Improving Context-Sensitive Dependency Pairs*

Beatriz Alarcón[1], Fabian Emmes[2], Carsten Fuhs[2], Jürgen Giesl[2], Raúl Gutiérrez[1], Salvador Lucas[1], Peter Schneider-Kamp[2], and René Thiemann[3]

[1] DSIC, Universidad Politécnica de Valencia, Spain
[2] LuFG Informatik 2, RWTH Aachen University, Germany
[3] Institute of Computer Science, University of Innsbruck, Austria

Abstract. Context-sensitive dependency pairs (CS-DPs) are currently the most powerful method for automated termination analysis of context-sensitive rewriting. However, compared to DPs for ordinary rewriting, CS-DPs suffer from two main drawbacks: (a) CS-DPs can be *collapsing*. This complicates the handling of CS-DPs and makes them less powerful in practice. (b) There does not exist a "*DP framework*" for CS-DPs which would allow one to apply them in a flexible and modular way. This paper solves drawback (a) by introducing a new definition of CS-DPs. With our definition, CS-DPs are always non-collapsing and thus, they can be handled like ordinary DPs. This allows us to solve drawback (b) as well, i.e., we extend the existing DP framework for ordinary DPs to context-sensitive rewriting. We implemented our results in the tool AProVE and successfully evaluated them on a large collection of examples.

1 Introduction

Context-sensitive rewriting [23, 24] models evaluations in programming languages. It uses a *replacement map* μ with $\mu(f) \subseteq \{1, ..., \text{arity}(f)\}$ for every function symbol f to specify the argument positions of f where rewriting may take place.

Example 1. Consider this context-sensitive term rewrite system (CS-TRS)

$$\begin{array}{rlrl}
\mathsf{gt}(0, y) &\to \mathsf{false} & \mathsf{p}(0) &\to 0 \\
\mathsf{gt}(\mathsf{s}(x), 0) &\to \mathsf{true} & \mathsf{p}(\mathsf{s}(x)) &\to x \\
\mathsf{gt}(\mathsf{s}(x), \mathsf{s}(y)) &\to \mathsf{gt}(x, y) & \mathsf{minus}(x, y) &\to \mathsf{if}(\mathsf{gt}(y, 0), \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y)), x) \\
\mathsf{if}(\mathsf{true}, x, y) &\to x & \mathsf{div}(0, \mathsf{s}(y)) &\to 0 \\
\mathsf{if}(\mathsf{false}, x, y) &\to y & \mathsf{div}(\mathsf{s}(x), \mathsf{s}(y)) &\to \mathsf{s}(\mathsf{div}(\mathsf{minus}(x, y), \mathsf{s}(y)))
\end{array} \quad (1)$$

with $\mu(\mathsf{if}) = \{1\}$ and $\mu(f) = \{1, \ldots, \text{arity}(f)\}$ for all other symbols f to model the usual behavior of if*: in $\mathsf{if}(t_1, t_2, t_3)$, one may evaluate t_1, but not t_2 or t_3. It will turn out that due to μ, this CS-TRS is indeed terminating. In contrast, if one allows arbitrary reductions, then the TRS would be non-terminating:*

* *Proc. LPAR '08*, LNAI, 2008. Authors from Valencia were partially supported by the EU (FEDER) and the Spanish MEC/MICINN, grants TIN 2007-68093-C02-02 and HA 2006-0007. B. Alarcón was partially supported by the Spanish MEC/MICINN, FPU grant AP2005-3399. R. Gutiérrez was partially supported by the Spanish MEC/MICINN, grant TIN 2004-7943-C04-02. Authors from Aachen were supported by the DAAD under grant D/06/12785 and by the DFG under grant GI 274/5-2.

$\mathsf{minus}(0,0) \to^+ \mathsf{if}(\mathsf{gt}(0,0), \mathsf{minus}(0,0), 0) \to^+ \mathsf{if}(..., \mathsf{if}(\mathsf{gt}(0,0), \mathsf{minus}(0,0), 0), ...) \to^+ ...$

There are two approaches to prove termination of context-sensitive rewriting. The first approach transforms CS-TRSs to ordinary TRSs, cf. [13, 26]. But transformations often generate complicated TRSs where all termination tools fail.

Therefore, it is more promising to adapt existing termination techniques from ordinary term rewriting to the context-sensitive setting. Such adaptions were done for classical methods like RPO or polynomial orders [8, 19, 25]. However, much more powerful techniques like the *dependency pair* (DP) method [6] are implemented in almost all current termination tools for TRSs. But for a long time, it was not clear how to adapt the DP method to context-sensitive rewriting.

This was solved first in [1]. The corresponding implementation in the tool MU-TERM [3] outperformed all previous tools for termination of CS rewriting.

Nevertheless, the existing results on CS-DPs [1, 2, 4, 20] still have major disadvantages compared to the DP method for ordinary rewriting, since CS-DPs can be *collapsing*. To handle such DPs, one has to impose strong requirements which make the CS-DP method quite weak and which make it difficult to extend refined termination techniques based on DPs to the CS case. In particular, the *DP framework* [14, 17, 21], which is the most powerful formulation of the DP method for ordinary TRSs, has not yet been adapted to the CS setting.

In this paper, we solve these problems. After presenting preliminaries in Sect. 2, we introduce a new notion of *non-collapsing* CS-DPs in Sect. 3. This new notion makes it much easier to adapt termination techniques based on DPs to context-sensitive rewriting. Therefore, Sect. 4 extends the *DP framework* to the context-sensitive setting and shows that existing methods from this framework only need minor changes to apply them to context-sensitive rewriting.

All our results are implemented in the termination prover AProVE [16]. As shown by the empirical evaluation in Sect. 5, our contributions improve the power of automated termination analysis for context-sensitive rewriting substantially.

2 Context-Sensitive Rewriting and CS-Dependency Pairs

See [7] and [23] for basics on term rewriting and context-sensitive rewriting, respectively. Let $\mathcal{P}os(s)$ be the set of *positions* of a term s. For a replacement map μ, we define the *active positions* $\mathcal{P}os^\mu(s)$: For $x \in \mathcal{V}$ let $\mathcal{P}os^\mu(x) = \{\varepsilon\}$ where ε is the root position. Moreover, $\mathcal{P}os^\mu(f(s_1, \ldots, s_n)) = \{\varepsilon\} \cup \{i\,p \mid i \in \mu(f),\ p \in \mathcal{P}os^\mu(s_i)\}$. We say that $s \trianglerighteq_\mu t$ holds if $t = s|_p$ for some $p \in \mathcal{P}os^\mu(s)$ and $s \triangleright_\mu t$ if $s \trianglerighteq_\mu t$ and $s \neq t$. Moreover, $s \triangleright_\mu^! t$ if $t = s|_p$ for some $p \in \mathcal{P}os(s) \setminus \mathcal{P}os^\mu(s)$. We denote the ordinary subterm relations by $\trianglerighteq$ and $\triangleright$.

