\mu \in CM_{\mathcal{R}}$, we often say that $\hookrightarrow_{\mathcal{R},\mu}$ performs *canonical* CSR [40, Sect. 5]. Canonical *CSR* is useful in (head) normalization and infinitary normalization with left-linear TRSs [40, Sect. 9].

4 Derivational Complexity in Term Rewriting

The following definition formalizes the standard notion of derivational complexity. The definition is given for abstract relations R rather than TRSs.

Definition 2 (Derivational complexity) Let R be a well-founded and *finitely branching* relation, and t be a term. Then, the *derivational height* of t is

$$\mathsf{dh}(t, \mathsf{R}) = \max\{n \mid (\exists u)\, t\, \mathsf{R}^n\, u\}$$

and, for $n \in \mathbb{N}$, the *derivational complexity* of R is

$$\mathsf{dc}_{\mathsf{R}}(n) = \max\{\mathsf{dh}(t, \mathsf{R}) \mid |t| \leq n\}$$

Obviously, $\mathsf{dh}(t, \mathsf{R})$, hence $\mathsf{dc}_{\mathsf{R}}(n)$, is defined only if R is well founded. Requiring that R is also *finitely branching*, though, is often necessary.

Example 8 Consider the terminating TRS $\mathcal{R} = \{a \to i \mid i \in \mathbb{N}\} \cup \{i + 1 \to i \mid i \in \mathbb{N}\}$. Note that $\to_{\mathcal{R}}$ is not finitely branching. In particular, $\mathsf{dh}(a, \to_{\mathcal{R}})$ is not defined because for all $n \in \mathbb{N}$, a admits a rewrite sequence $a \to n \to^* 0$ of length $n + 1$.

In this paper, we deal with TRSs $\mathcal{R} = (\mathcal{F}, R)$ and (subsets of) the rewrite relation $\to_{\mathcal{R}}$. Thus, along the paper we assume $\to_{\mathcal{R}}$ to be finitely branching (thus, $\hookrightarrow_{\mathcal{R},\mu}$ also is). A sufficient condition making $\to_{\mathcal{R}}$ finitely branching is requiring R to be *finite*.[3] This is often the case in most practical uses.

Notation 1 *In the following, when dealing with TRSs $\mathcal{R}$ and the rewrite relation $\to_{\mathcal{R}}$, we use $\mathsf{dc}_{\mathcal{R}}$ instead of $\mathsf{dc}_{\to_{\mathcal{R}}}$.*

4.1 Use of Ground Terms in Term Rewriting

For TRSs $\mathcal{R} = (\mathcal{F}, R)$ whose signature $\mathcal{F}$ contains a constant symbol a, we have $\mathcal{T}(\mathcal{F}) \neq \varnothing$. In this case, by closedness of term rewriting with respect to substitution application (see, e.g., [5, Definition 4.2.2]), every rewrite sequence $s_1 \to_{\mathcal{R}} s_2 \to_{\mathcal{R}} \cdots \to_{\mathcal{R}} s_n$ has a corresponding (ground) rewrite sequence $s_1' \to_{\mathcal{R}} s_2' \to_{\mathcal{R}} \cdots \to_{\mathcal{R}} s_n'$ of the same length which is obtained by replacing, for all $1 \leq i \leq n$, all variable occurrences in s_i by a to obtain s_i'. Thus, $\mathsf{dh}_{\mathcal{R}}$ remains unchanged if only *ground* terms are considered, i.e., we have

$$\mathsf{dc}_{\mathcal{R}}(n) = \max\{\mathsf{dh}(t, \to_{\mathcal{R}}) \mid t \in \mathcal{T}(\mathcal{F}), |t| \leq n\} \tag{11}$$

If $\mathcal{F}$ contains no constant symbol, we can add a fresh constant $\bullet$ to obtain $\mathcal{F}^{\bullet} = \mathcal{F} \cup \{\bullet\}$. It is also well known that this does not affect termination of $\mathcal{R}$ [56]. Given a term $t \in \mathcal{T}(\mathcal{F}, \mathcal{X})$,

---

[3] There are infinite (and terminating) TRSs which are finitely branching. Consider, for instance, the TRS $\mathcal{R}$ with set of rules $R = \{i + 1 \to i \mid i \in \mathbb{N}\}$.

a ground term $t^\bullet \in \mathcal{T}(\mathcal{F}^\bullet)$ is obtained by replacing all variables in t by $\bullet$. We have the following.

Proposition 1 *Let* $\bullet$ *be a constant symbol,* $\mathcal{R} = (\mathcal{F}, R)$ *be a TRS where* $\bullet \notin \mathcal{F}$ *and* $\mathcal{R}^\bullet = (\mathcal{F}^\bullet, R)$. *Then, for all* $n \in \mathbb{N}$, $dc_\mathcal{R}(n) = dc_{\mathcal{R}^\bullet}(n)$.

Proof Let $s \in \mathcal{T}(\mathcal{F}, \mathcal{X})$ and $s = s_1 \to_\mathcal{R} s_2 \to_\mathcal{R} \cdots \to_\mathcal{R} s_m$ for some terms $s_1, \ldots, s_m$. Since rewriting is closed under substitutions, for all $1 \leq i \leq m$, we can instantiate each variable in s_i with $\bullet$ to obtain $s^\bullet = s_1^\bullet \to_\mathcal{R} s_2^\bullet \to_\mathcal{R} \cdots \to_\mathcal{R} s_m^\bullet$. Since the rules in $\mathcal{R}$ and $\mathcal{R}^\bullet$ coincide, we have $s_1^\bullet \to_{\mathcal{R}^\bullet} s_2^\bullet \to_{\mathcal{R}^\bullet} \cdots \to_{\mathcal{R}^\bullet} s_m^\bullet$. Hence, $\mathsf{dh}(s, \to_\mathcal{R}) \leq \mathsf{dh}(s^\bullet, \to_{\mathcal{R}^\bullet})$. Since the signature of $\mathcal{R}^\bullet$, i.e., $\mathcal{F}^\bullet$, contains a constant symbol $\bullet$, and for all terms $t \in \mathcal{T}(\mathcal{F}, \mathcal{X})$, $|t| = |t^\bullet|$, by using (11), we conclude $dc_\mathcal{R}(n) \leq dc_{\mathcal{R}^\bullet}(n)$.

Now let $s = s_1 \to_{\mathcal{R}^\bullet} s_2 \to_{\mathcal{R}^\bullet} \cdots \to_{\mathcal{R}^\bullet} s_m$, for some $s \in \mathcal{T}(\mathcal{F}^\bullet, \mathcal{X})$. Since rules in $\mathcal{R}$ and $\mathcal{R}^\bullet$ coincide, we have $s_1 \to_\mathcal{R} s_2 \to_\mathcal{R} \cdots \to_\mathcal{R} s_m$. Now replace each occurrence of $\bullet$ in terms $s_i, 1 \leq i \leq m$ by a variable $x \in \mathcal{X}$ to obtain a term $s_i' \in \mathcal{T}(\mathcal{F}, \mathcal{X})$. Since $\bullet$ does not occur in any rule of $\mathcal{R}$, we have $s_1' \to_\mathcal{R} s_2' \to_\mathcal{R} \cdots \to_\mathcal{R} s_m'$. Thus, we have $\mathsf{dh}(s, \to_{\mathcal{R}^\bullet}) \leq \mathsf{dh}(s', \to_\mathcal{R})$. Since for all terms $t \in \mathcal{T}(\mathcal{F}^\bullet, \mathcal{X})$, $|t| = |t'|$, we obtain $dc_{\mathcal{R}^\bullet}(n) \leq dc_\mathcal{R}(n)$. Therefore, we obtain $dc_\mathcal{R}(n) = dc_{\mathcal{R}^\bullet}(n)$, as desired. $\square$

Remark 2 *(Non-empty set of ground terms)* Proposition 1 means that, whether the signature $\mathcal{F}$ of a TRS $\mathcal{R} = (\mathcal{F}, R)$ contains constant symbols or not, we can add a fresh constant $\bullet$ to $\mathcal{F}$ without changing the complexity bounds. In the following, it is often important to be able to guarantee that $\mathcal{T}(\mathcal{F}) \neq \varnothing$ holds, i.e., $\mathcal{F}$ contains a constant symbol. Thus, as a kind of *preprocessing*, we can systematically use $\mathcal{R}^\bullet = (\mathcal{F}^\bullet, R)$ instead of $\mathcal{R}$.

4.2 Derivational Complexity and Strategies

A *non-deterministic one-step rewriting strategy* is a function S that assigns a non-empty subset $S(t) \subseteq \mathcal{P}os_\mathcal{R}(t)$ of redexes to every reducible term t [6,44]. We write $s \to_S t$ if $s \xrightarrow{p} t$ and $p \in S(t)$. A strategy S is *normalizing* if for all normalizing terms t no infinite $\to_S$-sequence starts from t. The notions and notations in Definition 2 apply to normalizing strategies S for normalizing TRSs $\mathcal{R}$ to yield $\mathsf{dh}(t, \to_S)$ for normalizing terms t and $dc_{\to_S}(n)$ (or just $dc_S(n)$) as the derivational height and complexity of computations using a strategy S. Obviously, for terminating TRSs $\mathcal{R}$ and arbitrary strategies S,

$$dc_S(n) \leq dc_\mathcal{R}(n) \tag{12}$$

Remark 3 *(Strategies in derivational complexity analysis)* There are non-terminating, but normalizing TRSs $\mathcal{R}$ (where $dc_\mathcal{R}$ is not defined) for which a given strategy S is still normalizing and dc_S defined. This is a positive aspect of analyzing derivational complexity of strategies. Another aspect is that in real-life programming languages a particular strategy S (rather than full rewriting) is implemented and used. In this case, dc_S is a *more realistic* bound for practical uses than $dc_\mathcal{R}$.

Derivational complexity (bounds) for specific rewriting strategies like *innermost* rewriting (where only redexes not containing other redexes are contracted) have been investigated [29,49]. The obtained bounds often differ from full rewriting.

Example 9 If computations with $\mathcal{R}$ in Example 3 are restricted to *innermost rewriting*, AProVE yields BOUNDS(n^1, n^2), i.e., a *linear* lower bound (displayed as n^1 in the first argument) and a *quadratic* upper bound (n^2 in the second argument), instead of BOUNDS(n^1, INF), where a linear lower bound was found but no finite upper bound could be obtained.

Although, in general, *CSR* is *not* a rewriting strategy (as it is unable to reduce terms in μ-normal form which are not normal forms, see [40, Remark 1.2] for a discussion), in some cases we can rely on the following result to faithfully use bounds on $\mathsf{dc}_\mu(n)$ as bounds for (a subclass of) normalizing sequences for TRSs.

Theorem 1 [40, Theorem 5.8] *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_\mathcal{R}$. Let s, t be terms such that $\mathcal{P}os(t) = \mathcal{P}os^\mu(t)$. Then, $s \rightarrow^* t$ if and only if $s \hookrightarrow_\mu^* t$.*

Given a TRS $\mathcal{R} = (\mathcal{F}, R)$ with $\mathcal{F} = \mathcal{C} \uplus \mathcal{D}$, a defined symbol $f \in \mathcal{D}$ is called *completely defined* if no ground normal form contains such a symbol; $\mathcal{R}$ is *completely defined* if all defined symbols are completely defined. We have the following.

Corollary 1 *Let $\mathcal{R}$ be a completely defined left-linear TRS and $\mu \in CM_\mathcal{R}$ be such that, for all constructor symbols $c \in \mathcal{C}$, $\mu(c) = \mu_\top(c)$. Let s and t be terms, with t ground. Then, $s \rightarrow^! t$ if and only if $s \hookrightarrow_\mu^! t$.*

Proof Since $\mathcal{R}$ is completely defined, t is a constructor term and $\mathcal{P}os^\mu(t) = \mathcal{P}os(t)$. By Theorem 1, the conclusion follows.

Example 10 Note that $\mathcal{R}$ in Example 1, with defined symbols a, f, and h is completely defined as there is a rule that applies to the root of any term a, $f(t)$ and $h(t)$ for all terms t. Also, it is not difficult to prove, by induction on the structure of the ground normal form t, that $\mathcal{R}$ in Example 3 is completely defined.

As a consequence of Corollary 1, for left-linear, completely defined TRSs $\mathcal{R}$ and $\mu \in CM_\mathcal{R}$ satisfying $\mu(c) = \mu_\top(c)$ for all constructor symbols c, the set of ground normal forms of $\mathcal{R}$ and the set of ground μ-normal forms of $\mathcal{R}$ coincide. Thus, dealing with ground terms t, *CSR* can be seen as a (non-deterministic) one-step rewriting strategy given by $S_{CSR}(t) = \mathcal{P}os_\mathcal{R}^\mu(t)$. It is often the case that $\mathsf{dc}_{S_{CSR}}$ can be used when $\mathsf{dc}_\mathcal{R}$ is undefined (Example 18), or improves on $\mathsf{dc}_\mathcal{R}$ as suggested by (12), see Example 19.

Remark 4 *(Reinforcing completely definedness?)* In view of the previous discussion, one could think of adding rules of the shape $f(x_1, \ldots, x_k) \rightarrow \bot$ (where $\bot$ is a fresh constant symbol) to each defined symbol $f \in \mathcal{D}$ so that $\mathcal{R}$ becomes completely defined whilst the maximal length of rewriting sequences in such a new TRS $\mathcal{R}_\bot$ barely changes, i.e., $\mathsf{dc}_\mathcal{R}$ and $\mathsf{dc}_{\mathcal{R}_\bot}$ (asymptotically) coincide. However, the set $\mathsf{NF}_{\mathcal{R}_\bot}$ of normal forms of $\mathcal{R}_\bot$ would in general lack some normal forms of $\mathcal{R}$ due to the addition of rules as above. In this situation, the use of Corollary 1 with $\mathcal{R}_\bot$ to obtain bounds on $\mathsf{dc}_\mathcal{R}$ using (12) with S_{CSR} based on $\mathcal{R}_\bot$ and μ would be *misleading* as many $\mathcal{R}$-normal forms would not be really approximated using *CSR*.

5 Termination Proofs and Complexity Bounds

As remarked above, derivational complexity bounds are often obtained from the method to obtain a termination proof. We summarize the ones we use in the paper, namely polynomial and matrix interpretations which are particular cases of the *semantic* approach to prove termination of rewriting often called termination by interpretation [60] and also [50, Sect. 5.2.1].

5.1 Termination by Interpretation

Given a signature $\mathcal{F}$, an $\mathcal{F}$-algebra $\mathcal{A}$ consists of a set (by abuse, also denoted $\mathcal{A}$) together with a mapping $f^{\mathcal{A}} : \mathcal{A}^k \to \mathcal{A}$, for each k-ary symbol $f \in \mathcal{F}$. Terms t are interpreted by induction on their structure by using valuation mappings $\alpha : \mathcal{X} \to \mathcal{A}$ to give meaning to variables inside as follows: (i) $[\alpha]_{\mathcal{A}}(x) = \alpha(x)$ if $x \in \mathcal{X}$ and (ii) $[\alpha]_{\mathcal{A}}(f(t_1, \ldots, t_k)) = f^{\mathcal{A}}([\alpha]_{\mathcal{A}}(t_1), \ldots, [\alpha]_{\mathcal{A}}(t_k))$. If $\mathcal{A}$ is supplied with a well-founded relation $\succ$ on $\mathcal{A}$, then $(\mathcal{A}, \succ)$ is called a *well-founded $\mathcal{F}$-algebra*. Mappings $f^{\mathcal{A}}$ are often required to be *monotone*, i.e., for all $i \in \{1, \ldots, k\}$ and $x_1, \ldots, x_k, y_i \in \mathcal{A}$, if $x_i \succ y_i$, then $f^{\mathcal{A}}(x_1, \ldots, x_i, \ldots, x_k) \succ f^{\mathcal{A}}(x_1, \ldots, y_i, \ldots, x_k)$. In this case, $(\mathcal{A}, \succ)$ is a *monotone well-founded $\mathcal{F}$-algebra*. A well-founded monotone algebra $(\mathcal{A}, \succ)$ induces a *reduction ordering* $\succ_{\mathcal{A}}$ on terms: $s \succ_{\mathcal{A}} t$ iff for all valuations α, $[\alpha]_{\mathcal{A}}(s) \succ [\alpha]_{\mathcal{A}}(t)$. We say that $(\mathcal{A}, \succ)$ is compatible with a TRS $\mathcal{R}$ if $\ell \succ_{\mathcal{A}} r$ for all $\ell \to r \in \mathcal{R}$. A TRS is terminating if and only if it admits a compatible non-empty well-founded monotone algebra [60, Proposition 1].

5.1.1 Polynomial Interpretations

The use of polynomials for proving termination of rewriting goes back to Manna and Ness [42] and Lankford [35] (see also [12,39] and the references therein). In a *polynomial interpretation* (PI) $\mathcal{A}$, the domain $\mathcal{A}$ is an infinite interval of natural or positive real numbers, and the well-founded relation $\succ$ on $\mathcal{A}$ is the usual order over the naturals $>_{\mathbb{N}}$ (in the first case) or the ordering $>_\delta$ over the reals in the second case, where δ is a positive number and $x >_\delta y$ if and only if $x - y \geq \delta$ [39]. Mappings $f^{\mathcal{A}}$ are obtained, for each $x_1, \ldots, x_k \in \mathcal{A}$ by evaluating a polynomial $\sum_{\vec{\alpha} \in \mathbb{N}^k} f_{\vec{\alpha}} x^{\vec{\alpha}}$, where, for each tuple $\vec{\alpha}$ of k natural numbers $\alpha_1, \ldots, \alpha_k \in \mathbb{N}$, the *coefficient* $f_{\vec{\alpha}}$ is a number multiplying the monomial $x^{\vec{\alpha}} = x_1^{\alpha_1} \cdots x_k^{\alpha_k}$ of the polynomial.

Remark 5 *(Terms as polynomials)* Terms t interpreted by using a polynomial interpretation $\mathcal{A}$ yield *polynomials* $t^{\mathcal{A}}$ obtained by using variable names (ranging now on the domain of the polynomial interpretation) instead of valuation mappings α.

Example 11 *(Running example IV—polynomial termination)* The TRS $\mathcal{R}$ in Example 5 is proved terminating by using a reduction ordering $\succ_{\mathcal{A}}$ obtained from a polynomial interpretation $\mathcal{A}$ as follows [55, Example 11][4]

$$\text{ff}^{\mathcal{A}} = 0 \qquad \text{tt}^{\mathcal{A}} = 0 \qquad \neg^{\mathcal{A}}(x) = x^2 + 6x + 9$$
$$x \wedge^{\mathcal{A}} y = xy + 4x + 3y + 12 \qquad x \vee^{\mathcal{A}} y = x + y + 8 \qquad x \supset^{\mathcal{A}} y = x + y + 8$$

For instance, for rule (10), we have:[5]

$$\ell_{(10)}^{\mathcal{A}} = ((\neg x) \supset (\neg y))^{\mathcal{A}} = (x^2 + 6x + 9) + (y^2 + 6y + 9) + 8$$
$$= x^2 + y^2 + 6x + 6y + 26$$
$$r_{(10)}^{\mathcal{A}} = (y \supset (y \wedge x))^{\mathcal{A}} = y + (yx + 4y + 3x + 12) + 8$$
$$= xy + 3x + 5y + 20$$

---

[4] Steinbach's interpretation is $[\neg](x) = x^2$, $x[\wedge]y = xy + x$ and $x[\vee]y = x[\supset]y = x + y + 2$ with interpretation domain $\mathbb{N}_{\geq 3}$. The interpretation in Example 11 is obtained from it by normalizing the intepretation domain to $\mathbb{N}$ using the translation $n \mapsto n + 3$ [12, Sect. 3.4].

[5] Rather than considering infinitely many valuations of the variables of the rules, we leave the variables as such and then compare the obtained polynomials as usual (see Remark 5).

The subtraction $\ell^{\mathcal{A}}_{(10)} - r^{\mathcal{A}}_{(10)}$ of both polynomials yields $x^2 + y^2 - xy + 3x + y + 6 = (x - y)^2 + xy + 3x + y + 6$, which is clearly positive for all $x, y \geq 0$. This proves $\succ_{\mathcal{A}}$ compatible with rule (10), i.e., $\ell_{(10)} \succ_{\mathcal{A}} r_{(10)}$ holds.