A *CS-TRS* $(\mathcal{R}, \mu)$ consists of a finite TRS $\mathcal{R}$ and a replacement map μ. We have $s \hookrightarrow_{\mathcal{R},\mu} t$ iff there are $\ell \to r \in \mathcal{R}$, $p \in \mathcal{P}os^\mu(s)$, and a substitution σ with $s|_p = \sigma(\ell)$ and $t = s[\sigma(r)]_p$. This reduction is an *innermost* step (denoted $\overset{i}{\hookrightarrow}_{\mathcal{R},\mu}$) if all t with $s|_p \triangleright_\mu t$ are in normal form w.r.t. $(\mathcal{R}, \mu)$. A term s is in *normal form* w.r.t. $(\mathcal{R}, \mu)$ if there is no term t with $s \hookrightarrow_{\mathcal{R},\mu} t$. A CS-TRS $(\mathcal{R}, \mu)$ is *terminating* if $\hookrightarrow_{\mathcal{R},\mu}$ is well founded and *innermost terminating* if $\overset{i}{\hookrightarrow}_{\mathcal{R},\mu}$ is well founded.

Let $\mathcal{D} = \{root(\ell) \mid \ell \to r \in \mathcal{R}\}$ be the set of *defined symbols*. For every $f \in \mathcal{D}$, let $f^\sharp$ be a fresh *tuple symbol* of same arity, where we often write "F" instead of "$f^\sharp$". For $t = f(t_1, \ldots, t_n)$ with $f \in \mathcal{D}$, let $t^\sharp = f^\sharp(t_1, \ldots, t_n)$.

Definition 2 (CS-DPs [1]). *Let $(\mathcal{R}, \mu)$ be a CS-TRS. If $\ell \to r \in \mathcal{R}$, $r \trianglerighteq_\mu t$, and* $\mathrm{root}(t) \in \mathcal{D}$*, then $\ell^\sharp \to t^\sharp$ is an* ordinary dependency pair.[4] *If $\ell \to r \in \mathcal{R}$, $r \trianglerighteq_\mu x$ for a variable x, and $\ell \not\trianglerighteq_\mu x$, then $\ell^\sharp \to x$ is a* collapsing DP. *Let $\mathsf{DP}_o(\mathcal{R}, \mu)$ and $\mathsf{DP}_c(\mathcal{R}, \mu)$ be the sets of all ordinary resp. all collapsing DPs.*

Example 3. For the TRS of Ex. 1, we obtain the following CS-DPs.

$$\mathsf{GT}(\mathsf{s}(x), \mathsf{s}(y)) \to \mathsf{GT}(x, y) \quad (2)$$
$$\mathsf{IF}(\mathsf{true}, x, y) \to x \quad (3)$$
$$\mathsf{IF}(\mathsf{false}, x, y) \to y \quad (4)$$
$$\mathsf{M}(x, y) \to \mathsf{IF}(\mathsf{gt}(y, \mathsf{0}), \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y)), x) \quad (5)$$
$$\mathsf{M}(x, y) \to \mathsf{GT}(y, \mathsf{0}) \quad (6)$$
$$\mathsf{D}(\mathsf{s}(x), \mathsf{s}(y)) \to \mathsf{D}(\mathsf{minus}(x, y), \mathsf{s}(y)) \quad (7)$$
$$\mathsf{D}(\mathsf{s}(x), \mathsf{s}(y)) \to \mathsf{M}(x, y) \quad (8)$$

To prove termination, one has to show that there is no infinite *chain* of DPs. For ordinary rewriting, a sequence $s_1 \to t_1, s_2 \to t_2, \ldots$ of DPs is a *chain* if there is a substitution σ such that $t_i\sigma$ reduces to $s_{i+1}\sigma$.[5] If all $t_i\sigma$ are terminating, then the chain is *minimal* [14, 17, 22]. But due to the collapsing DPs, the notion of "chains" has to be adapted when it is used with CS-DPs [1]. If $s_i \to t_i$ is a collapsing DP (i.e., if $t_i \in \mathcal{V}$), then instead of $t_i\sigma \hookrightarrow^*_{\mathcal{R},\mu} s_{i+1}\sigma$ (and termination of $t_i\sigma$ for minimality), one requires that there is a term w_i with $t_i\sigma \trianglerighteq_\mu w_i$ and $w_i^\sharp \hookrightarrow^*_{\mathcal{R},\mu} s_{i+1}\sigma$. For minimal chains, $w_i^\sharp$ must be terminating.

Example 4. Ex. 1 has the chain (5), (3), (5) as $\mathsf{IF}(\mathsf{gt}(\mathsf{s}(y), \mathsf{0}), \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(\mathsf{s}(y))), x) \hookrightarrow^*_{\mathcal{R},\mu} \mathsf{IF}(\mathsf{true}, \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(\mathsf{s}(y))), x) \hookrightarrow_{(3),\mu} \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(\mathsf{s}(y)))$ *and* $(\mathsf{minus}(\mathsf{p}(x), \mathsf{p}(\mathsf{s}(y))))^\sharp = \mathsf{M}(\mathsf{p}(x), \mathsf{p}(\mathsf{s}(y)))$ *is an instance of the left-hand side of (5).*

A CS-TRS is terminating iff there is no infinite chain [1]. As in the non-CS case, the above notion of chains can also be adapted to *innermost* rewriting. Then a CS-TRS is innermost terminating iff there is no infinite innermost chain [4].

Due to the collapsing CS-DPs (and the corresponding definition of "chains"), it is not easy to extend existing techniques for proving absence of infinite chains to CS-DPs. Therefore, we now introduce a new improved definition of CS-DPs.

3 Non-Collapsing CS-Dependency Pairs

Ordinary DPs only consider active subterms of right-hand sides. So Rule (1) of Ex. 1 only leads to the DP (5), but not to $\mathsf{M}(x, y) \to \mathsf{M}(\mathsf{p}(x), \mathsf{p}(y))$. However, the inactive subterm $\mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y))$ of the right-hand side of (1) may become active again when applying the rule $\mathsf{if}(\mathsf{true}, x, y) \to x$. Therefore, Def. 2 creates a collapsing DP like (3) whenever a rule $\ell \to r$ has a *migrating variable* x with $r \trianglerighteq_\mu x$, but $\ell \not\trianglerighteq_\mu x$. Indeed, when instantiating the *collapse-variable* x in (3) with an instance of the "hidden term" $\mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y))$, one obtains a chain which simulates the rewrite sequence from $\mathsf{minus}(t_1, t_2)$ over $\mathsf{if}(..., \mathsf{minus}(\mathsf{p}(t_1), \mathsf{p}(t_2)), ...)$

[4] A refinement is to eliminate DPs where $\ell \rhd_\mu t$, cf. [1, 9].

[5] We always assume that different occurrences of DPs are variable-disjoint and consider substitutions whose domains may be infinite.

to $\mathsf{minus}(\mathsf{p}(t_1), \mathsf{p}(t_2))$, cf. Ex. 4. Our main observation is that collapsing DPs are only needed for certain instantiations of the variables. One might be tempted to allow only instantiations of collapse-variables by *hidden terms*.[6]

Definition 5 (Hidden Term). *Let $(\mathcal{R}, \mu)$ be a CS-TRS. We say that t is a* hidden term *if $root(t) \in \mathcal{D}$ and if there exists a rule $\ell \to r \in \mathcal{R}$ with $r \rhd_\mu t$.*

In Ex. 1, the only hidden term is $\mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y))$. But unfortunately, only allowing instantiations of collapse-variables with hidden terms would be unsound.

Example 6. Consider $\mu(\mathsf{g}) = \{1\}$, $\mu(\mathsf{a}) = \mu(\mathsf{b}) = \mu(\mathsf{f}) = \mu(\mathsf{h}) = \varnothing$ and the rules

$$\begin{array}{rcl} \mathsf{a} \to \mathsf{f}(\mathsf{g}(\mathsf{b})) & (9) & \mathsf{h}(x) \to x \\ \mathsf{f}(x) \to \mathsf{h}(x) & & \mathsf{b} \to \mathsf{a} \end{array}$$

The CS-TRS has the following infinite rewrite sequence:

$$\mathsf{a} \hookrightarrow_{\mathcal{R},\mu} \mathsf{f}(\mathsf{g}(\mathsf{b})) \hookrightarrow_{\mathcal{R},\mu} \underline{\mathsf{h}(\mathsf{g}(\mathsf{b})) \hookrightarrow_{\mathcal{R},\mu} \mathsf{g}(\mathsf{b})} \hookrightarrow_{\mathcal{R},\mu} \mathsf{g}(\mathsf{a}) \hookrightarrow_{\mathcal{R},\mu} \dots$$

We obtain the following CS-DPs according to Def. 2:

$$\begin{array}{rcll} \mathsf{A} \to \mathsf{F}(\mathsf{g}(\mathsf{b})) & & \mathsf{H}(x) \to x & (10) \\ \mathsf{F}(x) \to \mathsf{H}(x) & & \mathsf{B} \to \mathsf{A} & \end{array}$$

The only hidden term is b*, obtained from Rule (9). There is also an infinite chain that corresponds to the infinite reduction above. However, here the collapse-variable x in the DP (10) must be instantiated by* $\mathsf{g}(\mathsf{b})$ *and not by the hidden term* b*, cf. the underlined part above. So if one replaced (10) by* $\mathsf{H}(\mathsf{b}) \to \mathsf{b}$*, there would be no infinite chain anymore and one would falsely conclude termination.*

The problem in Ex. 6 is that rewrite rules may add additional symbols like g *above* hidden terms. This can happen if a term $\mathsf{g}(t)$ occurs at an inactive position in a right-hand side and if an instantiation of t could possibly reduce to a term containing a hidden term (i.e., if t has a defined symbol or a variable at an active position). Then we call $\mathsf{g}(\Box)$ a *hiding context*, since it can "hide" a hidden term. Moreover, the composition of hiding contexts is again a hiding context.

Definition 7 (Hiding Context). *Let $(\mathcal{R}, \mu)$ be a CS-TRS. The function symbol f* hides position i *if there is a rule $\ell \to r \in \mathcal{R}$ with $r \rhd_\mu f(r_1, \dots, r_i, \dots, r_n)$, $i \in \mu(f)$, and r_i contains a defined symbol or a variable at an active position. A context C is* hiding *iff $C = \Box$ or C has the form $f(t_1, \dots, t_{i-1}, C', t_{i+1}, \dots, t_n)$ where f hides position i and C' is a hiding context.*

Example 8. In Ex. 6, g *hides position 1 due to Rule (9). So the hiding contexts are* $\Box, \mathsf{g}(\Box), \mathsf{g}(\mathsf{g}(\Box)), \dots$ *In the TRS of Ex. 1,* minus *hides both positions 1 and 2 and* p *hides position 1 due to Rule (1). So the hiding contexts are* $\Box, \mathsf{p}(\Box), \mathsf{minus}(\Box, \Box), \mathsf{p}(\mathsf{p}(\Box)), \mathsf{minus}(\Box, \mathsf{p}(\Box)), \dots$

To remove collapsing DPs $s \to x$, we now restrict ourselves to instantiations of x with terms of the form $C[t]$ where C is a hiding context and t is a hidden term. So in Ex. 6, the variable x in the DP (10) should only be instantiated by

[6] A similar notion of *hidden symbols* was presented in [2, 4], but there one only used these symbols to improve one special termination technique (the *dependency graph*).

b, $\mathsf{g}(\mathsf{b})$, $\mathsf{g}(\mathsf{g}(\mathsf{b}))$, etc. To represent these infinitely many instantiations in a finite way, we replace $s \to x$ by new *unhiding* DPs (which "unhide" hidden terms).