In this paper, we often restrict the attention to interpretations with simply shaped polynomials, like *linear* polynomials $f^{\mathcal{A}}(x_1, \ldots, x_k) = f_1 x_1 + \cdots + f_k x_k + f_0$, where $f_i \geq 1$ for all $1 \leq i \leq k$ (due to the monotonicity requirements), or even *strongly* linear polynomials $f^{\mathcal{A}}(x_1, \ldots, x_k) = x_1 + \cdots + x_k + f_0$. Also, for simplicity we restrict the attention to polynomials over the naturals, i.e., the interpretation domain is $\mathbb{N}$, and hence, the polynomial coefficients f_i are natural numbers as well.

5.1.2 Matrix Interpretations

In *matrix interpretations* (over the naturals) $\mathcal{A}$ [18], the interpretation domain is $\mathbb{N}^d$, the set of tuples (or vectors) $\mathbf{x}$ of d natural numbers. Mappings $f^{\mathcal{A}}$ are expressions $\sum_{i=1}^{k} F_i x_i + \mathbf{f}_0$, where $\mathbf{f}_0$ is a vector of d natural numbers and $F_1, \ldots, F_k$ are d-square matrices of natural numbers. Monotonicity of $f^{\mathcal{A}}$ is guaranteed if, for all $1 \leq i \leq k$, the top leftmost entry $(F_i)_{11}$ is positive, i.e., $(F_i)_{11} \geq 1$.

Remark 6 *(Terms as affine mappings)* As usual, terms t interpreted by using a matrix interpretation $\mathcal{A}$ yield *affine mappings* $t^{\mathcal{A}}$ which are obtained by just replacing the use of valuation mappings α by the variable name (ranging now on the domain of the matrix interpretation), i.e., $x^{\mathcal{A}} = \mathbf{x}$ and $f(t_1, \ldots, t_k)^{\mathcal{A}} = f^{\mathcal{A}}(t_1^{\mathcal{A}}, \ldots, t_k^{\mathcal{A}})$.

The following order $\succ$ is often used (see also [47]): for all $x_1, \ldots, x_d, y_1, \ldots, y_d \in \mathbb{N}$,

$$(x_1, \ldots, x_d) \succ (y_1, \ldots, y_d) \text{ iff } x_1 >_{\mathbb{N}} y_1 \text{ and } x_i \geq_{\mathbb{N}} y_i \text{ for all } 2 \leq i \leq d \quad (13)$$

In a *triangular matrix interpretation* (TMI), all matrices F_i are upper triangular and for all k-ary symbols $f \in \mathcal{F}$, $1 \leq i \leq k$ and $1 \leq j \leq d$, $(F_i)_{jj} \leq 1$ (but $(F_i)_{11} = 1$ due to the monotonicity requirement); also, for all $1 \leq j \leq d$, $A_{ij} = 0$ if $j > i$.

Remark 7 Matrix interpretations were generalized to admit *real* entries and interpretation domains which are tuples of *non-negative* real natural numbers [1]. The ordering $\succ_\delta$ is used instead of $>_{\mathbb{N}}$ for strict comparisons of tuple components in (13). See also [48], which extended TMIs to matrix interpretations over the reals.

5.2 Derivational Complexity Bounds from Termination Proofs

Table 1 summarizes the best (general) upper bounds reported for the different techniques we use in the paper. Except for (arbitrary) matrix interpretations, all considered techniques provide bounds given by elementary functions[6].

The different categories of asymptotic bounds can be ordered as follows:

$$O(n) \subseteq O(n^d) \subseteq 2^{O(n)} \subseteq 2^{2^{O(n)}} \subseteq \cdots \subseteq PR(n)$$

Bounds obtained from other methods, e.g., lexicographic or multiset orderings [10,32,59], Knuth–Bendix orderings [34] or dependency pairs [3,21], have been analyzed too [33,45,46, 48]. In general, they exceed those in Table 1.

---

[6] See [27, Sect. 2] and also [53]. Elementary functions are bounded by 2^n or 2^{2^n} or …

Table 1 Bounds on $dc_{\mathcal{R}}(n)$ from termination techniques

Reduction ordering $\succ_{\mathcal{A}}$ from	Upper bound on $dc_{\mathcal{R}}(n)$	References
Strongly linear PI	$O(n)$	[46, Theorem 6]
TMI of dimension d	$O(n^d)$	[46, Theorem 6]
Linear PI	$2^{O(n)}$	[33, Proposition 2.2.1]
PI	$2^{2^{O(n)}}$	[33, Proposition 2.1.2]
Matrix interpretation	Primitive recursive (*PR*)	[18, Sect. 9]

Example 12 *(Running example IV—complexity bound)* For $\mathcal{R}$ in Example 5, in view of Table 1 and according to the polynomial termination proof in Example 11, the best upper bound we can obtain is doubly exponential, i.e., $2^{2^{O(n)}}$.

Example 13 *(Running example II—termination and complexity bound)* Consider the TRS $\mathcal{R}$ in Example 2. A proof of termination is obtained by using the reduction ordering $\succ_{\mathcal{A}}$ induced by the following TMI $\mathcal{A}$ with domain $\mathbb{N}^2$ [46, Example 4]:

$$\mathsf{app}^{\mathcal{A}}(x) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} y + \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

For the rule $\ell \to r$ in $\mathcal{R}$, we have:

$$
\begin{aligned}
\ell^{\mathcal{A}} = \mathsf{app}(\mathsf{app}(x, y), z)^{\mathcal{A}} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\mathbf{y} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\mathbf{z} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\mathbf{y} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\mathbf{z} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} x_1 + 2x_2 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 + y_2 \\ y_2 \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} x_1 + 2x_2 + y_1 + y_2 + z_1 + 1 \\ x_2 + y_2 + z_2 + 2 \end{bmatrix} \\
r^{\mathcal{A}} = \mathsf{app}(x, \mathsf{app}(y, z))^{\mathcal{A}} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\mathbf{y} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\mathbf{z} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\mathbf{y} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\mathbf{z} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} x_1 + x_2 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 + y_2 \\ y_2 \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} x_1 + x_2 + y_1 + y_2 + z_1 \\ x_2 + y_2 + z_2 + 2 \end{bmatrix}
\end{aligned}
$$

For all $1 \le i \le 2$ and $x_i, y_i, z_i \in \mathbb{N}$, $x_1 + 2x_2 + y_1 + y_2 + z_1 + 1 > x_1 + x_2 + y_1 + y_2 + z_1$ for the first component of the obtained vectors, and $x_2 + y_2 + z_2 + 2 \ge x_2 + y_2 + z_2 + 2$ for the second one. Thus, $\ell^{\mathcal{A}} \succ_{\mathcal{A}} r^{\mathcal{A}}$, and $dc_{\mathcal{R}}(n) \in O(n^2)$ (see Table 1).

6 Derivational Complexity of Context-Sensitive Rewriting

According to Definition 2, given a finitely branching TRS $\mathcal{R}$ and $\mu \in M_{\mathcal{R}}$ such that $\mathcal{R}$ is μ-terminating, the *derivational height and derivational complexity of* CSR are defined as:

$$\mathsf{dh}_{\mathcal{R},\mu}(t) = \mathsf{dh}(t, \hookrightarrow_{\mathcal{R},\mu}) \qquad \text{and} \qquad \mathsf{dc}_{\mathcal{R},\mu}(n) = \mathsf{dc}_{\hookrightarrow_{\mathcal{R},\mu}}(n).$$

For signatures containing a constant symbol, as for ordinary term rewriting (see Sect. 4.1), since *CSR* is also closed under substitutions [40, Sect. 4], we can consider ground terms only, i.e., we have:

$$\mathsf{dc}_{\mathcal{R},\mu}(n) = \max\{\mathsf{dh}(t, \to_{\mathcal{R}}) \mid t \in \mathcal{T}(\mathcal{F}), |t| \leq n\} \tag{14}$$

Following the discussion in Sect. 4.1, we also have the following version of Proposition 1 for *CSR* (with an analogous proof using the fact that *CSR* is closed under substitutions).

Proposition 2 *Let $\bullet$ be a constant symbol, $\mathcal{R} = (\mathcal{F}, R)$ be a TRS where $\bullet \notin \mathcal{F}$, $\mu \in M_{\mathcal{R}}$, and $\mathcal{R}^{\bullet} = (\mathcal{F}^{\bullet}, R)$. Then, for all $n \in \mathbb{N}$, $\mathsf{dc}_{\mathcal{R},\mu}(n) = \mathsf{dc}_{\mathcal{R}^{\bullet},\mu}(n)$.*

In the following, we often use dh_{μ} and dc_{μ} (rather than $\mathsf{dh}_{\mathcal{R},\mu}$ and $\mathsf{dc}_{\mathcal{R},\mu}$). Since $\hookrightarrow_{\mu} \subseteq \to$ for all $\mu \in M_{\mathcal{R}}$, bounds on $\mathsf{dc}_{\mathcal{R}}(n)$ are also bounds on $\mathsf{dc}_{\mu}(n)$, i.e.,

Proposition 3 *For all terms s and $n \in \mathbb{N}$, $\mathsf{dh}_{\mu}(s) \leq \mathsf{dh}(s, \to_{\mathcal{R}})$ and $\mathsf{dc}_{\mu}(n) \leq \mathsf{dc}_{\mathcal{R}}(n)$.*

For left-linear, non-collapsing, and $\mu_{\perp}$-terminating TRSs $\mathcal{R}$, if $\mu_{\perp}$ is canonical, then the derivational height of all terms is bounded by the size of $\mathcal{R}$.

Proposition 4 *Let $\mathcal{R} = (\mathcal{F}, R)$ be a finite, left-linear, non-collapsing TRS such that $\mu_{\perp} \in CM_{\mathcal{R}}$. If $\mathcal{R}$ is $\mu_{\perp}$-terminating, then for all terms s and $n \in \mathbb{N}$, $\mathsf{dh}_{\mu_{\perp}}(s) \leq |R|$ and $\mathsf{dc}_{\mu_{\perp}}(n) \in O(1)$.*

Proof In the following, for $\mathcal{R} = (\mathcal{F}, R)$, we let $R_{cons} = \{\ell \to r \in R \mid root(r) \in \mathcal{C}_{\mathcal{R}}\}$ and call them *the constructor rules* of $\mathcal{R}$. First, note that, since $\mu_{\perp}$ permits no reduction on the arguments of any function symbol, every μ-reduction sequence $s_1 \overset{p_1}{\hookrightarrow} s_2 \overset{p_2}{\hookrightarrow} \cdots \overset{p_n}{\hookrightarrow} s_{n+1}$ satisfies $p_1 = \cdots = p_n = \Lambda$. We proceed by induction on the number n of rules in $\mathcal{R}$. First, note that, since $\mathcal{R}$ is left-linear and $\mu_{\perp} \in CM_{\mathcal{R}}$, all rules in R are of the form $f(x_1, \ldots, x_k) \to r$ where $x_1, \ldots, x_k$ are different variables and r is a non-variable term (due to non-collapsingness of $\mathcal{R}$). This means that every term $f(t_1, \ldots, t_k)$ is $\mu_{\perp}$-reducible if f is a defined symbol. Therefore, there must be at least one *constructor rule* $\ell \to r$ where $root(r) \in \mathcal{C}_{\mathcal{R}}$. Otherwise, the application of a rule could always be followed by the application of another rule (at the root position) and $\mathcal{R}$ would not be $\mu_{\perp}$-terminating. If $n = 1$, then the only rule in $\mathcal{R}$ must be a constructor rule. Thus, if $s \hookrightarrow s'$, then $root(s') = c$. Since c is a constructor symbol and $\mu_{\perp}(c) = \varnothing$, no further reduction is possible. Thus, $\mathsf{dh}(s) \leq 1$, as desired. If $n > 1$, then consider a μ-rewrite sequence $s = s_1 \hookrightarrow \cdots \hookrightarrow s_{n+1}$. If a constructor rule $\alpha : \ell \to r \in R_{cons}$ is used in the sequence, then it can be used only at the end of the sequence, and therefore, the sequence $s_1 \hookrightarrow s_n$ of length $n - 1$ can be obtained using a TRS $\mathcal{R}' = (\mathcal{F}', R')$ which is as $\mathcal{R}$ except the constructor rule, i.e., $R' = R - \{\alpha\}$. Note that $\mathcal{R}'$ is left-linear, non-collapsing and $\mu_{\perp}$-terminating as well. Note also that $\mathcal{C}_{\mathcal{R}} \subseteq \mathcal{C}_{\mathcal{R}'}$ but $\mathcal{C}_{\mathcal{R}'}$ could include $root(\ell)$, which is defined in $\mathcal{R}$ but may become constructor in $\mathcal{R}'$ due to the removal of α. By the induction hypothesis, $n - 1 \leq |R'| - 1$, and hence, $n \leq |R|$ as required. If no constructor rule is used in the sequence, then, again, we can obtain the same sequence using a TRS $\mathcal{R}' = (\mathcal{F}', R')$ where $R' = R - R_{cons}$. By the induction hypothesis, $n \leq |R'| < |R|$.

Remark 8 *(Application to recursive program schemes)* A recursive program scheme (RPS [13,43]) is a TRS consisting of (left-linear) rules $f(x_1, \ldots, x_k) \rightarrow r$, one per defined function symbol $f \in \mathcal{D}$ [57, Definition 3.4.7]. Clearly, $\mu_{\perp} \in CM_{\mathcal{R}}$ for all RPSs $\mathcal{R}$. Note that $\mathcal{R}$ in Example 1 is an RPS.

6.1 Proving Termination of *CSR*

Several techniques for proving termination of rewriting have been generalized to *CSR* [40, Sect. 7.1]. Tools like AProVE and MU- TERM [28] (which we use in the proofs below) can be used to prove termination of *CSR*. Termination of *CSR* is characterized by the μ-reduction orderings $>$, which are μ-monotonic, closed under substitutions, and well-founded relations on terms [61]. A TRS is μ-terminating if it admits a compatible μ-reduction ordering $>$. This implies

$$\hookrightarrow_{\mathcal{R},\mu} \subseteq > \tag{15}$$

as it is implicit in the proof of [61, Proposition 1]. Also, μ-reduction orderings can be obtained from μ-*monotone*, well-founded $\mathcal{F}$-algebras $(\mathcal{A}, \succ)$ where mappings $f^{\mathcal{A}}$ are required to be μ-monotone, i.e., for all $f \in \mathcal{F}$, $i \in \mu(f)$, and $x_1, \ldots, x_k, y_i \in \mathcal{A}$, if $x_i \succ y_i$, then $f^{\mathcal{A}}(x_1, \ldots, x_i, \ldots, x_k) \succ f^{\mathcal{A}}(x_1, \ldots, y_i, \ldots, x_k)$ [61, Sect. 3]. The different termination techniques considered in Table 1 (among others) have been adapted to *CSR*: polynomial orderings in [26,39]; matrix interpretations in [2].

Remark 9 *(Guaranteeing μ-monotonicity)* In linear polynomial interpretations $\mathcal{A}$, μ-monotonicity is guaranteed if, for all k-ary symbols $f \in \mathcal{F}$ and $i \in \mu(f)$, the ith coefficient f_i of the linear polynomial $f^{\mathcal{A}}(x_1, \ldots, x_k)$ satisfies $f_i \geq 1$. Similarly, dealing with matrix interpretations $\mathcal{A}$, μ-monotonicity is guaranteed if, for all k-ary symbols $f \in \mathcal{F}$ and $i \in \mu(f)$, the leftmost uppermost entry $(F_i)_{11}$ of the ith matrix F_i of the affine expression $f^{\mathcal{A}}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ satisfies $(F_i)_{11} \geq 1$.

6.2 Polynomial Bounds on $dc_{\mu}(n)$ from Matrix and Polynomial Interpretations

In the following result, I_d is the identity d-square matrix.

Theorem 2 *Let $\mathcal{R}$ be a TRS, $\mu \in M_{\mathcal{R}}$ and let $\mathcal{A}$ be a μ-monotone TMI with domain $\mathbb{N}^d$ for some $d \in \mathbb{N}_{>0}$ such that $\succ_{\mathcal{A}}$ is compatible with $\mathcal{R}$.*

1. *If for all $f \in \mathcal{F}$, $f^{\mathcal{A}}(x_1, \ldots, x_k) = I_d x_i$ for some $1 \leq i \leq k$, or $f^{\mathcal{A}}(x_1, \ldots, x_k) = \mathbf{f}$ for some vector $\mathbf{f}$, then $dc_{\mu}(n) \in O(1)$;*
2. *otherwise, $dc_{\mu}(n) \in O(n^d)$.*

Proof We rework the proof of [46, Theorem 6] to make explicit the points which are different for *CSR*. According to (15), for all μ-rewrite sequences $s \hookrightarrow_{\mathcal{R},\mu}^k t$ of length k we have $[\alpha]_{\mathcal{A}}(s) \succ^k [\alpha]_{\mathcal{A}}(t)$, where $\succ$ is the ordering on tuples in (13). By definition of $\succ$, this implies $([\alpha]_{\mathcal{A}}(s))_1 \geq_{\mathbb{N}} k + ([\alpha]_{\mathcal{A}}(t))_1 \geq_{\mathbb{N}} k$, i.e., for all valuations α, the first entry $([\alpha]_{\mathcal{A}}(s))_1$ of the vector interpreting s is an upper bound of the length k of the sequence. In particular, this holds for α_0 defined by $\alpha_0(x) = \mathbf{0}$ for all $x \in \mathcal{X}$. For all upper triangular matrices $M \in \mathbb{N}^{d \times d}$ with diagonal entries 0 or 1, and $p \in \mathbb{N}$, [46, Lemma 5] proves that the entries of M^p satisfy

$$(M^p)_{ij} \leq (j - i)!(ap)^{j-i}, \text{ where } a = \max\{M_{ij} \mid 1 \leq i, j \leq d\}.$$

This implies that, for all $1 \leq i, j \leq d$, $(M^p)_{ij} \in O(p^{d-1})$. As noticed in the proof of [46, Theorem 6], whenever terms s are interpreted by using $\mathcal{A}$, such products of (possibly different) matrices occur in $s^{\mathcal{A}}$ in number p at most the depth δ_s of s minus one, i.e., $p \leq \delta_s - 1$ (for constant or variable terms (of depth 1), no product is required). Of course, $\delta_s \leq |s|$; hence, $p \leq |s|$. Furthermore, the authors observe that the number m of products to be considered as additive components of $[\alpha]_{\mathcal{A}}(s)$ is exactly $|s|$. By considering $M^0 = I_d$ as a 'degenerate' product accompanying the vector representing the interpretation of a constant or the valuation of the variable, this is acceptable. Note also that $(I_d)_{ij} \in O(p^{d-1})$ if $p > 0$. Thus, an addition $\sum_{i=1}^{m} N_i \mathbf{a}_i$ of $m = |s|$ vectors $N_i \mathbf{a}_i$ obtained from products N_i of at most $p \leq |s|$ matrices each with entries $(N_i)_{qr}$ bounded by $O(|s|^{d-1})$ is necessary to obtain the interpretation of s in $\mathcal{A}$ with respect to a given valuation α. Since the addition of $m = |s|$ values bounded by $O(|s|^{d-1})$ is bounded by $O(|s|^d)$, we obtain $([\alpha]_{\mathcal{A}}(s))_1 \in O(|s|^d)$ and, since $([\alpha]_{\mathcal{A}}(s))_1$ is an upper bound on the length of any μ-derivation starting from s, $\mathsf{dc}_\mu(n) \in O(n^d)$, which proves item 2.

As for the particular case considered in item 1, where for all $f \in \mathcal{F}$, $f^{\mathcal{A}} = I_d x_i$ for some $1 \leq i \leq k$ or $f^{\mathcal{A}} = \mathbf{f}$, note that the interpretation $[\alpha_0]_{\mathcal{A}}(t)$ of a term t using such kind of interpretation $\mathcal{A}$ is of the form $[\alpha_0]_{\mathcal{A}}(t) \leq \mathbf{b}$ for some vector $\mathbf{b} = \max(\{ f^{\mathcal{A}}(\mathbf{0}, \ldots, \mathbf{0}) \mid f \in \mathcal{F}\} \cup \{\mathbf{0}\})$. Since $\mathbf{b}$ depends on $\mathcal{A}$ only, we have $\mathsf{dh}(s, \hookrightarrow) \leq \mathbf{b}_1$ for all terms s, and hence, $\mathsf{dc}_\mu(n) \in O(1)$.