Definition 9 (Improved CS-DPs). *For a CS-TRS $(\mathcal{R}, \mu)$, if $\mathsf{DP}_c(\mathcal{R}, \mu) \neq \varnothing$, we introduce a fresh*[7] unhiding tuple symbol U *and the following* unhiding DPs*:*

- $s \to \mathsf{U}(x)$ *for every* $s \to x \in \mathsf{DP}_c(\mathcal{R}, \mu)$,
- $\mathsf{U}(f(x_1, \ldots, x_i, \ldots, x_n)) \to \mathsf{U}(x_i)$ *for every function symbol f of any arity n and every $1 \leq i \leq n$ where f hides position i, and*
- $\mathsf{U}(t) \to t^\sharp$ *for every hidden term t.*

Let $\mathsf{DP}_u(\mathcal{R}, \mu)$ be the set of all unhiding DPs (where $\mathsf{DP}_u(\mathcal{R}, \mu) = \varnothing$, if $\mathsf{DP}_c(\mathcal{R}, \mu) = \varnothing$). Then the set of improved CS-DPs *is $\mathsf{DP}(\mathcal{R}, \mu) = \mathsf{DP}_o(\mathcal{R}, \mu) \cup \mathsf{DP}_u(\mathcal{R}, \mu)$.*

Example 10. In Ex. 6, instead of (10) we get the unhiding DPs

$$\mathsf{H}(x) \to \mathsf{U}(x), \qquad \mathsf{U}(\mathsf{g}(x)) \to \mathsf{U}(x), \qquad \mathsf{U}(\mathsf{b}) \to \mathsf{B}.$$

Now there is indeed an infinite chain. In Ex. 1, instead of (3) and (4), we obtain:[8]

$$\mathsf{IF}(\mathsf{true}, x, y) \to \mathsf{U}(x) \quad (11)$$
$$\mathsf{IF}(\mathsf{false}, x, y) \to \mathsf{U}(y) \quad (12)$$
$$\mathsf{U}(\mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y))) \to \mathsf{M}(\mathsf{p}(x), \mathsf{p}(y)) \quad (13)$$
$$\mathsf{U}(\mathsf{p}(x)) \to \mathsf{P}(x) \quad (14)$$
$$\mathsf{U}(\mathsf{p}(x)) \to \mathsf{U}(x) \quad (15)$$
$$\mathsf{U}(\mathsf{minus}(x, y)) \to \mathsf{U}(x) \quad (16)$$
$$\mathsf{U}(\mathsf{minus}(x, y)) \to \mathsf{U}(y) \quad (17)$$

Clearly, the improved CS-DPs are never collapsing. Thus, now the definition of (minimal)[9] *chains* is completely analogous to the one for ordinary rewriting.

Definition 11 (Chain). *Let $\mathcal{P}$ and $\mathcal{R}$ be TRSs and let μ be a replacement map. We extend μ to tuple symbols by defining $\mu(f^\sharp) = \mu(f)$ for all $f \in \mathcal{D}$ and $\mu(\mathsf{U}) = \varnothing$.*[10] *A sequence of pairs $s_1 \to t_1, s_2 \to t_2, \ldots$ from $\mathcal{P}$ is a $(\mathcal{P}, \mathcal{R}, \mu)$-*chain *iff there is a substitution σ with $t_i\sigma \hookrightarrow^*_{\mathcal{R},\mu} s_{i+1}\sigma$ and $t_i\sigma$ is terminating w.r.t. $(\mathcal{R}, \mu)$ for all i. It is an* innermost $(\mathcal{P}, \mathcal{R}, \mu)$-chain *iff $t_i\sigma \overset{\mathsf{i}}{\hookrightarrow}^*_{\mathcal{R},\mu} s_{i+1}\sigma$, $s_i\sigma$ is in normal form, and $t_i\sigma$ is innermost terminating w.r.t. $(\mathcal{R}, \mu)$ for all i.*

Our main theorem shows that improved CS-DPs are still sound and complete.

Theorem 12 (Soundness and Completeness of Improved CS-DPs). *A CS-TRS $(\mathcal{R}, \mu)$ is terminating iff there is no infinite $(\mathsf{DP}(\mathcal{R}, \mu), \mathcal{R}, \mu)$-chain and innermost terminating iff there is no infinite innermost $(\mathsf{DP}(\mathcal{R}, \mu), \mathcal{R}, \mu)$-chain.*

Proof. We only prove the theorem for "full" termination. The proof for innermost termination is very similar and can be found in [5].

Soundness

$\mathcal{M}_{\infty,\mu}$ contains all *minimal non-terminating terms*: $t \in \mathcal{M}_{\infty,\mu}$ iff t is non-termi-

[7] Alternatively, one could also use different U-symbols for different collapsing DPs.

[8] We omitted the DP $\mathsf{U}(\mathsf{p}(y)) \to \mathsf{P}(y)$ that is "identical" to (14).

[9] Since we only regard *minimal* chains in the following, we included the "minimality requirement" in Def. 11, i.e., we require that all $t_i\sigma$ are (innermost) terminating. As in the DP framework for ordinary rewriting, this restriction to minimal chains is needed for several DP processors (e.g., for the reduction pair processor of Thm. 21).

[10] We define $\mu(\mathsf{U}) = \varnothing$, since the purpose of U is only to remove context around hidden terms. But during this removal, U's argument should not be evaluated.

nating and every r with $t \rhd_\mu r$ terminates. A term u has the *hiding property* iff

- $u \in \mathcal{M}_{\infty,\mu}$ and
- whenever $u \rhd_\mu\!\!\!\!\!/\; s \trianglerighteq_\mu t'$ for some terms s and t' with $t' \in \mathcal{M}_{\infty,\mu}$, then t' is an instance of a hidden term and $s = C[t']$ for some hiding context C.

We first prove the following claim:

$$\text{Let } u \text{ be a term with the hiding property and let } u \hookrightarrow_{\mathcal{R},\mu} v \trianglerighteq_\mu w \text{ with } w \in \mathcal{M}_{\infty,\mu}. \text{ Then } w \text{ also has the hiding property.} \tag{18}$$