Remark 10 *(More bounds based on matrix interpretations)* Important improvements on [46, Theorem 6] regarding the use of matrix interpretations to obtain complexity bounds on $\mathsf{dc}_{\mathcal{R}}(n)$ are discussed in [48]. We claim that (as Theorem 2(2) does for [46, Theorem 6]) they can be used as they are to obtain bounds on $\mathsf{dc}_\mu(n)$. The reason is that monotonicity plays no essential role in the proof of such results, which apply to matrix interpretations $\mathcal{A}$ generating an ordering $\succ_{\mathcal{A}}$ which is compatible with the rewrite relation. For $\rightarrow_{\mathcal{R}}$ monotonicity would be required on $\mathcal{A}$ (to guarantee termination, not compatibility!); for $\hookrightarrow_{\mathcal{R},\mu}$, μ-monotonicity would be required instead (to guarantee μ-termination now). Bounds obtained from the analysis of matrices in $\mathcal{A}$ would apply to $\mathsf{dc}_{\mathcal{R}}(n)$ or $\mathsf{dc}_\mu(n)$ without special changes.

For polynomial interpretations, we have the following result, where by a *hyperlinear* polynomial interpretation $\mathcal{A}$ we mean a strongly linear interpretation where for all $f \in \mathcal{F}$, $f^{\mathcal{A}} = x$ for some $x \in \mathcal{X}$ or $f^{\mathcal{A}} \in \mathbb{N}$.

Theorem 3 *Let $\mathcal{R}$ be a TRS and $\mu \in M_{\mathcal{R}}$. Let $\mathcal{A}$ be a polynomial interpretation and $\succ_{\mathcal{A}}$ be the corresponding μ-reduction ordering that is compatible with $\mathcal{R}$.*

1. *If $\mathcal{A}$ is hyperlinear, then $\mathsf{dc}_\mu(n) \in O(1)$.*
2. *If $\mathcal{A}$ is strongly linear, then $\mathsf{dc}_\mu(n) \in O(n)$.*
3. *If $\mathcal{A}$ is linear, then $\mathsf{dc}_\mu(n) \in 2^{O(n)}$.*
4. *Otherwise, $\mathsf{dc}_\mu(n) \in 2^{2^{O(n)}}$.*

Proof The first two items are particular cases of one-dimensional matrix interpretations. The proof of the last two cases is an immediate adaptations of the proofs of [33, Proposition 2.2.1] and [33, Proposition 2.1.2], respectively.

The bounds in Theorem 3 are *tight*, i.e., they cannot be improved.

Remark 11 Note that matrix interpretations $\mathcal{A}_M$ satisfying the condition in item 1 of Theorem 2 can always be treated as polynomial interpretations $\mathcal{A}_P$ as in item 1 of Theorem 3 by just letting $f^{\mathcal{A}_P}((\mathbf{x}_1)_1, \ldots, (\mathbf{x}_k)_1) = (f^{\mathcal{A}_M}(\mathbf{x}_1, \ldots, \mathbf{x}_k))_1$.

Example 14 Consider $\mathcal{R}$ and μ in Example 1. A proof of μ-termination for $\mathcal{R}$ is obtained with the following hyperlinear interpretation

$$\mathsf{a}^{\mathcal{A}} = 2 \qquad \mathsf{b}^{\mathcal{A}} = 0 \qquad \mathsf{f}^{\mathcal{A}}(x) = 3 \qquad \mathsf{g}^{\mathcal{A}}(x) = x \qquad \mathsf{h}^{\mathcal{A}}(x) = 1$$

By Theorem 3(1), $\mathsf{dc}_{\mu}(n) \in O(1)$.

Example 15 Consider $\mathcal{R}$ and μ in Example 2. A proof of μ-termination of $\mathcal{R}$ is obtained with the strongly linear interpretation $\mathcal{A}$ where $\mathsf{app}^{\mathcal{A}}(x, y) = x + 1$. By Theorem 3(2), $\mathsf{dc}_{\mu}(n) \in O(n)$.

Example 16 Consider $\mathcal{R}$ and μ in Example 3. The strongly linear interpretation

$$0^{\mathcal{A}} = 1 \qquad \mathsf{s}^{\mathcal{A}}(x) = x + 2 \qquad \mathsf{f}^{\mathcal{A}}(x, y) = x + 1$$

proves $\mathcal{R}$ μ-terminating. By Theorem 3(2), $\mathsf{dc}_{\mu}(n) \in O(n)$.

Example 17 Consider $\mathcal{R}$ and μ in Example 5. A proof of μ-termination of $\mathcal{R}$ is obtained with the linear interpretation

$$\neg^{\mathcal{A}}(x) = 2x + 2 \qquad x \wedge^{\mathcal{A}} y = 2y \qquad x \vee^{\mathcal{A}} y = x + 2y + 2 \qquad x \supset^{\mathcal{A}} y = x + y$$

By Theorem 3(3), $\mathsf{dc}_{\mu}(n) \in 2^{O(n)}$.

Remind that TRSs $\mathcal{R}$ in Examples 1 and 3 are left-linear and completely defined (see Example 10). Thus, they fulfill the conditions in Corollary 1. According to Remark 3, we notice the following.

Example 18 *(Running example I—bounds)* For $\mathcal{R}$ and μ in Example 1, $\mathsf{dc}_{\mu}(n) \in O(1)$ in Example 14 provides a suitable bound on the length of normalization sequences in $\mathcal{R}$. Note that $\mathsf{dc}_{\mathcal{R}}(n)$ is *not* defined as $\mathcal{R}$ is not terminating.

Example 19 *(Running example III—improved bounds)* For $\mathcal{R}$ and μ in Example 3, $\mathsf{dc}_{\mu}(n) \in O(n)$ in Example 16 improves on AProVE's bounds for rewriting and innermost rewriting (Examples 3 and 9).

7 Normalization via μ-Normalization

For left-linear TRSs $\mathcal{R}$ and $\mu \in CM_{\mathcal{R}}$, normalization of terms can be decomposed in the computation of a (possibly empty) initial subsequence of context-sensitive rewriting steps followed of a possibly empty sequence of additional reductions issued on maximal frozen subterms of μ-normal forms.

Theorem 4 [40, Theorem 6.8] *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_{\mathcal{R}}$. Let $s, t \in \mathcal{T}(\mathcal{F}, \mathcal{X})$ such that $s \rightarrow^! t$. There exists $u \in \mathcal{T}(\mathcal{F}, \mathcal{X})$ such that $s \hookrightarrow^!_{\mu} u \rightarrow^!_{\mu} t$ and $MRC^{\mu}(u) = MRC^{\mu}(t)$.*

This is the basis of the iterative *normalization via μ-normalization* ($N\mu N$) process norm_{μ} in [37, Sect. 4], further discussed in [40, Sect. 9.3], see also Fig. 1:

$$\mathsf{norm}_{\mu}(s) = C[\mathsf{norm}_{\mu}(s_1), \dots, \mathsf{norm}_{\mu}(s_n)] \tag{16}$$

where

$$u \text{ is a } \mu - \text{ normal form of } s \tag{17}$$

$$u = C[s_1, \dots, s_n] \text{ for } C[, \dots,] = MRC^{\mu}(u) \tag{18}$$

Remark 12 *(Deterministic/non-deterministic presentations of* norm_μ*)* In [37, Figure 1] $\mathsf{norm}_\mu(t)$ deterministically returns a (possibly empty) *set of terms* T, whereas the (simpler) presentation in [40, Sect. 9.3] *non-deterministically* returns one of the terms in T. For simplicity, the presentation of the above algorithm uses this style. In the sequel, we rather follow the set-theoretic presentation in the technical results.

In order to prepare the analysis of derivational complexity of $N\mu N$, we investigate the number of maximal frozen subterms $s_1, \ldots, s_n$ in (18) of the μ-normal forms u obtained in $N\mu N$'s μ-normalization steps (17). Note that $N\mu N$ will recursively continue on such terms in (16).

Definition 3 Let μ be a replacement map and t be a term. The number of *maximal frozen subterms* of t, which coincides with the number of holes of its maximal replacing context $MRC^\mu(t)$, is denoted as $\mathsf{nmf}_\mu(t) = |\mathcal{F}r^\mu(t)| = |minimal_{\leq}(\overline{\mathcal{P}os^\mu}(t))|$.

Example 20 For $\mathcal{R}$ and μ in Example 5, $\mathsf{nmf}_\mu(\ell_{(8)}) = |\mathcal{F}r^\mu(\ell_{(8)})| = |\{1\}| = 1$ and $\mathsf{nmf}_\mu(r_{(8)}) = |\mathcal{F}r^\mu(r_{(8)})| = |\{1.1, 2.1\}| = 2$.

Section 7.1 investigates bounds on $\mathsf{nmf}_\mu(t)$ for a given term t. Then, Sect. 7.2 explores how to approximate the number $\mathsf{nmf}_\mu(t)$ of maximal frozen subterms s' of the μ-normal forms t of a given term s.

7.1 The Number of Maximal Frozen Subterms of a Term

Definition 4 (*p-adic and p-active signature*) Let $\mathcal{F}$ be a signature and $\mu \in M_\mathcal{F}$. Let p be the least natural number such that $ar(f) \leq p$ for all $f \in \mathcal{F}$. Then $\mathcal{F}$ is called a *p-adic* signature. Let p be the least natural number such that $|\mu(f)| \leq p$ for all $f \in \mathcal{F}$. Then, $\mathcal{F}$ is called *p-active* (with respect to μ).

We often use monadic and dyadic instead of 1- and 2-adic. In the sequel, given a term t, $\mathcal{F}(t)$ is the set of symbols from $\mathcal{F}$ occurring in t. We have the following.

Proposition 5 *Let $\mathcal{F}$ be a signature, $\mu \in M_\mathcal{F}$, $t \in \mathcal{T}(\mathcal{F}, \mathcal{X})$, and $\mathcal{G} = \mathcal{F}(MRC^\mu(t))$.*

1. *If $\mathcal{G}$ is p-adic and 0-active, then $\mathsf{nmf}_\mu(t) = ar(root(t)) \leq p$.*
2. *If for all $f \in \mathcal{G}$, $\mu(f) = \mu_\top(f)$, then $\mathsf{nmf}_\mu(t) = 0$.*
3. *If $\mathcal{G}$ is p-adic for some $p > 1$, and 1-active, then $\mathsf{nmf}_\mu(t) \leq \frac{p-1}{p}|MRC^\mu(t)|$.*
4. *Otherwise, $\mathsf{nmf}_\mu(t) \leq |MRC^\mu(t)|$.*

Proof The first two cases and the last one are trivial. As for the third case, if $p > 1$ and $\mathcal{G}$ is 1-active, then since $|\mu(f)| \leq 1$ for all $f \in \mathcal{G}$, each level of the tree associated with t has (at most) a single active position and (at least) a single minimal frozen position. The structure of t allowing for more frozen positions is $t = f(x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_p)$ where f is a p-adic symbol with $\mu(f) = \{i\}$ for some $1 \leq i \leq p$, $x_1, \ldots, x_{i-1}, x_{i+1}, x_p$ are (non-necessarily distinct) variables, and u is either a variable or a term of the same structure. In this case, each level of the tree below the root adds (at most) $p - 1$ new minimal frozen positions to $\mathcal{F}r^\mu(t) = minimal(\overline{\mathcal{P}os^\mu}(t))$, and therefore, there are

$$\mathsf{nmf}_\mu(t) = (\delta_t - 1)(p - 1) \tag{19}$$

minimal frozen positions in t. We also have $|MRC^\mu(t)| = 1 + (p - 1) + |MRC^\mu(u)|$, i.e., $|MRC^\mu(t)| = (\delta_t - 1)p + 1$. Hence, $\delta_t - 1 = \frac{|MRC^\mu(t)|-1}{p}$ and using (19), we obtain $\mathsf{nmf}_\mu(t) = \frac{|MRC^\mu(t)|-1}{p}(p - 1) \leq \frac{p-1}{p}|MRC^\mu(t)|$.

The following fact is used below.

Proposition 6 *Let* $s, t \in \mathcal{T}(\mathcal{F}, \mathcal{X})$ *be such that* $s = \sigma(t)$ *for some substitution* σ. *Let* $\mu \in M_{\mathcal{R}}$. *Then,* $\mathsf{nmf}_{\mu}(s) = \mathsf{nmf}_{\mu}(t) + \sum_{x \in \mathcal{V}ar(t)} |\mathcal{P}os_x^{\mu}(t)| \mathsf{nmf}_{\mu}(\sigma(x))$.

Proof The set of positions in the frontier of s consists of the positions in the frontier of t together with those frozen positions introduced by instances $\sigma(x)$ of active positions of variables in t.

7.2 The Number of Maximal Frozen Subterms of μ-Normal Forms of Terms

In $N\mu N$, intermediate μ-normal forms t of terms s are obtained, and then, $N\mu N$ resumes on maximal frozen subterms of t. In this section, we investigate approximations to the maximum value of $\mathsf{nmf}_{\mu}(t)$ whenever t is a μ-normal form of s.

Definition 5 Let $\mathcal{R}$ be a finitely branching TRS and $\mu \in M_{\mathcal{R}}$ be such that $\mathcal{R}$ is μ-terminating. For all terms s, we let $\mathsf{NMF}_{\mu}^{\hookrightarrow !}(s) = \max_{s \hookrightarrow_{\mu}^! t} \mathsf{nmf}_{\mu}(t)$.

In order to use Proposition 5 to obtain bounds on $\mathsf{NMF}_{\mu}^{\hookrightarrow !}(s)$, we need to consider the function symbols occurring in the maximal replacing contexts of μ-normal forms of terms. For this purpose, we introduce the following notation:

$$\mathcal{F}_{\mu}^! = \{\mathcal{F}(MRC^{\mu}(t)) \mid (\exists s \in \mathcal{T}(\mathcal{F})) \ s \hookrightarrow_{\mu}^! t\}$$

We can approximate $\mathcal{F}_{\mu}^!$ as $\mathcal{F}$; however, *completely defined symbols* can be discarded as they cannot occur in the maximal replacing context of (ground) normal forms. This motivates the restriction to ground terms s in the definition of $\mathcal{F}_{\mu}^!$, thus guaranteeing groundness of t. As discussed in Sect. 4.1, if $\mathcal{T}(\mathcal{F}) = \varnothing$, we just need to add a fresh constant symbol $\bullet$ to $\mathcal{F}$ and deal with $\mathcal{R}^{\bullet}$ instead of $\mathcal{R}$.

Example 21 *[Running example V]* Consider the following (non-terminating) TRS $\mathcal{R}$

$$\mathsf{a} \rightarrow \mathsf{f}(\mathsf{c}(\mathsf{a}), \mathsf{b}) \tag{20}$$
$$\mathsf{f}(\mathsf{b}, y) \rightarrow y \tag{21}$$
$$\mathsf{f}(\mathsf{c}(x), y) \rightarrow y \tag{22}$$

together with $\mu = \mu_{\mathcal{R}}^{can}$, i.e., $\mu(\mathsf{f}) = \{1\}$ and $\mu(\mathsf{c}) = \varnothing$. Since a and f are completely defined (use induction on the structure of ground normal forms) $\mathcal{F}_{\mu}^! = \mathcal{C} = \{\mathsf{b}, \mathsf{c}\}$.

The following result is immediate from Proposition 5.

Corollary 2 *Let* $\mathcal{R} = (\mathcal{F}, R)$ *be a TRS,* $\mu \in M_{\mathcal{R}}$ *be such that* $\mathcal{F}_{\mu}^!$ *is p-adic, and s be a ground term. Then,*

1. *If* $\mathcal{F}_{\mu}^!$ *is 0-active, then* $\mathsf{NMF}_{\mu}^{\hookrightarrow !}(s) \leq p$.
2. *If for all* $f \in \mathcal{F}_{\mu}^!$, $\mu(f) = \mu_{\top}(f)$, *then* $\mathsf{NMF}_{\mu}^{\hookrightarrow !}(s) = 0$.

Proof Since s is ground, the μ-normalization of s leads to a term t in $\mathcal{T}(\mathcal{F}_{\mu}^!)$. Now we apply the first two items of Proposition 5 with $\mathcal{G} = \mathcal{F}_{\mu}^!$.

Example 22 For $\mathcal{R}$ and μ in Example 21, $\mathcal{F}_\mu^! = \{\mathsf{b}, \mathsf{c}\}$. Since $\mathcal{F}_\mu^!$ is monadic and 0-active, by Corollary 2(1), for all ground terms s, $\mathsf{NMF}_\mu^{\hookrightarrow^!}(s) \leq 1$.

Note that, if the size of s *decreases* when computing a μ-normal form t of s, then $\mathsf{NMF}_\mu^{\hookrightarrow^!}(s)$ is clearly bounded by $|s|$. In the following, we investigate this issue.

7.2.1 Size-Decreasing Relations

In this section, we investigate conditions guaranteeing that some relations (in particular the μ-normalization relation $\hookrightarrow_\mu^!$) are size-decreasing in the following sense:

Definition 6 (Size-decreasing relation) Consider a relation R on terms. We say that R is *size-decreasing* if for all terms $s, t \in \mathcal{T}(\mathcal{F}, \mathcal{X})$, s R t implies $|s| \geq |t|$.

Clearly, size-decreasingness of R implies size-decreasingness of R^+, R^*, $\mathsf{R}^!$ and also size-decreasingness of any subrelation $\mathsf{R}' \subseteq \mathsf{R}$. The following result is immediate from the notion of size-decreasingness and Proposition 5.

Proposition 7 *Let $\mathcal{R} = (\mathcal{F}, R)$ and $\mu \in M_\mathcal{R}$ be such that $\mathcal{F}$ is p-adic and $\hookrightarrow_\mu^!$ is size-decreasing. Let s be a ground term. If $p > 1$ and $\mathcal{F}_\mu^!$ is 1-active, then (i) $\mathsf{NMF}_\mu^{\hookrightarrow^!}(s) \leq \frac{p-1}{p}|s|$. Otherwise, (ii) $\mathsf{NMF}_\mu^{\hookrightarrow^!}(s) \leq |s|$.*

Proof By size-decreasingness of $\hookrightarrow_\mu^!$, $|s| \geq |t|$ for all μ-normal forms t of s. Since s is ground, t is also ground and $t \in \mathcal{T}(\mathcal{F}_\mu^!)$. By Proposition 5, with $\mathcal{G} = \mathcal{F}_\mu^!$, we have, for item (i),

$$\mathsf{NMF}_\mu^{\hookrightarrow^!}(s) = \max_{s \hookrightarrow^! t} \mathsf{nmf}_\mu(t) \leq \max_{s \hookrightarrow^! t} \frac{p-1}{p}|MRC^\mu(t)| = \frac{p-1}{p} \max_{s \hookrightarrow^! t} |MRC^\mu(t)|$$
$$\leq \frac{p-1}{p} \max_{s \hookrightarrow^! t} |t| \leq \frac{p-1}{p} \max_{s \hookrightarrow^! t} |s|$$

Item (ii) follows due to the fact that $|s| \geq |t|$ (size-decreasingness of $\hookrightarrow_\mu^!$). $\quad\blacksquare$

In the following, we investigate criteria to guaranteee size-decreasingness of $\hookrightarrow^!$.

7.2.2 Size-Decreasing TRSs

Definition 7 (Size-decreasing rules and TRSs) A rule $\ell \to r$ satisfying $|\ell| \geq |r|$ is called *size-decreasing*. A TRS $\mathcal{R}$ is size-decreasing if all rules are size-decreasing.

Example 23 The TRS $\mathcal{R}$ in Example 2 is size-decreasing. The TRSs $\mathcal{R}$ in Examples 1, 3, 5 and 21 are *not* size-decreasing.

The following observation is used below: given a term t and a substitution σ, $|\sigma(t)| = |t| + \sum_{x \in \mathcal{V}ar(t)} |\mathcal{P}os_x(t)|(|\sigma(x)| - 1)$.

Proposition 8 *Let s and t be terms and $\ell \to r$ be a* non-duplicating *rule. If $s \to_{\ell \to r} t$, then $|s| - |t| \geq |\ell| - |r|$.*

Proof Since $s \rightarrow_{\ell \rightarrow r} t$, we have $s|_p = \sigma(\ell)$ for some $p \in \mathcal{P}os(s)$ and substitution σ and $t = s[\sigma(r)]_p$. Therefore, since $|t| = |s| - |\sigma(\ell)| + |\sigma(r)|$, we have

$$
\begin{aligned}
|s| - |t| &= |\sigma(\ell)| - |\sigma(r)| \\
&= |\ell| + \sum_{x \in \mathcal{V}ar(\ell)} |\mathcal{P}os_x(\ell)|(|\sigma(x)| - 1) - |r| - \sum_{x \in \mathcal{V}ar(r)} |\mathcal{P}os_x(r)|(|\sigma(x)| - 1) \\
&= |\ell| - |r| + \sum_{x \in \mathcal{V}ar(\ell)} (|\mathcal{P}os_x(\ell)| - |\mathcal{P}os_x(r)|)(|\sigma(x)| - 1)
\end{aligned}
$$

Since $|\mathcal{P}os_x(\ell)| \geq |\mathcal{P}os_x(r)|$ and $|\sigma(x)| - 1 \geq 0$, we have $|s| - |t| \geq |\ell| - |r|$.

Proposition 9 *Let $\mathcal{R}$ be a size-decreasing TRS. Then, $\rightarrow^*$ is size-decreasing iff $\mathcal{R}$ is non-duplicating.*

Proof For the *if part*, we proceed by induction on the length of the sequence $s \rightarrow^* t$, to prove $|s| \geq |t|$. The base case is immediate. Let $s \rightarrow s' \rightarrow^* t$ with $s|_p = \sigma(\ell)$ and $s' = s[\sigma(r)]_p$. By Proposition 8, $|\sigma(\ell)| \geq |\sigma(r)|$; hence, $|s| \geq |s'|$. By the induction hypothesis, $|s| \geq |s'| \geq |t|$, i.e., $\rightarrow^*$ is size-decreasing. Now, for the *only if* part, we proceed by contradiction. Assume that both $\mathcal{R}$ and $\rightarrow^*$ are size-decreasing but $\mathcal{R}$ is duplicating. Then there is a rule $\ell \rightarrow r$ with $m = |\ell| - |r| \geq 0$ (by size-decreasingness of $\mathcal{R}$) and $|\mathcal{P}os_x(\ell)| < |\mathcal{P}os_x(r)|$ for some variable x. Let $n = |\mathcal{P}os_x(r)| - |\mathcal{P}os_x(\ell)| > 0$. Let t be a term of size $|t| = k \geq m + 2$ and σ be such that $\sigma(x) = t$ and $\sigma(y) = y$ for all $y \neq x$. Let $s = \sigma(\ell)$ and $s' = \sigma(r)$. Note that $s \rightarrow s'$. Now, note that $|s| = |\ell| + |\mathcal{P}os_x(\ell)|(k-1)$ and $|s'| = |r| + |\mathcal{P}os_x(r)|(k-1)$. Hence, $|s| - |s'| = |\ell| - |r| + (|\mathcal{P}os_x(\ell)| - |\mathcal{P}os_x(r)|)(k-1) = m - n(k-1) = m - n(m+1) < 0$ (because $n > 0$), i.e., $\rightarrow^*$ is not size-decreasing, a contradiction.

Example 24 The TRS $\mathcal{R}$ in Example 2 is size-decreasing and non-duplicating. By Proposition 9, $\rightarrow^*$ (and hence $\hookrightarrow_\mu^!$, for μ in the example) is size-decreasing.

Remark 13 *(Limitation)* Left-linear and non-duplicating TRSs are *linear*. Since left-linearity is necessary for $N\mu N$ to be sound, this leads to a strong restriction on the TRSs amenable for a complexity treatment that assumes non-duplicatingness.

7.2.3 Weakly Size-Decreasing TRSs

The following definition slightly relaxes size-decreasingness.

Definition 8 (Weak size-decreasingness) A TRS $\mathcal{R}$ is *weakly size-decreasing* if for all rules $\ell \rightarrow r \in \mathcal{R}$ with $|\ell| < |r|$, there are rules $\ell_1 \rightarrow r_1, \ldots, \ell_n \rightarrow r_n \in \mathcal{R}$ and substitutions σ_i such that (i) $r = \sigma_1(\ell_1)$, (ii) for all $1 \leq i < n$, $r_i = \sigma_{i+1}(\ell_{i+1})$ and (iii) $|r| - |\ell| \leq \sum_{i=1}^n |\ell_i| - |r_i|$.

Example 25 The TRS $\mathcal{R}$ in Example 21 is weakly size-decreasing: both (21) and (22) are size-decreasing, and (i) the right-hand side $r_{(20)}$ of rule (20) matches $\ell_{(22)}$ and (ii) $|r_{(20)}| - |\ell_{(20)}| = 3 \leq 3 = |\ell_{(22)}| - |r_{(22)}|$, as required in Definition 8.

For size-decreasing and non-duplicating TRSs, $\rightarrow^*$ is size-decreasing (Proposition 9), and hence, $\hookrightarrow^*$ and $\hookrightarrow^!$ also are. However, in general size-decreasingness of $\hookrightarrow^!$ does *not* imply that of $\hookrightarrow^*$. For weak size-decreasingness, we have the following.

Proposition 10 *Let $\mathcal{R}$ be a weakly size-decreasing and non-duplicating TRS and $\mu \in M_{\mathcal{R}}$ be such that $\mathcal{R}$ is μ-confluent and μ-terminating. Then, $\hookrightarrow^!$ is size-decreasing.*

Proof Since $\mathcal{R}$ is μ-terminating, there is a μ-reduction ordering $\succ$ on terms such that whenever $s \hookrightarrow^+ t$, we have $s \succ t$. We proceed by induction on s using $\succ$. As for the base case, we consider terms s which are μ-normal forms. Thus, $s = t$. For the induction step, if s is μ-reducible, assume $s \hookrightarrow_{\ell \to r} s' \hookrightarrow^! t$, hence $s \succ s'$ and either $s' \succ t$ or $s' = t$. By Proposition 8, $|s| - |s'| \geq |\ell| - |r|$ and by the induction hypothesis we have $|s'| \geq |t|$. We consider two cases. (A) If $\ell \to r$ is size-decreasing, then $|\ell| \geq |r|$ and $|s| \geq |s'| \geq |t|$. (B) If $\ell \to r$ is not size-decreasing, then there are rules $\ell_1 \to r_1, \ldots, \ell_n \to r_n \in \mathcal{R}$ for some $n > 0$ and substitutions σ_i such that $r = \sigma'(\ell_1)$, for all $1 \leq i < n$, $r_i = \sigma_{i+1}(\ell_{i+1})$, and $|r| - |\ell| \leq \sum_{i=1}^n |\ell_i| - |r_i|$. Note that $\sigma(r) = \sigma(\sigma_1(\ell_1))$, and therefore,

$$s'|_p = \sigma(r) = \sigma(\sigma_1(\ell_1)) \overset{\Lambda}{\hookrightarrow} \sigma(\sigma_1(r_1)) = \sigma(\sigma_1(\sigma_2(\ell_2))) \hookrightarrow^* \sigma(\sigma_1(\cdots(\sigma_n(\ell_n))\cdots))$$
$$\overset{\Lambda}{\hookrightarrow} \sigma(\sigma_1(\cdots(\sigma_n(r_n))\cdots))$$

Let $s'' = s'[\sigma(\sigma_1(\cdots(\sigma_n(r_n))\cdots))]_p$. Note that $s \hookrightarrow s' \hookrightarrow^+ s''$. By Proposition 8, for all $1 \leq i < n$, $|\sigma(\sigma_1(\cdots\sigma_i(\ell_i)\cdots))| - |\sigma(\sigma_1(\cdots\sigma_i(\sigma_{i+1}(\ell_{i+1}))\cdots))| \geq |\ell_i| - |r_i|$ and also $|\sigma(\sigma_1(\cdots\sigma_n(\ell_n)\cdots))| - |\sigma(\sigma_1(\cdots\sigma_n(r_n)\cdots))| \geq |\ell_n| - |r_n|$. Since

$$|s'| - |s''| = \sum_{i=1}^{n-1} |\sigma(\sigma_1(\cdots\sigma_i(\ell_i)\cdots))| - |\sigma(\sigma_1(\cdots\sigma_i(\sigma_{i+1}(\ell_{i+1}))\cdots))|$$

we obtain $|s'| - |s''| \geq \sum_{i=1}^{n-1} |\ell_i| - |r_i|$. Hence,

$$|s'| - |s| \leq |r| - |\ell| \leq \sum_{i=1}^{n-1} |\ell_i| - |r_i| \leq |s'| - |s''|$$

and therefore $|s| \geq |s''|$. By μ-confluence, $s'' \hookrightarrow^! t$. If $s'' = t$, then $|s''| \geq |t|$; hence, $|s| \geq |t|$. Otherwise, $s \succ s'' \succ t$; by the induction hypothesis $|s| \geq |t|$. $\square$

Example 26 Consider $\mathcal{R}$ and μ in Example 21. The following strongly linear polynomial interpretation

$$\mathsf{a}^{\mathcal{A}} = 2 \quad \mathsf{b}^{\mathcal{A}} = 0 \quad \mathsf{c}^{\mathcal{A}}(x) = 0 \quad \mathsf{f}^{\mathcal{A}}(x, y) = x + y + 1$$

proves $\mathcal{R}$ μ-terminating. Also, $\mathcal{R}$ is μ-confluent as there is no critical pair and no active variable in the left-hand side of a rule becomes frozen anywhere else in the rule (LHRV property, see [40, Sect. 8]). By Proposition 10, $\hookrightarrow^!$ is size-decreasing.

7.2.4 Beyond Size-Decreasingness

Definition 9 Let $\mathcal{R} = (\mathcal{F}, R)$ be a TRS and $\mu \in M_{\mathcal{R}}$. Given a rule $\ell \to r \in \mathcal{R}$, we let $F(\ell \to r) = \mathsf{nmf}_\mu(r) - \mathsf{nmf}_\mu(\ell)$. Then, $F(\mathcal{R}) = \max\{F(\ell \to r) \mid \ell \to r \in \mathcal{R}\}$. When no confusion arises, we drop $\mathcal{R}$ and rather use F instead of $F(\mathcal{R})$.

Intuitively, $F(\mathcal{R})$ shows how many maximal *frozen* subterms can be introduced by the application of a rewrite step with $\mathcal{R}$. In particular, if $F(\mathcal{R}) < 0$ we can think of rules as 'destroying' frozen positions.

Example 27 For $\mathcal{R}$ and μ in Example 5,

$$\mathcal{F}r^\mu(\ell_{(8)}) = \{1\} \qquad \mathcal{F}r^\mu(r) = \{1.1, 2.1\} \qquad \text{hence} \quad F((8)) = 2 - 1 = 1$$
$$\mathcal{F}r^\mu(\ell_{(9)}) = \varnothing \qquad \mathcal{F}r^\mu(r_{(9)}) = \varnothing \qquad \text{hence} \quad F((9)) = 0 - 0 = 0$$
$$\mathcal{F}r^\mu(\ell_{(10)}) = \varnothing \qquad \mathcal{F}r^\mu(r_{(10)}) = \{2.1\} \qquad \text{hence} \quad F((10)) = 1 - 0 = 1$$

Therefore, $F(\mathcal{R}) = 1$. For $\mathcal{R}$ and μ in Example 2, $F(\mathcal{R}) = -1$.

We have the following:

Proposition 11 *Let $\mathcal{R}$ be a TRS and $\mu \in M_\mathcal{R}$. For all $\ell \to r \in \mathcal{R}$, $\mathsf{nmf}_\mu(r) \leq F(\mathcal{R}) + \mathsf{nmf}_\mu(\ell)$.*

Proof By Definition 9, for all $\ell \to r \in \mathcal{R}$, we have $F(\mathcal{R}) \geq F(\ell \to r)$ and $F(\ell \to r) + \mathsf{nmf}_\mu(\ell) = \mathsf{nmf}_\mu(r)$. Hence, $\mathsf{nmf}_\mu(r) \leq F(\mathcal{R}) + \mathsf{nmf}_\mu(\ell)$.

Definition 10 [25, Definition 1] Let $\mathcal{R}$ be a TRS and $\mu \in M_\mathcal{R}$. A rule $\ell \to r$ is *non-μ-duplicating* iff for all $x \in Var(\ell)$, $|\mathcal{P}os_x^\mu(\ell)| \geq |\mathcal{P}os_x^\mu(r)|$. A TRS is called *non-$\mu$-duplicating* if all rules are; otherwise, it is called *μ-duplicating*.

Example 28 The TRS $\mathcal{R}$ with μ in Example 5 is non-μ-duplicating:

$$\mathcal{P}os_x^\mu(\ell_{(8)}) = \varnothing \quad \text{and} \quad \mathcal{P}os_x^\mu(r_{(8)}) = \varnothing$$
$$\mathcal{P}os_y^\mu(\ell_{(8)}) = \{2.1\} \quad \text{and} \quad \mathcal{P}os_y^\mu(r_{(8)}) = \{1.2\}$$
$$\mathcal{P}os_z^\mu(\ell_{(8)}) = \{2.2\} \quad \text{and} \quad \mathcal{P}os_z^\mu(r_{(8)}) = \{2.2\}$$

$$\mathcal{P}os_x^\mu(\ell_{(9)}) = \{1.1\} \quad \text{and} \quad \mathcal{P}os_x^\mu(r_{(9)}) = \{1\}$$
$$\mathcal{P}os_y^\mu(\ell_{(9)}) = \{2\} \quad \text{and} \quad \mathcal{P}os_y^\mu(r_{(9)}) = \{2\}$$

$$\mathcal{P}os_x^\mu(\ell_{(10)}) = \{1.1\} \quad \text{and} \quad \mathcal{P}os_x^\mu(r_{(10)}) = \{2.2\}$$
$$\mathcal{P}os_y^\mu(\ell_{(10)}) = \{2.1\} \quad \text{and} \quad \mathcal{P}os_y^\mu(r_{(10)}) = \{1\}$$

The following results put bounds on $\mathsf{nmf}_\mu(t)$ whenever t is obtained by μ-rewriting a term s. Size-decreasingness is not required, but bounds now depend on (bounds on) the derivational height of *CSR*.

Proposition 12 *Let $\mathcal{R}$ be a TRS and $\mu \in M_\mathcal{R}$ be such that $\mathcal{R}$ is non-μ-duplicating. Let $F = F(\mathcal{R})$ and s, t be terms such that $s \hookrightarrow^* t$. Then,*

1. *If $F > 0$, then $\mathsf{nmf}_\mu(t) \leq \mathsf{nmf}_\mu(s) + F\,\mathsf{dh}_\mu(s)$ and $\mathsf{NMF}_\mu^{\hookrightarrow!}(s) \leq \mathsf{nmf}_\mu(s) + F\,\mathsf{dh}_\mu(s)$.*
2. *If $F \leq 0$, then $\mathsf{nmf}_\mu(t) \leq \mathsf{nmf}_\mu(s)$ and $\mathsf{NMF}_\mu^{\hookrightarrow!}(s) \leq \mathsf{nmf}_\mu(s)$.*
3. *If $F < 0$, then $\mathsf{dh}_\mu(s) \leq \frac{\mathsf{nmf}_\mu(s)}{|F|}$.*

Proof We prove, by induction on the length n of $s \hookrightarrow^* t$, that $\mathsf{nmf}_\mu(t) \leq \mathsf{nmf}_\mu(s) + Fn$. Note that $n \leq \mathsf{dh}_\mu(s)$. If $n = 0$, then $s = t$ and it is immediate. If $n > 0$, let $s \hookrightarrow s' \hookrightarrow^* t$, i.e., there are $\ell \to r \in \mathcal{R}$, $p \in \mathcal{P}os^\mu(s)$ and a substitution σ such that $s|_p = \sigma(\ell)$ and $s' = s[\sigma(r)]_p$. By Proposition 6,

$$\mathsf{nmf}_\mu(s|_p) = \mathsf{nmf}_\mu(\ell) + \sum_{x \in Var(\ell)} |\mathcal{P}os_x^\mu(\ell)|\mathsf{nmf}_\mu(\sigma(x))$$
$$\mathsf{nmf}_\mu(s'|_p) = \mathsf{nmf}_\mu(r) + \sum_{x \in Var(r)} |\mathcal{P}os_x^\mu(r)|\mathsf{nmf}_\mu(\sigma(x)).$$

Therefore,

$$\mathsf{nmf}_\mu(s) = \mathsf{nmf}_\mu(s[\,]_p) + \mathsf{nmf}_\mu(s|_p)$$
$$= \mathsf{nmf}_\mu(s[\,]_p) + \mathsf{nmf}_\mu(\ell) + \sum_{x \in Var(\ell)} |\mathcal{P}os_x^\mu(\ell)|\mathsf{nmf}_\mu(\sigma(x))$$
$$\mathsf{nmf}_\mu(s') = (s'[\,]_p) + \mathsf{nmf}_\mu(s'|_p) = \mathsf{nmf}_\mu(s[\,]_p) + \mathsf{nmf}_\mu(s'|_p)$$
$$= \mathsf{nmf}_\mu(s[\,]_p) + \mathsf{nmf}_\mu(r) + \sum_{x \in Var(r)} |\mathcal{P}os_x^\mu(r)|\mathsf{nmf}_\mu(\sigma(x))$$

Hence,

$$\mathsf{nmf}_\mu(s) - \mathsf{nmf}_\mu(s') = \mathsf{nmf}_\mu(s|_p) - \mathsf{nmf}_\mu(s'|_p)$$
$$= \mathsf{nmf}_\mu(\ell) - \mathsf{nmf}_\mu(r) + \sum_{x \in Var(\ell)} (|\mathcal{P}os_x^\mu(\ell)| - |\mathcal{P}os_x^\mu(r)|)\mathsf{nmf}_\mu(\sigma(x)).$$

Since for all $\ell \to r \in \mathcal{R}$ and $x \in \mathcal{V}ar(\ell)$, $|\mathcal{P}os_x^\mu(\ell)| \geq |\mathcal{P}os_x^\mu(r)|$ holds, we have $\mathsf{nmf}_\mu(s) - \mathsf{nmf}_\mu(s') \geq \mathsf{nmf}_\mu(\ell) - \mathsf{nmf}_\mu(r)$. By Proposition 11, $\mathsf{nmf}_\mu(\ell) - \mathsf{nmf}_\mu(r) \geq -F$. By the induction hypothesis, $\mathsf{nmf}_\mu(t) \leq \mathsf{nmf}_\mu(s') + F(n-1)$. Thus,

$$\mathsf{nmf}_\mu(t) \leq \mathsf{nmf}_\mu(s') + Fn - F \leq \mathsf{nmf}_\mu(s') + Fn + \mathsf{nmf}_\mu(\ell) - \mathsf{nmf}_\mu(r) \leq \mathsf{nmf}_\mu(s) + Fn.$$

Now, since $n \leq \mathsf{dh}_\mu(s)$, we have (1) if $F \geq 0$, then $\mathsf{nmf}_\mu(t) \leq \mathsf{nmf}_\mu(s) + F\mathsf{dh}_\mu(s)$ and $\mathsf{NMF}_\mu^{\hookrightarrow!}(t) \leq \mathsf{nmf}_\mu(s) + F\mathsf{dh}_\mu(s)$, as desired. In particular, (2) if $F \leq 0$, then $\mathsf{nmf}_\mu(t) \leq \mathsf{nmf}_\mu(s)$ and $\mathsf{NMF}_\mu^{\hookrightarrow!}(s) \leq \mathsf{nmf}_\mu(s)$. Finally, (3) if $F < 0$, since $\mathsf{nmf}_\mu(t) \geq 0$, it must be $n|F| \leq \mathsf{nmf}_\mu(s)$ for all n representing the length of a μ-rewrite sequence starting from s. Thus, $\mathsf{dh}_\mu(s) \leq \frac{\mathsf{nmf}_\mu(s)}{|F|}$.

In the remainder of the paper, we investigate computational complexity of $N\mu N$:

1. Section 8 recalls and further develops a generalization of *CSR* called *layered CSR* (*LCSR* [38]) and provides a characterization of $N\mu N$ using *LCSR*.
2. Section 9 defines computational complexity of $N\mu N$ as the derivational complexity of *LCSR* and investigates derivational height of *LCSR*.
3. Section 10 connects derivational complexity of *LCSR* and derivational complexity of *CSR*.
4. Finally, Sect. 11 shows how to obtain bounds on derivational complexity of *LCSR* from bounds on $\mathsf{dc}_\mu(n)$.

8 Layered Context-Sensitive Rewriting

In [38], layered normalization of terms was formalized as a reduction relation called *layered context-sensitive rewriting*.