Let $w \rhd_\mu\!\!\!\!\!/\; s \trianglerighteq_\mu t'$ for some terms s and t' with $t' \in \mathcal{M}_{\infty,\mu}$. Clearly, this also implies $v \rhd_\mu\!\!\!\!\!/\; s$. If already $u \rhd s$, then we must have $u \rhd_\mu\!\!\!\!\!/\; s$ due to the minimality of u. Thus, t' is an instance of a hidden term and $s = C[t']$ for a hiding context C, since u has the hiding property. Otherwise, $u \not\rhd s$. There must be a rule $\ell \to r \in \mathcal{R}$, an active context D (i.e., a context where the hole is at an active position), and a substitution δ such that $u = D[\delta(\ell)]$ and $v = D[\delta(r)]$. Clearly, $u \not\rhd s$ implies $\delta(\ell) \not\rhd s$ and $D \not\rhd s$. Hence, $v \rhd_\mu\!\!\!\!\!/\; s$ means $\delta(r) \rhd_\mu\!\!\!\!\!/\; s$. (The root of s cannot be above $\Box$ in D since those positions would be active.) Note that s cannot be at or below a variable position of r, because this would imply $\delta(\ell) \rhd s$. Thus, s is an instance of a non-variable subterm of r that is at an inactive position. So there is a $r' \notin \mathcal{V}$ with $r \rhd_\mu\!\!\!\!\!/\; r'$ and $s = \delta(r')$. Recall that $s \trianglerighteq_\mu t'$, i.e., there is a $p \in \mathcal{P}os^\mu(s)$ with $s|_p = t'$. If p is a non-variable position of r', then $\delta(r'|_p) = t'$ and $r'|_p$ is a subterm with defined root at an active position (since $t' \in \mathcal{M}_{\infty,\mu}$ implies $\mathrm{root}(t') \in \mathcal{D}$). Hence, $r'|_p$ is a hidden term and thus, t' is an instance of a hidden term. Moreover, any instance of the context $C' = r'[\Box]_p$ is hiding. So if we define C to be $\delta(C')$, then $s = \delta(r') = \delta(r')[t']_p = \delta(C')[t'] = C[t']$ for the hiding context C. On the contrary, if p is not a non-variable position of r', then $p = p_1\, p_2$ where $r'|_{p_1}$ is a variable x. Now t' is an active subterm of $\delta(x)$ (more precisely, $\delta(x)|_{p_2} = t'$). Since x also occurs in ℓ, we have $\delta(\ell) \rhd \delta(x)$ and thus $u \rhd \delta(x)$. Due to the minimality of u this implies $u \rhd_\mu\!\!\!\!\!/\; \delta(x)$. Since $u \rhd_\mu\!\!\!\!\!/\; \delta(x) \trianglerighteq_\mu t'$, the hiding property of u implies that t' is an instance of a hidden term and that $\delta(x) = \overline{C}[t']$ for a hiding context $\overline{C}$. Note that since $r'|_{p_1}$ is a variable, the context C' around this variable is also hiding (i.e., $C' = r'[\Box]_{p_1}$). Thus, the context $C = \delta(C')[\overline{C}]$ is hiding as well and $s = \delta(r') = \delta(r')[\delta(x)[t']_{p_2}]_{p_1} = \delta(C')[\overline{C}[t']] = C[t']$.

Proof of Thm. 12 using Claim (18)

If $\mathcal{R}$ is not terminating, then there is a $t \in \mathcal{M}_{\infty,\mu}$ that is minimal w.r.t. $\trianglerighteq$. So there are t, t_i, s_i, t'_{i+1} such that

$$t \xhookrightarrow{>\varepsilon}{}^*_{\mathcal{R},\mu}\ t_1 \xrightarrow{\varepsilon}_{\mathcal{R}} s_1 \trianglerighteq_\mu t'_2 \xhookrightarrow{>\varepsilon}{}^*_{\mathcal{R},\mu}\ t_2 \xrightarrow{\varepsilon}_{\mathcal{R}} s_2 \trianglerighteq_\mu t'_3 \xhookrightarrow{>\varepsilon}{}^*_{\mathcal{R},\mu}\ t_3 \ldots \tag{19}$$

where $t_i, t'_i \in \mathcal{M}_{\infty,\mu}$ and all proper subterms of t (also at *inactive* positions) terminate. Here, "ε" (resp. "$> \varepsilon$") denotes reductions at (resp. strictly below) the root.

Note that (18) implies that all t_i have the hiding property. To see this, we use induction on i. Since t trivially has the hiding property (as it has no non-terminating proper subterms) and all terms in the reduction $t \xhookrightarrow{>\varepsilon}{}^*_{\mathcal{R},\mu}\ t_1$ are

from $\mathcal{M}_{\infty,\mu}$ (as both $t, t_1 \in \mathcal{M}_{\infty,\mu}$), we conclude that t_1 also has the hiding property by applying (18) repeatedly. In the induction step, if t_{i-1} has the hiding property, then one application of (18) shows that t'_i also has the hiding property. By applying (18) repeatedly, one then also shows that t_i has the hiding property.

Now we show that $t_i^\sharp \to^+_{\mathsf{DP}(\mathcal{R},\mu)} {t'_{i+1}}^\sharp$ and that all terms in the reduction $t_i^\sharp \to^+_{\mathsf{DP}(\mathcal{R},\mu)} {t'_{i+1}}^\sharp$ terminate w.r.t. $(\mathcal{R},\mu)$. As ${t'_{i+1}}^\sharp \xhookrightarrow{>\varepsilon}{}^*_{\mathcal{R},\mu} t_{i+1}^\sharp$, we get an infinite $(\mathsf{DP}(\mathcal{R},\mu),\mathcal{R},\mu)$-chain.

From (19) we know that there are $\ell_i \to r_i \in \mathcal{R}$ and $p_i \in \mathcal{P}os^\mu(s_i)$ with $t_i = \ell_i\sigma$, $s_i = r_i\sigma$, and $s_i|_{p_i} = r_i\sigma|_{p_i} = t'_{i+1}$ for all i. First let $p_i \in \mathcal{P}os(r_i)$ with $r_i|_{p_i} \notin \mathcal{V}$. Then $\ell_i^\sharp \to (r_i|_{p_i})^\sharp \in \mathsf{DP}_o(\mathcal{R},\mu)$ and $t_i^\sharp = \ell_i^\sharp\sigma \to_{\mathsf{DP}_o(\mathcal{R},\mu)} (r_i|_{p_i})^\sharp\sigma = {t'_{i+1}}^\sharp$. Moreover, as $t_i, t'_{i+1} \in \mathcal{M}_{\infty,\mu}$, the terms $t_i^\sharp$ and ${t'_{i+1}}^\sharp$ are terminating.

Now let p_i be at or below the position of a variable x_i in r_i. By minimality of t_i, x_i only occurs at inactive positions of ℓ_i. Thus, $\ell_i^\sharp \to \mathsf{U}(x_i) \in \mathsf{DP}_u(\mathcal{R},\mu)$ and $r_i = C_i[x_i]$ where C_i is an active context. Recall that $t_i = \ell_i\sigma$ has the hiding property and that $t_i \rhd_\mu \sigma(x_i) \trianglerighteq_\mu t'_{i+1}$. Thus, we have $\sigma(x_i) = C[t'_{i+1}]$ for a hiding context C and moreover, t'_{i+1} is an instance of a hidden term. Hence we obtain:

$$
\begin{array}{lll}
t_i^\sharp = \sigma(\ell_i^\sharp) & & \\
\to_{\mathsf{DP}_u(\mathcal{R},\mu)} \mathsf{U}(\sigma(x_i)) & \text{since } \ell_i^\sharp \to \mathsf{U}(x_i) \in \mathsf{DP}_u(\mathcal{R},\mu) \\
= \mathsf{U}(C[t'_{i+1}]) & \text{for a hiding context } C \\
\to^*_{\mathsf{DP}_u(\mathcal{R},\mu)} \mathsf{U}(t'_{i+1}) & \text{since } \mathsf{U}(C[x]) \to^*_{\mathsf{DP}_u(\mathcal{R},\mu)} \mathsf{U}(x) \text{ for any hiding context } C \\
\to_{\mathsf{DP}_u(\mathcal{R},\mu)} {t'_{i+1}}^\sharp & \text{since } t'_{i+1} \text{ is an instance of a hidden term and} \\
 & \mathsf{U}(t) \to_{\mathsf{DP}_u(\mathcal{R},\mu)} t^\sharp \text{ for any instance } t \text{ of a hidden term}
\end{array}
$$