Definition 11 (Layered Context-Sensitive Rewriting)[38, Definition 1] Let $\mathcal{R}$ be a TRS, $\mu \in M_\mathcal{R}$ and $s, t \in \mathcal{T}(\mathcal{F}, \mathcal{X})$. We write $s \overset{p}{\hookrightarrow}_\mu t$ (or just $s \hookrightarrow_\mu t$, or $s \hookrightarrow t$) if either

1. $s \overset{p}{\hookrightarrow}_\mu t$, or
2. s is a μ-normal form, $s = C[s_1, \ldots, s_i, \ldots, s_n]$, where $C[, \ldots,] = MRC^\mu(t)$, $t = C[s_1, \ldots, t_i, \ldots, s_n]$, $s_i = s|_q$ for some $i \in \{1, \ldots, n\}$ and $q \in \mathcal{P}os(s)$, $p = q.p'$ for some position p', and $s_i \overset{p'}{\hookrightarrow}_\mu t_i$.

Example 29 Consider $\mathcal{R}$, μ, and the μ-normal form u in Example 2. Note that $MRC^\mu(u) = \mathsf{app}(x_1, \square)$. The normalization sequence with $\hookrightarrow$ finishes (5):

$$u = \mathsf{app}(x_1, \underline{\mathsf{app}(\mathsf{app}(x_2, x_3), x_4)}) \overset{2}{\hookrightarrow}_{\mathcal{R},\mu} \mathsf{app}(x_1, \mathsf{app}(x_2, \mathsf{app}(x_3, x_4)))$$

In contrast to *CSR*, where $\mathsf{NF}_\mathcal{R} \subseteq \mathsf{NF}_\mathcal{R}^\mu$ but, in general, $\mathsf{NF}_\mathcal{R}^\mu \not\subseteq \mathsf{NF}_\mathcal{R}$ (see Example 2), $\hookrightarrow_{\mathcal{R},\mu}$-normal forms are normal forms.

Proposition 13 *Let $\mathcal{R}$ be a TRS and $\mu \in M_\mathcal{R}$. Then, $\mathsf{NF}_{\hookrightarrow_{\mathcal{R},\mu}} = \mathsf{NF}_\mathcal{R}$.*

Proof Since normal forms are irreducible, $\mathsf{NF}_\mathcal{R} \subseteq \mathsf{NF}_{\hookrightarrow_{\mathcal{R},\mu}}$ trivially holds. Now let $t \in \mathsf{NF}_{\hookrightarrow_{\mathcal{R},\mu}}$. By definition of *LCSR*, $t \in \mathsf{NF}_\mathcal{R}^\mu$ and if $t = C[t_1, \ldots, t_n]$ for $C[, \ldots,] = MRC^\mu(t)$, then $t_1, \ldots, t_n \in \mathsf{NF}_{\hookrightarrow_{\mathcal{R},\mu}}$. We prove by structural induction that $t \in \mathsf{NF}_\mathcal{R}$. If t is a constant

or a variable, it is obvious. Otherwise, since $t_1, \ldots, t_n \in \mathsf{NF}_{\hookrightarrow_{\mathcal{R},\mu}}$ and $C[\ldots,] \neq \square$, by the induction hypothesis $t_1, \ldots, t_n \in \mathsf{NF}_{\mathcal{R}}$. Thus, if t is not a normal form, there must be a position $p \in \mathcal{P}os_{\mathcal{F}}(C[\ldots,])$, i.e., $p \in \mathcal{P}os^{\mu}(t)$ such that $t|_p$ is a redex. Hence, $t \notin \mathsf{NF}_{\mathcal{R}}^{\mu}$, a contradiction.

For left-linear TRSs $\mathcal{R}$ and canonical replacement maps $\mu \in CM_{\mathcal{R}}$, layered *CSR* sequences can be decomposed into an initial sequence of μ-rewriting steps which eventually obtains a μ-normal form. After that, the terms in the sequence remain μ-normal forms.

Proposition 14 *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_{\mathcal{R}}$. Consider a (finite or infinite) $\hookrightarrow_{\mu}$-sequence $t_1 \hookrightarrow_{\mu} t_2 \hookrightarrow_{\mu} \cdots$. If t_i is a μ-normal form for some $i \geq 1$, then there is $m \geq 1$ such that for all $1 \leq j < m$, s_j is μ-reducible and for all $j \geq m$, t_j is a μ-normal form.*

Proof Let m be such that t_m is the first μ-normal form in the sequence (note that $m \leq i$). By definition of *LCSR*, for all j, $1 \leq j < m$, t_j is μ-reducible. And for all t_j with $j \geq m$, $t_m \rightarrow^* t_j$. By [40, Theorem 6.2], t_j is a μ-normal form.

Proposition 14 may fail to hold if left-linearity or $\mu \in CM_{\mathcal{R}}$ is not required.

Example 30 Consider the following TRSs [40, Example 5.5]:

$$\mathcal{R}: \qquad \mathsf{a} \rightarrow \mathsf{c} \qquad\qquad \mathcal{R}': \qquad \mathsf{a} \rightarrow \mathsf{c}$$
$$\mathsf{f(c)} \rightarrow \mathsf{b} \qquad\qquad\qquad \mathsf{g}(x,x) \rightarrow \mathsf{b}$$

Let $\mu(\mathsf{f}) = \varnothing$. Note that $\mu \notin CM_{\mathcal{R}}$. Although $\mathsf{f(a)}$ is a μ-normal form, we have $\mathsf{f(\underline{a})} \hookrightarrow_{\mu} \mathsf{f(c)}$ and $\mathsf{f(c)}$ is *not* a μ-normal form. Now, consider the non-left-linear TRS $\mathcal{R}'$ and $\mu'(\mathsf{g}) = \varnothing$. Note that $\mu' = \mu_{\mathcal{R}'}^{can}$. Although $\mathsf{g(c,a)}$ is a μ'-normal form, we have $\mathsf{g(c,\underline{a})} \hookrightarrow_{\mathcal{R}',\mu'} \mathsf{g(c,c)}$ and $\mathsf{g(c,c)}$ is not a μ'-normal form. Thus, in both cases, the statement claimed in Proposition 14 does not hold.

In the following auxiliary result, which we use later, $sprefix_t(p)$ is the strict prefix of position p in a term t, i.e., the (possibly empty) list of symbols traversed when going from the root of t to position p (excluding p) [40, Sect. 3]. The result establishes that, for left-linear TRSs and $\mu \in CM_{\mathcal{R}}$ such that $\mathcal{R}$ is μ-terminating, in infinite *LCSR* sequences there are strict prefixes of arbitrary length which remain unchanged after some point in the sequence.

Proposition 15 *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_{\mathcal{R}}$ be such that $\mathcal{R}$ is μ-terminating. For all $t \in \mathcal{T}(\mathcal{F}, \mathcal{X})$ starting an infinite $\hookrightarrow_{\mu}$-sequence $t = t_1 \hookrightarrow_{\mu} t_2 \hookrightarrow_{\mu} \cdots$ and for all $N \geq 0$, there are $i \geq 0$ and $p \in \mathcal{P}os(t_i)$ such that $|p| \geq N$ and for all u such that $t_i \rightarrow^* u$, $p \in \mathcal{P}os(u)$ and $sprefix_{t_i}(p) = sprefix_u(p)$.*

Proof By induction on N. If $N = 0$, it is trivial. If $N > 0$, then by Proposition 14, and since $\mathcal{R}$ is μ-terminating, we can assume that there is $m \geq 1$ such that $t_i \hookrightarrow t_{i+1}$ for all i, $1 \leq i < m$ and $t_m \in \mathsf{NF}_{\mathcal{R}}^{\mu}$. By [40, Theorems 6.2 & 6.3], if $C[] = MRC^{\mu}(t_m)$, then, for all $i \geq m$, $t_i = C[t_{i1}, \ldots, t_{ik}] \in \mathsf{NF}_{\mathcal{R}}^{\mu}$ for terms $t_{ij} \in \mathcal{T}(\mathcal{F}, \mathcal{X})$. Furthermore, no reduct of t_i (with this or other rewrite sequence) contains a μ-replacing redex. Note that $C[]$ is not empty. Therefore, for all $i \geq m$, $|t_i| \geq N_1$ where $N_1 = min(\{|p_j| \mid p_j \in \mathcal{P}os_{\square}(C)\}) > 0$ is the minimum depth of a hole symbol $\square$ in $C[]$. Let $t_m = C[t_{m1}, \ldots, t_{mk}]$. There is j, $1 \leq j \leq k$ such that t_{mj} starts an infinite sequence. Assume that $t_m|_{p'} = t_{mj}$ for some $p' \in \mathcal{P}os_{\square}(C)$. By the induction hypothesis, there is $n \geq m$ and $q \in \mathcal{P}os(t_n)$ such that $|q| \geq N - N_1$ and for all v such

that $t_n|_{p'} \to^* v$, $q \in \mathcal{P}os(v)$ and $sprefix_{t_n|_{p'}}(q) = sprefix_v(q)$. Now, we let $p = p'.q$. Note that, since $t_m \in \mathsf{NF}_{\mathcal{R}}^{\mu}$, by [40, Theorem 6.2], $t_n = C[t_{m1}, \ldots, t_{mj-1}, t_n|_{p'}, \ldots, t_{mk}] \in \mathsf{NF}_{\mathcal{R}}^{\mu}$ and for all u such that $t_n \to^* u$, we will have $p \in \mathcal{P}os(u)$ and $sprefix_{t_n}(p) = sprefix_u(p)$ as desired.

As a consequence of Proposition 15, terms in infinite $\hookrightarrow\!\!\!\twoheadrightarrow_\mu$-sequences steadily increase their depth (and size).

Corollary 3 *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_{\mathcal{R}}$ be such that $\mathcal{R}$ is μ-terminating. For all $t \in \mathcal{T}(\mathcal{F}, \mathcal{X})$ starting an infinite $\hookrightarrow\!\!\!\twoheadrightarrow_\mu$-sequence $t = t_1 \hookrightarrow\!\!\!\twoheadrightarrow_\mu t_2 \hookrightarrow\!\!\!\twoheadrightarrow_\mu \cdots$ and for all $M, N \geq 0$, there is $i \geq 0$ such that for all u with $t_i \to^* u$, $\delta_u \geq M$ and $|u| \geq N$.*

8.1 Normalization via μ-Normalization and Layered *CSR*

We use the following theorem in the proof of the main result in this section which characterizes $N\mu N$ as $\hookrightarrow\!\!\!\twoheadrightarrow_\mu$-normalization (Theorem 6).

Theorem 5 *[38, Theorem 5] Let $\mathcal{R} = (\mathcal{F}, R)$ be a left-linear TRS, $\mu \in CM_{\mathcal{R}}$, and $s, t \in \mathcal{T}(\mathcal{F}, \mathcal{X})$. Then, $s \to^! t$ if and only if $s \hookrightarrow\!\!\!\twoheadrightarrow_\mu^! t$.*

Theorem 5 does not hold for arbitrary reducts (i.e., we cannot change $\to^!$ by $\to^*$ and $\hookrightarrow\!\!\!\twoheadrightarrow^!$ by $\hookrightarrow\!\!\!\twoheadrightarrow^*$) [38, Example 4].

Theorem 6 *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_{\mathcal{R}}$. For all terms s, t,*

$$t \in norm_\mu(s) \Leftrightarrow s \hookrightarrow\!\!\!\twoheadrightarrow_\mu^! t.$$

Proof ($\Rightarrow$) If $t \in norm_\mu(s)$, then $s \to^! t$. By Theorem 5, $t \hookrightarrow\!\!\!\twoheadrightarrow_\mu^! u$. ($\Leftarrow$) If $s \hookrightarrow\!\!\!\twoheadrightarrow_\mu^! t$, then, since t is a normal form (hence a μ-normal form), by Proposition 14, there is $m > 0$ such that $s = s_1 \hookrightarrow_\mu s_2 \hookrightarrow_\mu \cdots \hookrightarrow s_m \hookrightarrow\!\!\!\twoheadrightarrow_\mu^! t$ with s_m a normal form and s_i μ-reducible for all $1 \leq i < m$. We proceed by induction on the lexicographic combination of $>_{\mathbb{N}}$ and $\rhd$ on pairs of positive natural numbers and terms. For the base case, if $m = 1$, then s is a μ-normal form which is either a constant or a variable. By [40, Theorem 5.8], $s \hookrightarrow_\mu^! t$ and $\overline{\mathcal{P}os^\mu}(t) = \varnothing$ (hence $MRC^\mu(t) = t$). Therefore, according to (17) and (18), $norm_\mu(s) = t$. For the induction case, assume that $m > 1$. We consider two cases:

1. If s_2 is a μ-normal form, by [40, Theorem 6.3] we have $C[, \ldots,] = MRC^\mu(s_2) = MRC^\mu(t)$, and therefore, $s_2 = C[s'_1, \ldots, s'_n]$ and $t = C[t_1, \ldots, t_n]$ for some terms $s'_1, \ldots, s'_n, t_1, \ldots, t_n$, and for all $1 \leq i \leq n$, $s'_i \hookrightarrow_\mu^! t_i$. Since $C[, \ldots,] \neq \square$, by the induction hypothesis, for all $1 \leq i \leq n$, $t_i \in norm_\mu(s'_i)$ and according to (17), $t \in norm_\mu(s_2)$; hence, $t \in norm_\mu(s)$.
2. If s_2 is not a μ-normal form, then by the induction hypothesis, $t \in norm_\mu(s_2)$; hence, $t \in norm_\mu(s)$.

Left-linearity of $\mathcal{R}$ and $\mu \in CM_{\mathcal{R}}$ are necessary for Theorem 6 to hold.

Example 31 Consider $\mathcal{R}$ and μ in Example 30. We have $f(c) \in norm_\mu(f(a))$, but, since $f(c)$ is not a normal form, it can be reduced with *LCSR*. Furthermore, $f(a) \hookrightarrow\!\!\!\twoheadrightarrow_{\mathcal{R},\mu}^! b$, but $b \notin norm_\mu(f(a))$. With $\mathcal{R}'$ and μ' we have $g(c, c) \in norm_\mu(g(c, a))$, but $g(c, c)$ can be reduced with *LCSR*. Also, $g(c, a) \hookrightarrow\!\!\!\twoheadrightarrow_{\mathcal{R},\mu}^! b$, but $b \notin norm_\mu(g(c, a))$.

Theorem 6 enables the use of *LCSR* in a reduction-based description of $N\mu N$ which can be used to define computational complexity of $N\mu N$ as derivational complexity of $\hookrightarrow\!\!\!\twoheadrightarrow_\mu$. In order to do this, we need to guarantee that *LCSR* terminates when used with a CS-TRS $(\mathcal{R}, \mu)$. In the following section, we investigate this problem.

8.2 Termination of Layered *CSR*

We introduce a short notation for termination of $\hookrightarrow\!\!\!\twoheadrightarrow_{\mathcal{R},\mu}$.

Definition 12 Let $\mathcal{R}$ be a TRS and $\mu \in M_\mathcal{R}$. We say that $\mathcal{R}$ is $\widehat{\mu}$-terminating if $\hookrightarrow\!\!\!\twoheadrightarrow_{\mathcal{R},\mu}$ is terminating.

Termination of *LCSR* provides a sufficient condition for *normalization* of TRSs:

Proposition 16 *Let $\mathcal{R}$ be a TRS and $\mu \in M_\mathcal{R}$. If $\mathcal{R}$ is $\widehat{\mu}$-terminating, then $\mathcal{R}$ is normalizing.*

Proof Since $\mathcal{R}$ is $\widehat{\mu}$-terminating, for every term s there is t such that $s \hookrightarrow\!\!\!\twoheadrightarrow^!_\mu t$. By Proposition 13, t is a normal form, i.e., $\mathcal{R}$ is normalizing. ∎

Terminating TRSs are $\widehat{\mu}$-terminating for all $\mu \in M_\mathcal{R}$. And $\widehat{\mu}$-terminating TRSs are μ-terminating. We also have the following sufficient condition for $\widehat{\mu}$-termination.

Theorem 7 *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_\mathcal{R}$. If $\mathcal{R}$ is confluent, normalizing and μ-terminating, then $\mathcal{R}$ is $\widehat{\mu}$-terminating.*

Proof If $\mathcal{R}$ is not $\widehat{\mu}$-terminating, there is an infinite $\hookrightarrow\!\!\!\twoheadrightarrow_\mu$-sequence $t = t_1 \hookrightarrow\!\!\!\twoheadrightarrow t_2 \hookrightarrow\!\!\!\twoheadrightarrow \cdots$. Since $\mathcal{R}$ is normalizing and confluent, t has a single normal form $\bar{t}$ of size $N = |\bar{t}|$. By Corollary 3, there is $i \geq 1$ such that $|t_i| \geq N + 1$. By confluence, $t_i \rightarrow^* \bar{t}$ and by Corollary 3, $|\bar{t}| \geq N + 1$, leading to a contradiction. ∎

The following example shows that confluence cannot be dropped from Theorem 7.

Example 32 Consider the non-confluent TRS $\mathcal{R} = \{a \rightarrow b, a \rightarrow c(a)\}$ [36, Example 39] together with $\mu = \mu_\perp \in CM_\mathcal{R}$. Although $\mathcal{R}$ is μ-terminating and normalizing, we have the following infinite sequence of *LCSR*: $\underline{a} \hookrightarrow\!\!\!\twoheadrightarrow_\mu c(\underline{a}) \hookrightarrow\!\!\!\twoheadrightarrow_\mu c(c(\underline{a})) \hookrightarrow\!\!\!\twoheadrightarrow_\mu \cdots$.

Confluence and normalization of TRSs can often be automatically proved by using available tools, see CoCoWeb [31], for a unified platform of confluence tools, and FORT [52] for normalization (of right-ground TRSs).

Example 33 For $\mathcal{R}$ and μ in Example 21, the μ-termination proof in Example 26 is obtained by MU- TERM[7]. Confluence is proved by all CoCoWeb tools. Finally, $\mathcal{R}$ can easily be proved normalizing. Thus, $\mathcal{R}$ is $\widehat{\mu}$-terminating

8.3 Layered *CSR* as a Rewriting Strategy

In contrast to *CSR*, $\hookrightarrow\!\!\!\twoheadrightarrow_{\mathcal{R},\mu}$-normal forms are normal forms (Proposition 13), and therefore, *LCSR* always reduces terms which are not normal forms. Thus, *LCSR* can be seen as a (non-deterministic) one-step rewriting strategy

$$S_{LCSR}(t) = \{p \in \mathcal{P}os_\mathcal{R}(t) \mid (\exists u) \ t \xrightarrow{p}\!\!\!\!\!{}_\mu u\}$$

---

[7] Actually, MU- TERM automatically computes a strongly linear interpretation over the rationals as follows: $a^\mathcal{A} = 1$, $b^\mathcal{A} = 0$, $c^\mathcal{A}(x) = 0$, and $f^\mathcal{A}(x, y) = x + y + \frac{1}{2}$, from which we easily obtain the interpretation over the naturals in Example 26.

Now, the following generalization of Theorem 7 is immediate (the proof is analogous to that of Theorem 7).

Theorem 8 *Let $\mathcal{R}$ be a left-linear and confluent TRS, and $\mu \in CM_\mathcal{R}$. If $\mathcal{R}$ is μ-terminating, then S_{LCSR} is normalizing.*

Theorem 8 formalizes the observation in [40, Sect. 9.3], claiming that norm_μ is a *mapping* from terms to terms whenever $\mathcal{R}$ is left-linear and confluent, and $\mu \in CM_\mathcal{R}$ makes $\mathcal{R}$ μ-terminating.

9 Derivational Height of Layered Context-Sensitive Rewriting

Theorem 6 describes $N\mu N$ as normalization using *LCSR*. Thus, we investigate computational complexity of $N\mu N$ as the derivational complexity of *LCSR*:

Definition 13 Let $\mathcal{R}$ be a TRS and $\mu \in M_\mathcal{R}$ be such that $\mathcal{R}$ is $\widehat{\mu}$-terminating. Let

$$\mathsf{dh}_{\widehat{\mu}}(t) = \mathsf{dh}(t, \hookrightarrow\!\!\!\twoheadrightarrow_{\mathcal{R},\mu}) \qquad and \qquad \mathsf{dc}_{\widehat{\mu}}(n) = \max\{\mathsf{dh}_{\widehat{\mu}}(t) \mid |t| \le n\}$$

denote the derivational height and derivational complexity of *LCSR*, respectively.

In the remainder of the paper, we obtain bounds on $\mathsf{dc}_{\widehat{\mu}}(n)$ from bounds on $\mathsf{dc}_\mu(n)$. According to Proposition 14, for left-linear TRSs $\mathcal{R}$ and $\mu \in CM_\mathcal{R}$, rewrite sequences in *LCSR* can be decomposed in a (possibly empty) initial subsequence of context-sensitive rewriting steps followed of a possibly empty sequence of additional reductions issued on μ-normal forms. If $\mathcal{R}$ is μ-terminating, then the initial context-sensitive sequence must be finite.