All terms in the reduction above are terminating. The reason is that again $t_i, t'_{i+1} \in \mathcal{M}_{\infty,\mu}$ implies that $t_i^\sharp$ and ${t'_{i+1}}^\sharp$ are terminating. Moreover, all terms $\mathsf{U}(\ldots)$ are normal forms since $\mu(\mathsf{U}) = \varnothing$ and since U does not occur in $\mathcal{R}$.

Completeness

Let there be an infinite chain $v_1 \to w_1, v_2 \to w_2, \ldots$ of improved CS-DPs. First, let the chain have an infinite tail consisting only of DPs of the form $\mathsf{U}(f(x_1, \ldots, x_i, \ldots, x_n)) \to \mathsf{U}(x_i)$. Since $\mu(\mathsf{U}) = \varnothing$, there are terms t_i with $\mathsf{U}(t_1) \xrightarrow{\varepsilon}_{\mathsf{DP}(R,\mu)} \mathsf{U}(t_2) \xrightarrow{\varepsilon}_{\mathsf{DP}(R,\mu)} \ldots$ Hence, $t_1 \rhd_\mu t_2 \rhd_\mu \ldots$ which contradicts the well-foundedness of $\rhd_\mu$.

Now we regard the remaining case. Here the chain has infinitely many DPs $v \to w$ with $v = \ell^\sharp$ for a rule $\ell \to r \in \mathcal{R}$. Let $v_i \to w_i$ be such a DP and let $v_j \to w_j$ with $j > i$ be the *next* such DP in the chain. Let σ be the substitution used for the chain. We show that then $v_i^\flat\sigma \hookrightarrow^*_{\mathcal{R},\mu} C[v_j^\flat\sigma]$ for an active context C. Here, $(f^\sharp(t_1, \ldots, t_n))^\flat = f(t_1, \ldots, t_n)$ for all $f \in \mathcal{D}$. Doing this for all such DPs implies that there is an infinite reduction w.r.t. $(\mathcal{R},\mu)$.

If $v_i \to w_i \in \mathsf{DP}_o(R,\mu)$ then the claim is trivial, because then $j = i+1$ and $v_i^\flat\sigma \hookrightarrow_{\mathcal{R},\mu} C[w_i^\flat\sigma] \hookrightarrow^*_{\mathcal{R},\mu} C[v_{i+1}^\flat\sigma]$ for some active context C.

Otherwise, $v_i \to w_i$ has the form $v_i \to \mathsf{U}(x)$. Then $v_i^\flat\sigma \hookrightarrow_{\mathcal{R},\mu} C_1[\sigma(x)]$ for an active context C_1. Moreover, $\mathsf{U}(\sigma(x))$ reduces to $\mathsf{U}(\delta(t))$ for a hidden term t and a δ by removing hiding contexts. Since hiding contexts are active, $\sigma(x) = C_2[\delta(t)]$ for an active context C_2. Finally, $t^\sharp\delta \xhookrightarrow{>\varepsilon}{}^*_{\mathcal{R},\mu} v_j\sigma$ and thus, $t\delta \xhookrightarrow{>\varepsilon}{}^*_{\mathcal{R},\mu} v_j^\flat\sigma$. By defining $C = C_1[C_2]$, we get $v_i^\flat\sigma \hookrightarrow^+_{\mathcal{R},\mu} C[v_j^\flat\sigma]$. □

4 CS Dependency Pair Framework

By Thm. 12, (innermost) termination of a CS-TRS is equivalent to absence of infinite (innermost) chains. For ordinary rewriting, the *DP framework* is the most recent and powerful collection of methods to prove absence of infinite chains automatically. Due to our new notion of (non-collapsing) CS-DPs, adapting the DP framework to the context-sensitive case now becomes much easier.[11]

In the DP framework, termination techniques operate on *DP problems* instead of TRSs. Def. 13 adapts this notion to context-sensitive rewriting.

Definition 13 (CS-DP Problem and Processor). *A* CS-DP problem *is a tuple $(\mathcal{P}, \mathcal{R}, \mu, e)$, where $\mathcal{P}$ and $\mathcal{R}$ are TRSs, μ is a replacement map, and $e \in \{\mathbf{t}, \mathbf{i}\}$ is a flag that stands for* **t**ermination *or* **i**nnermost termination. *We also call $(\mathcal{P}, \mathcal{R}, \mu)$-chains "$(\mathcal{P}, \mathcal{R}, \mu, \mathbf{t})$-chains" and we call innermost $(\mathcal{P}, \mathcal{R}, \mu)$-chains "$(\mathcal{P}, \mathcal{R}, \mu, \mathbf{i})$-chains". A* CS-DP problem $(\mathcal{P}, \mathcal{R}, \mu, e)$ *is* finite *if there is no infinite $(\mathcal{P}, \mathcal{R}, \mu, e)$-chain.*

A CS-DP processor *is a function Proc that takes a CS-DP problem as input and returns a possibly empty set of CS-DP problems. The processor Proc is* sound *if a CS-DP problem d is finite whenever all problems in Proc(d) are finite.*

For a CS-TRS $(\mathcal{R}, \mu)$, the termination proof starts with the *initial DP problem* $(\mathsf{DP}(\mathcal{R}, \mu), \mathcal{R}, \mu, e)$ where e depends on whether one wants to prove termination or innermost termination. Then sound DP processors are applied repeatedly. If the final processors return empty sets, then (innermost) termination is proved. Since innermost termination is usually easier to show than full termination, one should use $e = \mathbf{i}$ whenever possible. As shown in [12], termination and innermost termination coincide for CS-TRSs $(\mathcal{R}, \mu)$ where $\mathcal{R}$ is *orthogonal* (i.e., left-linear and without critical pairs). So $(DP(\mathcal{R}, \mu), \mathcal{R}, \mu, \mathbf{i})$ would be the initial DP problem for Ex. 1, even when proving full termination. In Sect. 4.1 - 4.3, we recapitulate 3 important DP processors and extend them to context-sensitive rewriting.

4.1 Dependency Graph Processor

The first processor decomposes a DP problem into several sub-problems. To this end, one determines which pairs can follow each other in chains by constructing a *dependency graph*. In contrast to related definitions for collapsing CS-DPs in [1, 4], Def. 14 is analogous to the corresponding definition for non-CS rewriting.

Definition 14 (CS-Dependency Graph). *For a CS-DP problem $(\mathcal{P}, \mathcal{R}, \mu, e)$, the nodes of the* $(\mathcal{P}, \mathcal{R}, \mu, e)$-dependency graph *are the pairs of $\mathcal{P}$, and there is an arc from $v \to w$ to $s \to t$ iff $v \to w, s \to t$ is a $(\mathcal{P}, \mathcal{R}, \mu, e)$-chain.*

Example 15. Fig. 1 shows the dependency graph for Ex. 1, for both $e \in \{\mathbf{t}, \mathbf{i}\}$.[12]

[11] For this reason, we omitted the proofs in this section and refer to [5] for all proofs.

[12] To improve readability, we omitted nodes (6) and (14) from the graph. There are arcs from the nodes (8) and (13) to (6) and from all nodes (11), (12), (15), (16), (17) to (14). But (6) and (14) have no outgoing arcs and thus, they are not on any cycle.

A set $\mathcal{P}' \neq \varnothing$ of DPs is a *cycle* if for every $v \to w, s \to t \in \mathcal{P}'$, there is a non-empty path from $v \to w$ to $s \to t$ traversing only pairs of $\mathcal{P}'$. A cycle $\mathcal{P}'$ is a *strongly connected component* ("SCC") if $\mathcal{P}'$ is not a proper subset of another cycle.

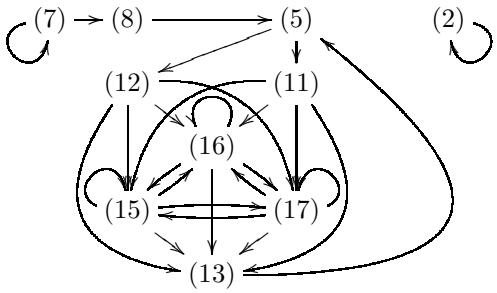

Fig. 1. Dependency graph for Ex. 1

One can prove termination separately for each SCC. Thus, the following processor (whose soundness is obvious and completely analogous to the non-context-sensitive case) modularizes termination proofs.

Theorem 16 (CS-Dependency Graph Processor). *For $d = (\mathcal{P}, \mathcal{R}, \mu, e)$, let $Proc(d) = \{(\mathcal{P}_1, \mathcal{R}, \mu, e), \ldots, (\mathcal{P}_n, \mathcal{R}, \mu, e)\}$, where $\mathcal{P}_1, \ldots, \mathcal{P}_n$ are the SCCs of the $(\mathcal{P}, \mathcal{R}, \mu, e)$-dependency graph. Then Proc is sound.*

Example 17. The graph in Fig. 1 has the three SCCs $\mathcal{P}_1 = \{(2)\}$, $\mathcal{P}_2 = \{(7)\}$, $\mathcal{P}_3 = \{(5), (11)\text{-}(13), (15)\text{-}(17)\}$. Thus, the initial DP problem $(\mathsf{DP}(\mathcal{R}, \mu), \mathcal{R}, \mu, \mathbf{i})$ is transformed into the new problems $(\mathcal{P}_1, \mathcal{R}, \mu, \mathbf{i})$, $(\mathcal{P}_2, \mathcal{R}, \mu, \mathbf{i})$, $(\mathcal{P}_3, \mathcal{R}, \mu, \mathbf{i})$.

As in the non-context-sensitive setting, the CS-dependency graph is not computable and thus, one has to use estimations to over-approximate the graph. For example, [1, 4] adapted the estimation of [6] that was originally developed for ordinary rewriting: $\textsc{Cap}^\mu(t)$ replaces all active subterms of t with defined root symbol by different fresh variables. Multiple occurrences of the same such subterm are also replaced by pairwise different variables. $\textsc{Ren}^\mu(t)$ replaces all active occurrences of variables in t by different fresh variables (i.e., no variable occurs at several active positions in $\textsc{Ren}^\mu(t)$). So $\textsc{Ren}^\mu(\textsc{Cap}^\mu(\mathsf{IF}(\mathsf{gt}(y, \mathsf{0}), \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y)), x))) = \textsc{Ren}^\mu(\mathsf{IF}(z, \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y)), x)) = \mathsf{IF}(z', \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y)), x)$.

To estimate the CS-dependency graph in the case $e = \mathbf{t}$, one draws an arc from $v \to w$ to $s \to t$ whenever $\textsc{Ren}^\mu(\textsc{Cap}^\mu(w))$ and s unify.[13] If $e = \mathbf{i}$, then one can modify $\textsc{Cap}^\mu$ and $\textsc{Ren}^\mu$ by taking into account that instantiated subterms at active positions of the left-hand side must be in normal form, cf. [4]. $\textsc{Cap}^\mu_v(w)$ is like $\textsc{Cap}^\mu(w)$, but the replacement of subterms of w by fresh variables is not done if the subterms also occur at active positions of v. Similarly, $\textsc{Ren}^\mu_v(w)$ is like $\textsc{Ren}^\mu(w)$, but the renaming of variables in w is not done if the variables also occur active in v. Now we draw an arc from $v \to w$ to $s \to t$ whenever $\textsc{Ren}^\mu_v(\textsc{Cap}^\mu_v(w))$ and s unify by an mgu θ where $v\theta$ and $s\theta$ are in normal form.[14]

It turns out that for the TRS of Ex. 1, the resulting estimated dependency graph is identical to the "real" graph in Fig. 1.

[13] Here (and also later in the instantiation processor of Sect. 4.3), we always assume that $v \to w$ and $s \to t$ are renamed apart to be variable-disjoint.

[14] These estimations can be improved further by adapting existing refinements to the context-sensitive case. However, different to the non-context-sensitive case, for $e = \mathbf{i}$ it is not sufficient to check only for unification of $\textsc{Cap}^\mu_v(w)$ and s (i.e., renaming variables with $\textsc{Ren}^\mu_v$ is also needed). This can be seen from the non-innermost terminating CS-TRS $(\mathcal{R}, \mu)$ from [4, Ex. 8] with $\mathcal{R} = \{\mathsf{f}(\mathsf{s}(x), x) \to \mathsf{f}(x, x), \mathsf{a} \to \mathsf{s}(\mathsf{a})\}$

4.2 Reduction Pair Processor

There are several processors to simplify DP problems by applying suitable *well-founded orders* (e.g., the *reduction pair processor* [17, 21], the *subterm criterion processor* [22], etc.). Due to the absence of collapsing DPs, most of these processors are now straightforward to adapt to the context-sensitive setting. In the following, we present the reduction pair processor with *usable rules*, because it is the only processor whose adaption is more challenging. (The adaption is similar to the one in [4, 20] for the CS-DPs of Def. 2.)