Proposition 17 *Let $\mathcal{R}$ be a terminating, left-linear TRS and $\mu \in CM_\mathcal{R}$. For all terms t and $n \in \mathbb{N}$, $dh_{\widehat{\mu}}(t) \le dh(t, \to_\mathcal{R})$ and $dc_{\widehat{\mu}}(n) \le dc_\mathcal{R}(n)$.*

Remark 14 In the following, we use sums ($\sum$) and products ($\prod$) of numbers indexed by finite but potentially empty collections of indices. As usual, we give values 0 and 1, respectively, to sums and products of empty collections of numbers.

The following result connects dh_μ and $\mathsf{dh}_{\widehat{\mu}}$ on the basis of the definition of *LCSR*.

Proposition 18 *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_\mathcal{R}$ be such that $\mathcal{R}$ is $\widehat{\mu}$-terminating. Then,*

$$dh_{\widehat{\mu}}(s) \le dh_\mu(s) + \max_{\substack{s \hookrightarrow^!_\mu u = C[s_1, \ldots, s_{nmf_\mu(u)}] \\ C[\ldots,] = MRC^\mu(u)}} \sum_{i=1}^{nmf_\mu(u)} dh_{\widehat{\mu}}(s_i). \tag{23}$$

Proof Since $\mathcal{R}$ is $\widehat{\mu}$-terminating, it is also μ-terminating. By Proposition 14, every maximal normalization sequence $s = s_1 \hookrightarrow\!\!\!\twoheadrightarrow_\mu s_2 \hookrightarrow\!\!\!\twoheadrightarrow_\mu \cdots \hookrightarrow\!\!\!\twoheadrightarrow_\mu s_n = t$, where t is a normal form, can be written as follows:

$$s = s_1 \hookrightarrow_\mu s_2 \hookrightarrow_\mu \cdots \hookrightarrow_\mu s_m \hookrightarrow\!\!\!\twoheadrightarrow_\mu \cdots \hookrightarrow\!\!\!\twoheadrightarrow_\mu s_n = t$$

for some $m \ge 1$ such that $s_m, \ldots, s_n$ are μ-normal forms. Let $C[\ldots,] = MRC^\mu(s_m)$ a context with p holes. By [40, Theorem 6.3], for all $m \le i \le n$ we have $s_i = C[s_{i1}, \ldots, s_{ip}]$ for some terms $s_{i1}, \ldots, s_{ip}$. In particular, $t = C[t_1, \ldots, t_p]$ where $t_j = s_{nj}$ for all $1 \le j \le p$.

Moreover, for all $1 \leq j \leq p$, $s_{mj} \hookrightarrow^!_\mu t_j$. Therefore, $\mathsf{dh}_{\widehat{\mu}}(s) = n \leq m + \sum_{j=1}^p \mathsf{dh}_{\widehat{\mu}}(s_{mj})$. Since $m \leq \mathsf{dh}_\mu(s)$, we obtain

$$\mathsf{dh}_{\widehat{\mu}}(s) \leq \mathsf{dh}_\mu(s) + \sum_{j=1}^p \mathsf{dh}_{\widehat{\mu}}(s_{mj}). \tag{24}$$

Since there can be several μ-normal forms u of s, with different maximal μ-replacing contexts $MRC^\mu(u)$ with $\mathsf{nmf}_\mu(u)$ holes, from (24) we finally obtain (23).

Note that, in (23), we can replace the different values $\mathsf{nmf}_\mu(u)$ for μ-normal forms u of s by $\mathsf{NMF}_\mu^{\hookrightarrow !}(s)$ in Definition 5 to remove the big sum in (23):

$$\mathsf{dh}_{\widehat{\mu}}(s) \leq \mathsf{dh}_\mu(s) + \mathsf{NMF}_\mu^{\hookrightarrow !}(s) \max_{\substack{s \hookrightarrow^!_\mu u = C[s_1,\ldots,s_{\mathsf{nmf}_\mu(u)}] \\ C[,\ldots,] = MRC^\mu(u)}} \mathsf{dh}_{\widehat{\mu}}(s_i). \tag{25}$$

The advantage is that we can use the bounds on $\mathsf{NMF}_\mu^{\hookrightarrow !}(s)$ developed in Sect. 7.2 (Corollary 2 and Proposition 12) to obtain a translation of (25) into a bound on $\mathsf{dc}_{\widehat{\mu}}(n)$. However, we first have to deal with the maximal frozen subterms occurring in (25). The set of maximal frozen subterms s' which can be obtained from μ-normal forms u of s is:

$$\mathsf{MFRZ}_\mu(s) = \{u|_p \mid s \hookrightarrow^!_\mu u, \, p \in \mathcal{F}r^\mu(u)\}$$

Example 34 Consider the following TRS $\mathcal{R}$ from [15, Example 5]

$$\mathsf{f}(\mathsf{f}(x)) \to \mathsf{f}(\mathsf{g}(\mathsf{f}(x)))$$

with $\mu(\mathsf{f}) = \{1\}$ and $\mu(\mathsf{g}) = \varnothing$. We have the *innermost* μ-normalization sequence:

$$s = \mathsf{f}(\mathsf{f}(\underline{\mathsf{f}(\mathsf{f}(x)))}) \hookrightarrow \mathsf{f}(\underline{\mathsf{f}(\mathsf{f}(\mathsf{g}(\mathsf{f}(x))))}) \hookrightarrow \underline{\mathsf{f}(\mathsf{f}(\mathsf{g}(\mathsf{f}(\mathsf{g}(\mathsf{f}(x))))))}$$
$$\hookrightarrow \mathsf{f}(\mathsf{g}(\mathsf{f}(\mathsf{g}(\mathsf{f}(\mathsf{g}(\mathsf{f}(x))))))) = u_1$$

We have $\mathcal{F}r^\mu(u_1) = \{1.1\}$ and $s_1 = \mathsf{f}(\mathsf{g}(\mathsf{f}(\mathsf{g}(\mathsf{f}(x)))))$ is the maximal frozen subterm in u_1. The *outermost* μ-normalization of s is:

$$s = \underline{\mathsf{f}(\mathsf{f}(\mathsf{f}(\mathsf{f}(x))))} \hookrightarrow \mathsf{f}(\mathsf{g}(\mathsf{f}(\mathsf{f}(\mathsf{f}(x))))) = u_2$$

where the maximal frozen subterm of u_2 is $s_2 = \mathsf{f}(\mathsf{f}(\mathsf{f}(x)))$. There also are 'intermediate' μ-normalization sequences for s:

$$s = \mathsf{f}(\underline{\mathsf{f}(\mathsf{f}(\mathsf{f}(x))))} \hookrightarrow \underline{\mathsf{f}(\mathsf{f}(\mathsf{g}(\mathsf{f}(\mathsf{f}(x)))))} \hookrightarrow \mathsf{f}(\mathsf{g}(\mathsf{f}(\mathsf{g}(\mathsf{f}(\mathsf{f}(x)))))) = u_3$$
$$s = \mathsf{f}(\underline{\mathsf{f}(\mathsf{f}(\mathsf{f}(x))))} \hookrightarrow \underline{\mathsf{f}(\mathsf{f}(\mathsf{g}(\mathsf{f}(x))))} \hookrightarrow \mathsf{f}(\mathsf{g}(\mathsf{f}(\mathsf{f}(\mathsf{g}(\mathsf{f}(x)))))) = u_4$$

where the maximal frozen subterms of u_3 and u_4 are $s_3 = \mathsf{f}(\mathsf{g}(\mathsf{f}(\mathsf{f}(x))))$, and $s_4 = \mathsf{f}(\mathsf{f}(\mathsf{g}(\mathsf{f}(x))))$, respectively. Thus, $\mathsf{MFRZ}_\mu(s) = \{s_1, s_2, s_3, s_4\}$.

By using the previous notations, Proposition 18 leads to the following:

Theorem 9 *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_\mathcal{R}$ be such that $\mathcal{R}$ is $\widehat{\mu}$-terminating. Let s be a term. Then,*

$$dh_{\widehat{\mu}}(s) \leq dh_\mu(s) + NMF_\mu^{\hookrightarrow !}(s) \max_{s' \in MFRZ_\mu(s)} dh_{\widehat{\mu}}(s'). \tag{26}$$

For non-μ-duplicating TRSs, according to Proposition 12, we have

Corollary 4 *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_{\mathcal{R}}$ be such that $\mathcal{R}$ is non-μ-duplicating and $\widehat{\mu}$-terminating, and $F = F(\mathcal{R})$. Let s be a term. Then,*

1. *If $F > 0$, then $dh_{\widehat{\mu}}(s) \leq dh_{\mu}(s) + (nmf_{\mu}(s) + F dh_{\mu}(s)) \max_{s' \in MFRZ_{\mu}(s)} dh_{\widehat{\mu}}(s')$.*
2. *If $F = 0$, then $dh_{\widehat{\mu}}(s) \leq dh_{\mu}(s) + nmf_{\mu}(s) \max_{s' \in MFRZ_{\mu}(s)} dh_{\widehat{\mu}}(s')$.*
3. *If $F < 0$, then, $dh_{\widehat{\mu}}(s) \leq nmf_{\mu}(s)(\frac{1}{|F|} + \max_{s' \in MFRZ_{\mu}(s)} dh_{\widehat{\mu}}(s'))$.*

9.1 Use of Ground Terms in Layered Term Rewriting

In sharp contrast to the unrestricted case (see Sect. 4.1), derivational complexity of *LCSR* differs for ground and non-ground terms.

Example 35 Consider the following left-linear TRS $\mathcal{R}$ where a is a constant symbol:

$$f(a, y) \to a \tag{27}$$

with $\mu(f) = \{1\}$. Note that $\mu \in CM_{\mathcal{R}}$. For $t = f(x, f(f(a, a), a))$ (which is a μ-normal form), we have

$$t = f(x, f(\underline{f(a, a)}, a)) \hookrightarrow_{\mu} f(x, \underline{f(a, a)})$$
$$\hookrightarrow_{\mu} f(x, \underline{a})$$

of length 2. This is the only possible $\hookrightarrow_{\mu}$-sequence on t. However, if we make t ground by replacing the variable occurrence of x by a to obtain a term t' of the same size, we are only able to obtain a *shorter* sequence

$$t' = \underline{f(a, f(f(a, a), a))} \hookrightarrow_{\mu} a$$

consisting of a single context-sensitive rewriting step due to the fact that t' is *not* a μ-normal form now.

Thus, it is unclear whether an equality like (11) holds for *LCSR*.

In particular, Example 35 shows that *LCSR* is *not* closed under substitutions, even for left-linear TRSs $\mathcal{R}$ and $\mu \in CM_{\mathcal{R}}$: We have $t = f(x, f(\underline{f(a, a)}, a)) \hookrightarrow_{\mu} f(x, f(a, a)) = u$, but the instance t' of t does *not* $\hookrightarrow_{\mu}$-reduce to the corresponding instance $u' = f(a, f(a, a))$ of u.

For this reason, in the following we also consider *ground* terms in the analysis of derivational complexity of *LCSR*. Accordingly, we introduce the following

$$\mathsf{gdc}_{\widehat{\mu}}(n) = \max\{\mathsf{dh}_{\widehat{\mu}}(t) \mid t \in \mathcal{T}(\mathcal{F}), |t| \leq n\} \tag{28}$$

Clearly, for all $n \in \mathbb{N}$, $\mathsf{gdc}_{\widehat{\mu}}(n) \leq \mathsf{dc}_{\widehat{\mu}}(n)$.

10 Derivational Complexity of Normalization via μ-Normalization

We can use (26) to obtain bounds on $\mathsf{dc}_{\widehat{\mu}}(n)$ by replacing (i) $\mathsf{dh}_{\widehat{\mu}}(s)$ by $\mathsf{dc}_{\widehat{\mu}}(n)$ for terms s whose size is bounded by n, (ii) $\mathsf{dh}_{\mu}(s)$ by $\mathsf{dc}_{\mu}(n)$, (iii) $\mathsf{NMF}_{\mu}^{\hookrightarrow!}(s)$ by

$$\mathsf{BNMF}_{\mu}^{\hookrightarrow!}(n) = \max\{\mathsf{NMF}_{\mu}^{\hookrightarrow!}(s) \mid |s| \leq n\}$$

to obtain an upper bound on the (maximum) number $\mathsf{NMF}_{\mu}^{\hookrightarrow!}(s)$ of holes in $MRC^{\mu}(u)$ to be considered in each intermediate μ-normalization phase of $N\mu N$ when applied to terms s of

size $|s| \leq n$ to yield a μ-normal form u, and (iv) $\mathsf{dh}_{\widehat{\mu}}(s')$ by $\mathsf{dc}_{\widehat{\mu}}(n-1)$, thus assuming a decrease of $|s'|$ with respect to $|s|$ for all $s' \in \mathsf{MFRZ}_\mu(s)$ to obtain

$$\mathsf{dc}_{\widehat{\mu}}(n) \leq \mathsf{dc}_\mu(n) + \mathsf{BNMF}_\mu^{\hookrightarrow!}(n)\mathsf{dc}_{\widehat{\mu}}(n-1)$$

from which an explicit description of $\mathsf{dc}_{\widehat{\mu}}(n)$ in terms of $\mathsf{dc}_\mu(n)$ can be obtained. However, Example 34 shows that terms s' in $\mathsf{MFRZ}_\mu(s)$ can be *bigger* than s; for instance, $|s_1| = 6 > 5 = |s|$. Thus, in general, terms in $\mathsf{MFRZ}_\mu(s)$ do *not* decrease in size. Hence, item (iv) in the previous list may fail to be sound.

Definition 14 We say that MFRZ_μ is (strictly) *size-decreasing* if for all terms s and $s' \in \mathsf{MFRZ}_\mu(s)$, $|s| > |s'|$.

Proposition 19 *Let $\mathcal{R}$ be a TRS and $\mu \in M_\mathcal{R}$. If $\hookrightarrow^!$ is size-decreasing, then MFRZ_μ is size-decreasing.*

Proof If $t' \in \mathsf{MFRZ}_\mu(s)$, then there is a term t such that $s \hookrightarrow^! t$ and t' is a frozen (i.e., strict) subterm of t. By size-decreasingness of $\hookrightarrow^!$, $|s| \geq |t|$. Since $t \rhd t'$, $|s| > |t'|$.

Example 36 Consider $\mathcal{R}$ in Example 2 and μ in Example 15. As shown in Example 24, $\hookrightarrow^!_{\mathcal{R},\mu}$ is size-decreasing. By Proposition 19, MFRZ_μ is size-decreasing.

The following alternative result establishes conditions guaranteeing a decrease in the size of frozen subterms obtained at the end of *any* μ-rewrite sequence.

Proposition 20 *Let $\mathcal{R}$ be a TRS and $\mu \in M_\mathcal{R}$ be such that, for all $\ell \to r \in \mathcal{R}$, $\overline{\mathcal{P}os^\mu_\mathcal{F}}(r) = \varnothing$ and $\mathcal{V}ar^{\not\mu}(r) - \mathcal{V}ar^\mu(r) \subseteq \mathcal{V}ar^{\not\mu}(\ell)$. Then, for all terms s such that $s \hookrightarrow^*_\mu t$ and frozen subterms t' of t, $s \rhd_{\not\mu} t'$ and $|s| > |t'|$. Hence, MFRZ_μ is size-decreasing.*

Proof If t contains no frozen subterm, then the conclusion vacuously follows. Thus, assume that t contains a frozen subterm t', i.e., $t \rhd_{\not\mu} t'$. We proceed by induction on the length n of the sequence $s \hookrightarrow^*_\mu t$. If $n = 0$, then $s = t$, and therefore, $s \rhd_{\not\mu} t'$. If $n > 0$, then consider $s \hookrightarrow_\mu s' \hookrightarrow^*_\mu t$. We have $s|_p = \sigma(\ell)$ for some $p \in \mathcal{P}os^\mu(s)$, $\ell \to r \in \mathcal{R}$, and substitution σ. We also have $s' = s[\sigma(r)]_p$. By the induction hypothesis, $s' \rhd_{\not\mu} t'$, i.e., there is a position $q \in \overline{\mathcal{P}os^\mu}(s')$ such that $s'|_q = t'$. If $q \parallel p$, then $q \in \overline{\mathcal{P}os^\mu}(s)$ and $s \rhd_{\not\mu} t'$. Otherwise, $p < q$ and, since $\overline{\mathcal{P}os^\mu_\mathcal{F}}(r) = \varnothing$, there is a variable $x \in \mathcal{V}ar(r)$ such that either (i) $x \in \mathcal{V}ar^\mu(r)$ and $\sigma(x) \rhd_{\not\mu} t'$, or (ii) $x \in \mathcal{V}ar^{\not\mu}(r) - \mathcal{V}ar^\mu(r)$ and $\sigma(x) \unrhd t'$. In the first case (i), we have $\sigma(\ell) \rhd_{\not\mu} t'$, and hence, $s \rhd_{\not\mu} t'$. In the second case (ii), since $\mathcal{V}ar^{\not\mu}(r) - \mathcal{V}ar^\mu(r) \subseteq \mathcal{V}ar^{\not\mu}(\ell)$, we also have $s \rhd_{\not\mu} t'$. Since frozen positions always are below the root of terms, s can be written $s = f(s_1, \ldots, s_k)$ and we have $s_i \unrhd t'$ for some $1 \leq i \leq n$, i.e., $|s| > |s_i| \geq |t'|$.

Example 37 For $\mathcal{R}$ and μ in Example 5, $\hookrightarrow^!$ is *not* size-decreasing. For instance, $s = x \wedge (y \vee z) \hookrightarrow_{(8),\mu} (x \wedge y) \vee (x \wedge z) = t$ and t is a μ-normal form, but $|s| = 5 < 7 = |t|$. However, the conditions in Proposition 20 are fulfilled, thus proving MFRZ_μ size-decreasing: subterms at frozen positions in right-hand sides of rules are variables; as for frozen variables in left- and right-hand sides,

$$\mathcal{V}ar^{\not\mu}(r_{(8)}) - \mathcal{V}ar^\mu(r_{(8)}) = \{x\} - \{y, z\} = \{x\} = \{x\} = \mathcal{V}ar^{\not\mu}(\ell_{(8)})$$
$$\mathcal{V}ar^{\not\mu}(r_{(9)}) - \mathcal{V}ar^\mu(r_{(9)}) = \varnothing - \{x, y\} = \varnothing \subseteq \varnothing = \mathcal{V}ar^{\not\mu}(\ell_{(9)})$$
$$\mathcal{V}ar^{\not\mu}(r_{(10)}) - \mathcal{V}ar^\mu(r_{(10)}) = \{y\} - \{x, y\} = \varnothing \subseteq \varnothing = \mathcal{V}ar^{\not\mu}(\ell_{(10)})$$

11 Bounds on $dc_{\widehat{\mu}}(n)$

In this section, we use the previous results to obtain bounds on $dc_{\widehat{\mu}}(n)$.

Theorem 10 *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_{\mathcal{R}}$ be such that $\mathcal{R}$ is $\widehat{\mu}$-terminating and $MFRZ_\mu$ is size-decreasing. Then,*

$$dc_{\widehat{\mu}}(n) \leq \sum_{i=1}^{n} \left(\prod_{j=i+1}^{n} BNMF_\mu^{\leftrightarrow!}(j) \right) dc_\mu(i) \tag{29}$$

Proof By induction on n. If $n = 0$, then $dc_{\widehat{\mu}}(n) = 0$ and (29) trivially follows. If $n > 0$, then, by Theorem 9,

$$dh_{\widehat{\mu}}(s) \leq dh_\mu(s) + NMF_\mu^{\leftrightarrow!}(s) \max_{s' \in MFRZ_\mu(s)} dh_{\widehat{\mu}}(s')$$

By size-decreasingness of $MFRZ_\mu$, for all terms s and $s' \in MFRZ_\mu(s)$, $|s| > |s'|$, and

$$dc_{\widehat{\mu}}(n) \leq dc_\mu(n) + \max\{NMF_\mu^{\leftrightarrow!}(s) \max_{s' \in MFRZ_\mu(s)} dh_{\widehat{\mu}}(s')\} \mid |s| \leq n\}$$

$$\leq dc_\mu(n) + BNMF_\mu^{\leftrightarrow!}(n) \max\{dh_{\widehat{\mu}}(t) \mid |t| \leq n - 1\}$$

$$= dc_\mu(n) + BNMF_\mu^{\leftrightarrow!}(n)dc_{\widehat{\mu}}(n - 1)$$

By the induction hypothesis, $dc_{\widehat{\mu}}(n - 1) \leq \sum_{i=1}^{n-1} \left(\prod_{j=i+1}^{n-1} BNMF_\mu^{\leftrightarrow!}(j) \right) dc_\mu(i)$. Therefore,

$$dc_{\widehat{\mu}}(n) \leq dc_\mu(n) + BNMF_\mu^{\leftrightarrow!}(n) \sum_{i=1}^{n-1} \left(\prod_{j=i+1}^{n-1} BNMF_\mu^{\leftrightarrow!}(j) \right) dc_\mu(i)$$

$$= \sum_{i=1}^{n} \left(\prod_{j=i+1}^{n} BNMF_\mu^{\leftrightarrow!}(j) \right) dc_\mu(i)$$

In the following, given a replacement map $\mu \in M_{\mathcal{F}}$ and $n \in \mathbb{N}$, we let

$$bnmf_\mu(n) = \max_{t \in \mathcal{T}(\mathcal{F}, \mathcal{X}), |t| \leq n} nmf_\mu(t).$$

The following corollary of Proposition 5 provides some bounds for $bnmf_\mu(n)$.

Corollary 5 *Let $\mathcal{F}$ be a p-adic signature and $\mu \in M_{\mathcal{F}}$. Then, for all $n \in \mathbb{N}$,*

1. *If $\mu = \mu_\perp$, then $bnmf_\mu(n) \leq p$.*
2. *If $\mu = \mu_\top$, then $bnmf_\mu(n) = 0$.*
3. *If $p > 1$, and $\mathcal{F}$ is 1-active, then $bnmf_\mu(n) \leq \frac{p-1}{p}n$.*
4. *Otherwise, $bnmf_\mu(n) \leq n$.*

The following result characterizes some bounds on $bnmf_\mu(n)$.

Proposition 21 *Let $\mathcal{F}$ be a p-adic signature and $\mu \in M_{\mathcal{F}}$. Then, for all $n \in \mathbb{N}$,*

1. *$bnmf_\mu(n) = 0$ if and only if $\mu = \mu_\top$.*
2. *$bnmf_\mu(n) \leq p$ if and only if $\mu = \mu_\top$ or $\mu(f) = \varnothing$ for all $f \in \mathcal{F}$ with $ar(f) \geq 2$.*

Proof 1. The *if* part follows by Corollary 5(2). As for the *only if* part, we proceed by contradiction. Assume that, for all $n \in \mathbb{N}$, $\mathsf{bnmf}_\mu(n) = 0$ but $\mu \neq \mu_\top$. Then there is a k-ary symbol f with $k \geq 1$ such that $\mu(f) \subset \{1, \ldots, k\}$. Hence, the term $f(x_1, \ldots, x_k)$ of size $k+1$ contains at least a frozen subterm. Hence, $\mathsf{bnmf}_\mu(k+1) \neq 0$, a contradiction.

2. Regarding the *if* part, if $\mu = \mu_\top$, then, by Corollary 5(2), we have $\mathsf{bnmf}_\mu(n) = 0 \leq p$. Otherwise, we prove by structural induction on terms t that $\mathsf{nmf}_\mu(t) \leq p$. If t is a constant or a variable, then $\mathsf{nmf}_\mu(t) = 0 \leq p$. If $t = f(t_1, \ldots, t_k)$ for some symbol $f \in \mathcal{F}$ and $t_1, \ldots, t_k$, we consider two cases: (a) If f is a monadic symbol, then $t = f(t')$ and by the induction hypothesis, $\mathsf{nmf}_\mu(t') \leq p$. If $\mu(f) = \{1\}$, then, by the induction hypothesis, $\mathsf{nmf}_\mu(t) = \mathsf{nmf}_\mu(t') \leq p$; if $\mu(f) = \varnothing$, then $\mathsf{nmf}_\mu(t) = 1 \leq p$ (b) If $ar(f) > 1$, then $\mu(f) = \varnothing$ and $\mathsf{nmf}_\mu(t) = ar(f) \leq p$. Thus, for all terms t, $\mathsf{nmf}_\mu(t) \leq p$, i.e., $\mathsf{bnmf}_\mu(n) \leq p$.

Regarding the *only if* part, assume that $\mathsf{bnmf}_\mu(n) \leq p$, but (c) $\mu \neq \mu_\top$ and (d) $\mu(f) \neq \varnothing$ for some $f \in \mathcal{F}$ with $ar(f) \geq 2$. Then, by (d) there is $g \in \mathcal{F}$ with $ar(g) \geq 2$ and $|\mu(g)| \neq \varnothing$, i.e., $i_1 \in \mu(g)$ for some $1 \leq i_1 \leq ar(g)$. Let $i_2 \in \{1, \ldots, ar(g)\} - \{i_1\}$ be another argument index of g. Without loss of generality, we assume $i_1 < i_2$. By (c), there is $f \in \mathcal{F}$ with $j \notin \mu(f)$ for some $1 \leq j \leq ar(f)$. Let $s = f(y_1, \ldots, y_{i-1}, y, y_{i+1}, \ldots, y_q)$ and

$$t = g(x_1, \ldots, x_{i_1-1}, x, x_{i_1+1}, \ldots, x_{i_2-1}, x, x_{i_2+1}, \ldots x_p)$$

Note that y is frozen in s and x is active in position i_1 of t. Define $t_0 = s$ and for all $m > 0$ $t_m = \sigma_m(t)$, where $\sigma_m(x) = t_{m-1}$. We consider two cases: (A) If $i_2 \in \mu(g)$, then t contains at least two active positions (i_1 and i_2) and for all $m \geq 0$, t_m contains at least 2^m maximal frozen subterms (introduced by the occurrences of variable y in s when distributed in t_m). (B) If $i_2 \notin \mu(g)$, then for all $m \geq 0$, t_m contains at least $m+1$ maximal frozen subterms. Therefore, for a sufficiently big value of m, we have $\mathsf{nmf}_\mu(t_m) > p$, thus contradicting the bound for $\mathsf{bnmf}_\mu(n)$.

For non-μ-duplicating TRSs, we have the following.

Theorem 11 *Let $\mathcal{R}$ be a left-linear TRS and $\mu \in CM_\mathcal{R}$ be such that $\mathcal{R}$ is $\widehat{\mu}$-terminating and non-μ-duplicating, and $MFRZ_\mu$ is size-decreasing. Let $F = F(\mathcal{R})$.*

1. *If $F > 0$, then $\mathsf{dc}_{\widehat{\mu}}(n) \leq \sum_{i=1}^n \left(\prod_{j=i+1}^n (\mathsf{bnmf}_\mu(j) + F\mathsf{dc}_\mu(j)) \right) \mathsf{dc}_\mu(i)$.*
2. *If $F = 0$, then $\mathsf{dc}_{\widehat{\mu}}(n) \leq \sum_{i=1}^n \left(\prod_{j=i+1}^n \mathsf{bnmf}_\mu(j) \right) \mathsf{dc}_\mu(i)$.*
3. *If $F < 0$, then $\mathsf{dc}_{\widehat{\mu}}(n) \leq \sum_{i=1}^n \prod_{j=i+1}^n \mathsf{bnmf}_\mu(j)$.*

Proof Proceed like in Theorem 10, using Corollary 4 instead of Theorem 9 for the three considered cases.

The following result shows how asymptotic bounds on $\mathsf{dc}_{\widehat{\mu}}(n)$ and $\mathsf{gdc}_{\widehat{\mu}}(n)$ are obtained from asymptotic bounds on $\mathsf{dc}_\mu(n)$. Remind that we assume TRSs $\mathcal{R} = (\mathcal{F}, R)$ where $\mathcal{F}$ contains a constant symbol (see (14) for *CSR*); use $\mathcal{R}^\bullet = (\mathcal{F}^\bullet, R)$ otherwise, as explained in the previous sections, see Proposition 2. This is important when bounds depend on $\mathcal{F}_\mu^!$, which is defined assuming the existence of ground terms.

Corollary 6 *Let $\mathcal{R} = (\mathcal{F}, R)$ be a left-linear TRS such that $\mathcal{T}(\mathcal{F}) \neq \varnothing$ and $\mu \in CM_\mathcal{R}$ be such that $\mathcal{R}$ is $\widehat{\mu}$-terminating and $MFRZ_\mu$ is size-decreasing. Let $\mathsf{dc}_\mu(n) \in O(\mathfrak{f}(n))$.*

1. *If $\mathcal{F}_\mu^!$ is 0-adic, then $\mathsf{gdc}_{\widehat{\mu}}(n) \in O(\mathfrak{f}(n))$.*
2. *If for all $f \in \mathcal{F}_\mu^!$, $\mu(f) = \mu_\top(f)$, then $\mathsf{gdc}_{\widehat{\mu}}(n) \in O(\mathfrak{f}(n))$.*

3. If $\mathcal{F}_\mu^!$ is monadic and 0-active, then $gdc_{\widehat{\mu}}(n) \in O(n\mathfrak{f}(n))$.
4. If $\mathcal{F}_\mu^!$ is p-adic with $p > 1$, then

    (a) If $\mathcal{F}_\mu^!$ is 0-active, then $gdc_{\widehat{\mu}}(n) \in O(p^n\mathfrak{f}(n))$
    (b) If $\hookrightarrow^!$ is size-decreasing, then $gdc_{\widehat{\mu}}(n) \in O(n^n\mathfrak{f}(n))$

5. If $\mathcal{R}$ is non-μ-duplicating, then let $\eta_F \in \{-1, 0, 1\}$ denote the sign of $F(\mathcal{R})$.

    (a) If $F(\mathcal{R}) > 0$, then
        i. if $bnmf_\mu(n) \le b$ for some $b \in \mathbb{N}$, then $dc_{\widehat{\mu}}(n) \in n(\mathfrak{f}(n))^n 2^{O(n)}$,
        ii. otherwise, $dc_{\widehat{\mu}}(n) \in O((n\mathfrak{f}(n))^n)$.
    (b) If $F(\mathcal{R}) \le 0$, then
        i. if $bnmf_\mu(n) = 0$ for all $n \in \mathbb{N}$, then $dc_{\widehat{\mu}}(n) \in O((\mathfrak{f}(n))^{1+\eta_F})$,
        ii. if there is $b \ge 1$ such that $bnmf_\mu(n) \le b$ for all $n \in \mathbb{N}$, then $dc_{\widehat{\mu}}(n) \in O(nb^n\mathfrak{f}(n))$,
        iii. otherwise, $dc_{\widehat{\mu}}(n) \in O(n^n(\mathfrak{f}(n))^{1+\eta_F})$.

Proof First note that, since $\mathcal{T}(\mathcal{F}) \ne \varnothing$, we have $\mathcal{F}_\mu^! \ne \varnothing$.

1. For 0-adic signatures *CSR* and *LCSR* coincide, i.e., $gdc_{\widehat{\mu}}(n) = dc_\mu(n)$.
2. By Theorem 10 and Corollary 2(2), which applies to ground terms, and taking into account (14), $gdc_{\widehat{\mu}}(n) \le dc_\mu(n)$, i.e., $gdc_{\widehat{\mu}}(n) \in O(\mathfrak{f}(n))$.
3. By Theorem 10 and Corollary 2(1), $gdc_{\widehat{\mu}}(n) \le n dc_\mu(n)$, i.e., $gdc_{\widehat{\mu}}(n) \in O(n\mathfrak{f}(n))$.
4. (a) By Theorem 10 and Corollary 2(1), $gdc_{\widehat{\mu}}(n) \le \sum_{i=1}^{n}(\prod_{j=i+1}^{n} p) dc_\mu(i) = \sum_{i=1}^{n} p^{n-i} dc_\mu(i) = \sum_{i=0}^{n-1} p^i dc_\mu(i+1) \le \frac{1-p^n}{1-p} dc_\mu(n) \le p^n dc_\mu(n)$, i.e., $gdc_{\widehat{\mu}}(n) \in O(p^n\mathfrak{f}(n))$.
    (b) By Theorem 10 and Proposition 7, $gdc_{\widehat{\mu}}(n) \le \sum_{i=1}^{n}(\prod_{j=i+1}^{n} n) dc_\mu(i) = \sum_{i=1}^{n} n^{n-i} dc_\mu(i) \le n^n dc_\mu(n)$, i.e., $gdc_{\widehat{\mu}}(n) \in O(n^n\mathfrak{f}(n))$.
5. (a) If $F(\mathcal{R}) > 0$, by Theorem 11,

$$dc_{\widehat{\mu}}(n) \le dc_\mu(n) + \sum_{i=1}^{n}\left(\prod_{j=i+1}^{n}(bnmf_\mu(j) + F dc_\mu(j))\right) dc_\mu(i)$$

Thus, $dc_{\widehat{\mu}}(n) \le dc_\mu(n) + n(bnmf_\mu(n) + F dc_\mu(n))^{n-1} dc_\mu(n)$. If $bnmf_\mu(n) \le b$, then

$$
\begin{aligned}
dc_{\widehat{\mu}}(n) &\le dc_\mu(n) + n(bnmf_\mu(n) + F dc_\mu(n))^{n-1} dc_\mu(n) \\
&\le dc_\mu(n) + n(b + F dc_\mu(n))^{n-1} dc_\mu(n) \\
&\le n(1 + b + F)^{n-1}(dc_\mu(n))^n \\
&\in n(\mathfrak{f}(n))^n 2^{O(n)}
\end{aligned}
$$

Otherwise, we similarly obtain $dc_{\widehat{\mu}}(n) \in O((n\mathfrak{f}(n))^n)$.

    (b) By Theorem 11, if $F = 0$, then $dc_{\widehat{\mu}}(n) \le \sum_{i=1}^{n}\left(\prod_{j=i+1}^{n} bnmf_\mu(j)\right) dc_\mu(i)$; if $F < 0$, $dc_{\widehat{\mu}}(n) \le \sum_{i=1}^{n} \prod_{j=i+1}^{n} bnmf_\mu(j)$.
    As for the first case, (i) $bnmf_\mu(n) = 0$, if $F = 0$, we obtain $dc_{\widehat{\mu}}(n) \le dc_\mu(n) \in O(\mathfrak{f}(n))$; if $F < 0$, then $dc_{\widehat{\mu}}(n) \le \sum_{i=1}^{n} \prod_{j=i+1}^{n} 0 = 1 \in O(1)$. Both can be written at once as $dc_{\widehat{\mu}}(n) \in O((\mathfrak{f}(n))^{1+\eta_F})$. The other cases are similar.

Example 38 (*Running example II—bounds*) For $\mathcal{R}$ and μ in Example 2, $\mathcal{F}$ is 1-active and by Proposition 5, $bnmf_\mu(n) \le \frac{n}{2}$; $\mathcal{R}$ is non-μ-duplicating (Example 24); and $F(\mathcal{R}) = -1$

(Example 27). The strongly linear interpretation in Example 15 proves $\mathsf{dc}_\mu(n) \in O(n)$. By Corollary 6(5(b)iii), $\mathsf{dc}_{\widehat{\mu}}(n) \in O(n^n)$. This bound is *worse* than the quadratic bound for $\mathsf{dc}_{\mathcal{R}}(n)$ obtained in Example 13 (also by AProVE).

Example 39 *(Running example IV—improved bounds)* For $\mathcal{R}$ and μ in Example 5, $\mathcal{R}$ is non-μ-duplicating (Example 28); $F(\mathcal{R}) = 1$ (Example 27); and the linear interpretation in Example 17 proves $\mathsf{dc}_\mu(n) \in 2^{O(n)}$. By Corollary 6(5(a)ii), there is $c > 0$ such that $\mathsf{dc}_{\widehat{\mu}}(n) \in O((n2^{cn})^n)$. Since

$$\log_2((n2^{cn})^n) = n(\log_2 n + cn) \leq 2cn^2 \in O(n^2)$$

we get $\mathsf{dc}_{\widehat{\mu}}(n) \in 2^{O(n^2)}$, improving on the doubly exponential bound for $\mathsf{dc}_{\mathcal{R}}(n)$ obtained from the polynomial interpretation in Example 11 (see Example 12).

Example 40 *(Running example V—bounds)* For $\mathcal{R}$ and μ in Example 21, $\mathcal{R}$ is proved μ-terminating by using a strongly linear polynomial interpretation (Example 26). By Theorem 3(2), $\mathsf{dc}_\mu(n) \in O(n)$. Furthermore, $\mathcal{R}$ is $\widehat{\mu}$-terminating (Example 33). Since $\mathcal{F}^!_\mu = \{\mathsf{b}, \mathsf{c}\}$ is monadic and 0-active (Example 22), by Corollary 6 (3), $\mathsf{gdc}_{\widehat{\mu}}(n) \in O(n^2)$. Note that $\mathcal{R}$ is not terminating, and $\mathsf{dc}_{\mathcal{R}}(n)$ is undefined.

12 Related Work

As far as we know, this is the first paper explicitly defining and investigating derivational complexity of *CSR* and *NμN*. However, Hirokawa and Moser [30] pioneering the use of *CSR* in complexity analysis by applying *CSR* to investigate *runtime complexity analysis* of TRSs [29]. In runtime complexity analysis, all considered initial expressions s are *basic* terms $s \in \mathcal{T}_b(\mathcal{F}, \mathcal{X})$, i.e., terms $s = f(t_1, \ldots, t_k)$ where $t_1, \ldots, t_k$ are *constructor* terms. Given a terminating TRS, the runtime complexity of basic terms of size at most n is denoted $\mathsf{rc}_{\mathcal{R}}(n)$ and bounds on $\mathsf{rc}_{\mathcal{R}}(n)$ are obtained from matrix interpretations, among other techniques. Bounds on runtime complexity can be very different from bounds on derivational complexity. For instance, [29, Example 1] shows a TRS whose derivational complexity is *exponentially* bounded whereas its runtime complexity is proved *linear* in [30, Example 18]. Hirokawa and Moser show how to define replacement maps μ so that the compatibility of the ordering $\succ_{\mathcal{A}}$ induced by specific μ-monotone matrix interpretations $\mathcal{A}$ (restricted versions of TMIs, actually) can be used to correctly induce bounds on $\mathsf{rc}_{\mathcal{R}}(n)$. As a matter of fact, they were providing the *first analysis of runtime complexity of* CSR based on matrix interpretations: a μ-monotone matrix interpretation $\mathcal{A}$ whose associated ordering $\succ_{\mathcal{A}}$ is compatible with a TRS $\mathcal{R}$ actually proves μ-termination of $\mathcal{R}$. As discussed in the proofs of our Theorem 2, the validity of the polynomial bounds obtained from matrix interpretations $\mathcal{A}$ in [30, Sect. 2] does not depend on any monotonicity assumption. Thus, given a replacement map μ, they actually provide bounds on $\mathsf{rc}_{\mathcal{R},\mu}(n) = \max\{\mathsf{dh}(s, \hookrightarrow_\mu) \mid s \in \mathcal{T}_b(\mathcal{F}, \mathcal{X}), |s| \leq n\}$, the runtime complexity bound for *CSR*. Therefore, the results in [30, Sect. 2] can be used to obtain bounds on $\mathsf{rc}_{\mathcal{R},\mu}(n)$ for arbitrary replacement maps μ.[8]

---

[8] Although Hirokawa and Moser were not aware of this, in personal e-mail communication on early July 2020, they agreed with this analysis of their results.

13 Conclusions

We have investigated derivational complexity of *CSR*, with special emphasis in the analysis of derivational complexity (bounds) for normalization via μ-normalization, a normalization procedure based on *CSR* which has been recently implemented for use in Maude by using its strategy language (see Remark 1). An appropriate definition of derivational complexity for *CSR*, $dc_\mu(n)$, is given in Sect. 6. Some results showing how to obtain bounds on derivational complexity of *CSR* are given in Proposition 4 (on a purely syntactic basis), and Theorems 2 and 3, for proofs of μ-termination based on matrix and polynomial interpretations, respectively. Normalization via μ-normalization can be used not only with terminating TRSs but also with μ-terminating TRSs for which the procedure is guaranteed to terminate. Section 7 gives a characterization of normalization via μ-normalization using *layered CSR*, which permits to reuse the standard definition of derivational complexity for abstract reduction relations. Theorem 7 provides a criterion for termination of layered *CSR* which can be used to prove termination of $N\mu N$ to make sense of derivational complexity analysis of non-terminating TRSs. We provide full descriptions of the length $dh_{\widehat{\mu}}(s)$ of $N\mu N$ computations starting from a term s from the length of the intermediate context-sensitive rewrite sequences (Proposition 18 and Theorem 9). Such a description is used to characterize the derivational complexity $dc_{\widehat{\mu}}(n)$ of $N\mu N$ when used to evaluate terms of size at most n (Theorems 10 and 11). These results finally yield asymptotic bounds on $dc_{\widehat{\mu}}(n)$ and $gdc_{\widehat{\mu}}(n)$. In Corollary 6, we show how to obtain them from bounds on $dc_\mu(n)$.

Overall, we can benefit from *CSR* to improve bounds on computational complexity of normalization in different ways. First, dealing with *non-terminating* but μ-terminating and normalizing TRSs,

1. If *CSR* can be seen as a rewriting strategy (because μ-normal forms and normal forms coincide), then $dc_\mu(n)$ can be used instead of the undefined $dc_\mathcal{R}(n)$. The running example I (Example 1) illustrates this use.
2. If μ-normal forms are not normal forms, then $N\mu N$ can be used to obtain them, provided that $\mathcal{R}$ is $\widehat{\mu}$-terminating, as in running example V (Example 21).

Dealing with terminating TRSs, we can also obtain some advantages:

1. If μ-normal forms and normal forms coincide, then $dc_\mu(n)$ can be used instead of $dc_\mathcal{R}(n)$. The running example III (Example 3) illustrates this.
2. Otherwise, we can use $N\mu N$ to obtain normal forms. Sometimes $dc_{\widehat{\mu}}(n)$ yields better estimations than $dc_\mathcal{R}(n)$ (e.g., the running example IV (Example 5). In other cases, no improvement is obtained (e.g., running example II, Example 2).

Future work

Extending this research to also consider *lower* bounds on derivational complexity of *CSR* and $N\mu N$ would be interesting. Also considering runtime complexity of $N\mu N$ is interesting. More research would be necessary to remove the size-decreasingness requirement in some of our results as it imposes quite hard requirements on TRSs to obtain explicit asymptotic bounds. The point is evident by comparing, e.g., Theorem 10 with Theorem 9, establishing the (recursive) relationship between derivational length of *CSR* and $N\mu N$, where size-decreasingness is not necessary. This suggests that size-decreasingness is not an intrinsic feature of the description of derivational complexity of $N\mu N$. Unfortunately, the attempts to get rid from this restriction failed to date.

References

1. Alarcón, B., Lucas, S., Navarro-Marset, R.: Proving termination with matrix interpretations over the reals. In: Geser, A., Waldmann, J. (eds.). Proceedings of the 10th International Workshop on Termination, WST 2009), HTWK Leipzig, pp. 12–15 (2009a)
2. Alarcón, B., Lucas, S., Navarro-Marset, R.: Using matrix interpretations over the reals in proofs of termination. In: Lucio, P., Moreno, G., Peña, R. (eds.). Actas de las IX Jornadas de Programación y Lenguajes (PROLE 2009), SISTEDES, pp. 268–2–77 (2009b)
3. Arts, T., Giesl, J.: Termination of term rewriting using dependency pairs. Theor. Comput. Sci. **236**(1–2), 133–178 (2000). https://doi.org/10.1016/S0304-3975(99)00207-8
4. Arts, T., Giesl, J.: A collection of examples for termination of term rewriting using dependency pairs. Tech. Rep. AIB-2001-09, RWTH Aachen, Germany (2001)
5. Baader, F., Nipkow, T.: Term Rewriting and All That. Cambridge University Press (1998)
6. Barendregt, H.P., van Eekelen, M.C.J.D., Glauert, J.R.W., Kennaway, R., Plasmeijer, M.J., Sleep, M.R.: Term graph rewriting. In: de Bakker, J.W., Nijman, A.J., Treleaven, P.C. (eds.). PARLE (2), Springer, Lecture Notes in Computer Science, vol. 259, pp. 141–158 (1987)
7. Bonfante, G., Cichon, A., Marion, J., Touzet, H.: Complexity classes and rewrite systems with polynomial interpretation. In: Gottlob, G., Grandjean, E., Seyr, K. (eds.). Computer Science Logic, 12th International Workshop, CSL '98, Annual Conference of the EACSL, Brno, Czech Republic, August 24–28, 1998, Proceedings, Springer, Lecture Notes in Computer Science, vol. 1584, pp. 372–384, (1998) https://doi.org/10.1007/10703163_25
8. Bonfante, G., Cichon, A., Marion, J., Touzet, H.: Algorithms with polynomial interpretation termination proof. J. Funct. Prog. **11**(1), 33–53 (2001)
9. Cichon, A., Lescanne, P.: Polynomial interpretations and the complexity of algorithms. In: Kapur, D. (ed.). Automated Deduction: CADE-11, 11th International Conference on Automated Deduction, Saratoga Springs, NY, USA, June 15–18, 1992, Proceedings, Springer, Lecture Notes in Computer Science, vol. 607, pp. 139–147, (1992) https://doi.org/10.1007/3-540-55602-8_161
10. Cichon, E.: Termination orderings and complexity characterizations. In: Aczell, S., Wainer (eds.). Proof Theory: a selection of papers from the Leeds Proof Theory Programme, 1990, Cambridge University Press, Lecture Notes in Computer Science, vol. 607, pp. 171–194 (1992)
11. Clavel, M., Durán, F., Eker, S., Lincoln, P., Martí-Oliet, N., Meseguer, J., Talcott, C.L.: All About Maude: A High-Performance Logical Framework, How to Specify, Program and Verify Systems in Rewriting Logic, Lecture Notes in Computer Science, vol. 4350. Springer (2007). https://doi.org/10.1007/978-3-540-71999-1
12. Contejean, E., Marché, C., Tomás, A.P., Urbain, X.: Mechanically proving termination using polynomial interpretations. J. Autom. Reason. **34**(4), 325–363 (2005). https://doi.org/10.1007/s10817-005-9022-x
13. Courcelle, B.: Recursive applicative program schemes. In: van Leeuwen, J. (ed.). Handbook of Theoretical Computer Science, Volume B: Formal Models and Semantics, Elsevier and MIT Press, pp. 459–492 (1990) https://doi.org/10.1016/b978-0-444-88074-1.50014-7
14. Dershowitz, N.: A note on simplification orderings. Inf. Process. Lett. **9**(5), 212–215 (1979). https://doi.org/10.1016/0020-0190(79)90071-1
15. Dershowitz, N.: 33 examples of termination. In: Comon, H., Jouannaud, J. (eds.). Term Rewriting, French Spring School of Theoretical Computer Science, Font Romeux, France, May 17–21, 1993, Advanced

Course, Springer, Lecture Notes in Computer Science, vol. 909, pp. 16–26, (1993) https://doi.org/10.1007/3-540-59340-3_2

16. Durán, F., Escobar, S., Lucas, S.: New Evaluation Commands for Maude Within Full Maude. In: Martí-Oliet, N. (ed.). Proceedings of the Fifth International Workshop on Rewriting Logic and Its Applications, WRLA 2004, Barcelona, Spain, March 27–28, 2004, Elsevier, Electronic Notes in Theoretical Computer Science, vol. 117, pp. 263–284, (2004) https://doi.org/10.1016/j.entcs.2004.06.014

17. Durán, F., Eker, S., Escobar, S., Martí-Oliet, N., Meseguer, J., Rubio, R., Talcott, C.L.: Programming and symbolic computation in Maude. J. Log. Algebra Method Prog. (2020). https://doi.org/10.1016/j.jlamp.2019.100497

18. Endrullis, J., Waldmann, J., Zantema, H.: Matrix interpretations for proving termination of term rewriting. J. Autom. Reason. **40**(2–3), 195–220 (2008)

19. Fasel, J.H., Hudak, P., Jones, S.L.P., Wadler, P.: SIGPLAN notices special issue on the functional programming language haskell. ACM SIGPLAN Not. **27**(5), 1 (1992)

20. Futatsugi, K., Nakagawa, A.T.: An overview of CAFE specification environment: an algebraic approach for creating, verifying, and maintaining formal specifications over networks. In: First IEEE International Conference on Formal Engineering Methods, ICFEM 1997, Hiroshima, Japan, November 12–14, 1997, Proceedings, IEEE Computer Society, pp. 170–182, (1997) https://doi.org/10.1109/ICFEM.1997.630424

21. Giesl, J., Thiemann, R., Schneider-Kamp, P., Falke, S.: Mechanizing and improving dependency pairs. J. Autom. Reason. **37**(3), 155–203 (2006). https://doi.org/10.1007/s10817-006-9057-7

22. Giesl, J., Aschermann, C., Brockschmidt, M., Emmes, F., Frohn, F., Fuhs, C., Hensel, J., Otto, C., Plücker, M., Schneider-Kamp, P., Ströder, T., Swiderski, S., Thiemann, R.: Analyzing Program Termination and Complexity Automatically with AProVE. J. Autom. Reason. **58**(1), 3–31 (2017). https://doi.org/10.1007/s10817-016-9388-y

23. Giesl, J., Rubio, A., Sternagel, C., Waldmann, J., Yamada, A.: The termination and complexity competition. In: Beyer, D., Huisman, M., Kordon, F., Steffen, B. (eds.). Tools and Algorithms for the Construction and Analysis of Systems: 25 Years of TACAS: TOOLympics, Held as Part of ETAPS 2019, Prague, Czech Republic, April 6–11, 2019, Proceedings, Part III, Springer, Lecture Notes in Computer Science, vol. 11429, pp. 156–166, (2019) https://doi.org/10.1007/978-3-030-17502-3_10

24. Goguen, J.A., Winkler, T., Meseguer, J., Futatsugi, K., Jouannaud, J.P.: Introducing obj Software Engineering with OBJ: algebraic specification in action:169–236 (2000)

25. Gramlich, B., Lucas, S.: Modular termination of context-sensitive rewriting. In: Proceedings of the 4th International ACM SIGPLAN Conference on Principles and Practice of Declarative Programming, October 6-8, 2002, Pittsburgh, PA, USA (Affiliated with PLI 2002), ACM, pp. 50–61, (2002a) https://doi.org/10.1145/571157.571163

26. Gramlich, B., Lucas, S.: Simple termination of context-sensitive rewriting. In: Fischer, B., Visser, E. (eds.). Proceedings of the 2002 ACM SIGPLAN Workshop on Rule-Based Programming, Pittsburgh, Pennsylvania, USA, 2002, ACM, pp. 29–42, (2002b) https://doi.org/10.1145/570186.570189

27. Grzegorczyk, A.: Some classes of recursive functions. J. Symbol. Logic **20**(1), 71–72 (1955). https://doi.org/10.2307/2268082

28. Gutiérrez, R., Lucas, S.: MU-TERM: Verify termination properties automatically (system description). In: Peltier, N., Sofronie-Stokkermans, V. (eds.). Automated Reasoning: 10th International Joint Conference, IJCAR 2020, Paris, France, July 1–4, 2020, Proceedings, Part II, Springer, Lecture Notes in Computer Science, vol. 12167, pp. 436–447, (2020) https://doi.org/10.1007/978-3-030-51054-1_28

29. Hirokawa, N., Moser, G.: Automated complexity analysis based on the dependency pair method. In: Armando, A., Baumgartner, P., Dowek, G. (eds.). Automated Reasoning, 4th International Joint Conference, IJCAR 2008, Sydney, Australia, August 12–15, 2008, Proceedings, Springer, Lecture Notes in Computer Science, vol. 5195, pp. 364–379, (2008) https://doi.org/10.1007/978-3-540-71070-7_32

30. Hirokawa, N., Moser, G.: Automated complexity analysis based on context-sensitive rewriting. In: Dowek, G. (ed.). Rewriting and Typed Lambda Calculi: Joint International Conference, RTA-TLCA 2014, Held as Part of the Vienna Summer of Logic, VSL 2014, Vienna, Austria, July 14–17, 2014. Proceedings, Springer, Lecture Notes in Computer Science, vol. 8560, pp. 257–271, (2014) https://doi.org/10.1007/978-3-319-08918-8_18

31. Hirokawa, N., Nagele, J., Middeldorp, A.: Cops and CoCoWeb: Infrastructure for Confluence Tools. In: Galmiche, D., Schulz, S., Sebastiani, R. (eds.). Automated Reasoning: 9th International Joint Conference, IJCAR 2018, Held as Part of the Federated Logic Conference, FloC 2018, Oxford, UK, July 14–17, 2018, Proceedings, Springer, Lecture Notes in Computer Science, vol. 10900, pp. 346–353, (2018) https://doi.org/10.1007/978-3-319-94205-6_23

32. Hofbauer, D.: Termination proofs by multiset path orderings imply primitive recursive derivation lengths. Theor. Comput. Sci. **105**(1), 129–140 (1992). https://doi.org/10.1016/0304-3975(92)90289-R

33. Hofbauer, D., Lautemann, C.: Termination proofs and the length of derivations (preliminary version). In: Dershowitz, N. (ed.). Rewriting Techniques and Applications, 3rd International Conference, RTA-89, Chapel Hill, North Carolina, USA, April 3–5, 1989, Proceedings, Springer, Lecture Notes in Computer Science, vol. 355, pp. 167–177, (1989) https://doi.org/10.1007/3-540-51081-8_107

34. Knuth, D.E., Bendix, P.E.: Simple word problems in universal algebra. In: Leech, J. (ed.). Computational Problems in Abstract Algebra, Pergamon Press, pp. 263–297 (1970)

35. Lankford, D.S.: On proving term rewriting systems are noetherian. Louisiana Technological University, Ruston, LA, Tech. rep (1979)

36. Lucas, S.: Context-sensitive rewriting strategies. Inf. Comput. **178**(1), 294–343 (2002a). https://doi.org/10.1006/inco.2002.3176

37. Lucas, S.: Termination of (canonical) context-sensitive rewriting. In: Tison, S. (ed.). RTA, Springer, Lecture Notes in Computer Science, vol. 2378, pp. 296–310 (2002b)

38. Lucas, S.: A note on completeness of conditional context-sensitive rewriting. In: 5th International Workshop on Reduction Strategies in Rewriting and Programming, pp. 3–15. University of Kyoto, WRS (2005a)

39. Lucas, S.: Polynomials over the reals in proofs of termination: from theory to practice. ITA **39**(3), 547–586 (2005b). https://doi.org/10.1051/ita:2005029

40. Lucas, S.: Context-sensitive rewriting. ACM Comput. Surv. **53**(4), 78:1–78:36 (2020). https://doi.org/10.1145/3397677

41. Lucas, S.: Applications and extensions of context-sensitive rewriting. J. Logical Algebraic Methods Prog. **121**, 100680 (2021). https://doi.org/10.1016/j.jlamp.2021.100680

42. Manna, Z., Ness, S.: On the termination of Markov algorithms. In: Proceedings of the Third Hawaii International Conference on System Science, pp. 789–792 (1970)

43. McCarthy, J.: Towards a mathematical science of computation. In: Information Processing, Proceedings of the 2nd IFIP Congress 1962, Munich, Germany, August 27–September 1, 1962, North-Holland, pp. 21–28 (1962)

44. Middeldorp, A.: Call by need computations to root-stable form. In: Lee, P., Henglein, F., Jones, N.D. (eds.). Conference Record of POPL'97: The 24th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, Papers Presented at the Symposium, Paris, France, 15–17 January 1997, ACM Press, pp. 94–105 (1997) https://doi.org/10.1145/263699.263711

45. Moser, G., Schnabl, A.: The derivational complexity induced by the dependency pair method. Logical Methods Comput. Sci. (2011). https://doi.org/10.2168/LMCS-7(3:1)2011

46. Moser, G., Schnabl, A., Waldmann, J.: Complexity analysis of term rewriting based on matrix and context dependent interpretations. In: Hariharan, R., Mukund, M., Vinay, V. (eds.). IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2008, December 9–11, 2008, Bangalore, India, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, LIPIcs, vol. 2, pp. 304–315 (2008). https://doi.org/10.4230/LIPIcs.FSTTCS.2008.1762

47. Neurauter, F., Middeldorp, A.: Revisiting matrix interpretations for proving termination of term rewriting. In: Schmidt-Schauß, M. (ed.). Proceedings of the 22nd International Conference on Rewriting Techniques and Applications, RTA 2011, May 30–June 1, 2011, Novi Sad, Serbia, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, LIPIcs, vol. 10, pp. 251–266 (2011), https://doi.org/10.4230/LIPIcs.RTA.2011.251

48. Neurauter, F., Zankl, H., Middeldorp, A.: Revisiting matrix interpretations for polynomial derivational complexity of term rewriting. In: Fermüller, C.G., Voronkov, A. (eds.). Logic for Programming, Artificial Intelligence, and Reasoning - 17th International Conference, LPAR-17, Yogyakarta, Indonesia, October 10–15, 2010. Proceedings, Springer, Lecture Notes in Computer Science, vol. 6397, pp. 550–564 (2010), https://doi.org/10.1007/978-3-642-16242-8_39

49. Noschinski, L., Emmes, F., Giesl, J.: Analyzing innermost runtime complexity of term rewriting by dependency pairs. J. Autom. Reason. **51**(1), 27–56 (2013). https://doi.org/10.1007/s10817-013-9277-6

50. Ohlebusch, E.: Advanced Topics in Term Rewriting. Springer (2002)

51. Péchoux, R.: Synthesis of sup-interpretations: a survey. Theor. Comput. Sci. **467**, 30–52 (2013). https://doi.org/10.1016/j.tcs.2012.11.003

52. Rapp, F., Middeldorp, A.: FORT 2.0. In: Galmiche, D., Schulz, S., Sebastiani, R. (eds.). Automated Reasoning: 9th International Joint Conference, IJCAR 2018, Held as Part of the Federated Logic Conference, FloC 2018, Oxford, UK, July 14–17, 2018, Proceedings, Springer, Lecture Notes in Computer Science, vol. 10900, pp. 81–88 (2018), https://doi.org/10.1007/978-3-319-94205-6_6

53. Ritchie, R.W.: Classes of predictably computable functions. Trans. Am. Math. Soc. **106**, 139–173 (1963)

54. Rubio, R., Martí-Oliet, N., Pita, I., Verdejo, A.: Parameterized Strategies Specification in Maude. In: Fiadeiro, J.L., Tutu, I. (eds.). Recent Trends in Algebraic Development Techniques: 24th IFIP WG 1.3 International Workshop, WADT 2018, Egham, UK, July 2-5, 2018, Revised Selected Papers, Springer,

Lecture Notes in Computer Science, vol. 11563, pp. 27–44 (2018), https://doi.org/10.1007/978-3-030-23220-7_2

55. Steinbach, J.: Proving polynomials positive. In: Shyamasundar, R.K. (ed.). Foundations of Software Technology and Theoretical Computer Science, 12th Conference, New Delhi, India, December 18–20, 1992, Proceedings, Springer, Lecture Notes in Computer Science, vol. 652, pp. 191–202 (1992), https://doi.org/10.1007/3-540-56287-7_105

56. Sternagel, C., Thiemann, R.: Signature extensions preserve termination: an alternative proof via dependency pairs. In: Dawar, A., Veith, H. (eds.). Computer Science Logic, 24th International Workshop, CSL 2010, 19th Annual Conference of the EACSL, Brno, Czech Republic, August 23–27, 2010. Proceedings, Springer, Lecture Notes in Computer Science, vol. 6247, pp. 514–528 (2010), https://doi.org/10.1007/978-3-642-15205-4_39

57. Terese.: Term Rewriting Systems, Cambridge tracts in theoretical computer science, vol. 55. Cambridge University Press (2003)

58. Waldmann, J.: Polynomially bounded matrix interpretations. In: Lynch, C. (ed.). Proceedings of the 21st International Conference on Rewriting Techniques and Applications, RTA 2010, July 11–13, 2010, Edinburgh, Scotland, UK, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, LIPIcs, vol. 6, pp. 357–372 (2010), https://doi.org/10.4230/LIPIcs.RTA.2010.357

59. Weiermann, A.: Termination proofs for term rewriting systems by lexicographic path orderings imply multiply recursive derivation lengths. Theor. Comput.Sci. **139**(1 & 2), 355–362 (1995). https://doi.org/10.1016/0304-3975(94)00135-6

60. Zantema, H.: Termination of term rewriting: interpretation and type elimination. J. Symb. Comput. **17**(1), 23–50 (1994). https://doi.org/10.1006/jsco.1994.1003

61. Zantema, H.: Termination of context-sensitive rewriting. In: Comon, H. (ed.). Rewriting Techniques and Applications, 8th International Conference, RTA-97, Sitges, Spain, June 2–5, 1997, Proceedings, Springer, Lecture Notes in Computer Science, vol. 1232, pp. 172–186 (1997), https://doi.org/10.1007/3-540-62950-5_69