To prove that a DP problem is finite, the reduction pair processor generates constraints which should be satisfied by a *μ-reduction pair* $(\succsim, \succ)$ [1]. Here, $\succsim$ is a stable μ-monotonic quasi-order, $\succ$ is a stable well-founded order, and $\succsim$ and $\succ$ are compatible (i.e., $\succ \circ \succsim \subseteq \succ$ or $\succsim \circ \succ \subseteq \succ$). Here, *$\mu$-monotonicity* means that $s_i \succsim t_i$ implies $f(s_1, ..., s_i, ..., s_n) \succsim f(s_1, ..., t_i, ..., s_n)$ whenever $i \in \mu(f)$.

For a DP problem $(\mathcal{P}, \mathcal{R}, \mu, e)$, the generated constraints ensure that some rules in $\mathcal{P}$ are strictly decreasing (w.r.t. $\succ$) and all remaining rules in $\mathcal{P}$ and $\mathcal{R}$ are weakly decreasing (w.r.t. $\succsim$). Requiring $\ell \succsim r$ for all $\ell \to r \in \mathcal{R}$ ensures that in a chain $s_1 \to t_1, s_2 \to t_2, ...$ with $t_i\sigma \hookrightarrow^*_{\mathcal{R},\mu} s_{i+1}\sigma$, we have $t_i\sigma \succsim s_{i+1}\sigma$ for all i. Hence, if a reduction pair satisfies the constraints, then one can delete the strictly decreasing pairs from $\mathcal{P}$ as they cannot occur infinitely often in chains.

To improve this idea, it is desirable to require only a weak decrease of *certain* instead of *all* rules. In the non-context-sensitive setting, when proving innermost termination, it is sufficient if just the *usable rules* are weakly decreasing [6]. The same is true when proving full termination, provided that $\succsim$ is *$\mathcal{C}_\varepsilon$-compatible*, i.e., $\mathsf{c}(x, y) \succsim x$ and $\mathsf{c}(x, y) \succsim y$ holds for a fresh function symbol c [17, 22].

For a term containing a symbol f, all f-rules are *usable*. Moreover, if the f-rules are usable and f *depends on* h (denoted $f \blacktriangleright_{\mathcal{R}} h$) then the h-rules are usable as well. Here, $f \blacktriangleright_{\mathcal{R}} h$ if $f = h$ or if there is a symbol g with $g \blacktriangleright_{\mathcal{R}} h$ and g occurs in the right-hand side of an f-rule. The usable rules of a DP problem are defined to be the usable rules of the right-hand sides of the DPs.

As in [4, 20], Def. 18 adapts[15] the concept of usable rules to the CS setting, resulting in $\mathcal{U}^{\blacktriangleright}(\mathcal{P}, \mathcal{R}, \mu)$. But as shown in [20], for CS rewriting it is also helpful to consider an alternative definition of "dependence" $\rhd_{\mathcal{R},\mu}$ where f also depends on symbols from *left-hand sides* of f-rules. Let $\mathcal{F}^\mu(t)$ (resp. $\mathcal{F}^{\not\mu}(t)$) contain all function symbols occurring at active (resp. inactive) positions of a term t.

Definition 18 (CS-Usable Rules). *Let $Rls(f) = \{\ell \to r \in \mathcal{R} \mid \mathrm{root}(\ell) = f\}$. For any symbols f, h and CS-TRS $(\mathcal{R}, \mu)$, let $f \blacktriangleright_{\mathcal{R},\mu} h$ if $f = h$ or if there is a symbol g with $g \blacktriangleright_{\mathcal{R},\mu} h$ and a rule $\ell \to r \in Rls(f)$ with $g \in \mathcal{F}^\mu(r)$. Let $f \rhd_{\mathcal{R},\mu} h$ if $f = h$ or if there is a symbol g with $g \rhd_{\mathcal{R},\mu} h$ and a rule $\ell \to r \in Rls(f)$ with*

and $\mu(\mathsf{f}) = \{1\}$, $\mu(\mathsf{s}) = \varnothing$. Clearly, $\mathrm{CAP}^\mu_{\mathsf{F}(\mathsf{s}(x),x)}(\mathsf{F}(x, x)) = \mathsf{F}(x, x)$ does not unify with $\mathsf{F}(\mathsf{s}(y), y)$. In contrast, $\mathrm{REN}^\mu_{\mathsf{F}(\mathsf{s}(x),x)}(\mathrm{CAP}^\mu_{\mathsf{F}(\mathsf{s}(x),x)}(\mathsf{F}(x, x))) = \mathsf{F}(x', x)$ unifies with $\mathsf{F}(\mathsf{s}(y), y)$. Thus, without using $\mathrm{REN}^\mu_{\mathsf{F}(\mathsf{s}(x),x)}$ one would conclude that the dependency graph has no cycle and wrongly prove (innermost) termination.

[15] The adaptions can also be extended to refined definitions of usable rules [15, 17].

$g \in \mathcal{F}^{\not\mu}(\ell) \cup \mathcal{F}(r)$. *We define two forms of* usable rules*:*

$$\mathcal{U}^{\blacktriangleright}(\mathcal{P},\mathcal{R},\mu) = \bigcup_{s\to t\in\mathcal{P}, f\in\mathcal{F}^{\mu}(t), f \blacktriangleright_{\mathcal{R},\mu} g} Rls(g)$$
$$\mathcal{U}^{\triangleright}(\mathcal{P},\mathcal{R},\mu) = \bigcup_{s\to t\in\mathcal{P}, f\in\mathcal{F}^{\not\mu}(s)\cup\mathcal{F}(t), f \triangleright_{\mathcal{R},\mu} g} Rls(g) \cup \bigcup_{\ell\to r\in\mathcal{R}, f\in\mathcal{F}^{\not\mu}(r), f \triangleright_{\mathcal{R},\mu} g} Rls(g)$$

Example 19. We continue Ex. 17. $\mathcal{U}^{\blacktriangleright}(\mathcal{P}_1,\mathcal{R},\mu) = \varnothing$ *for* $\mathcal{P}_1 = \{(2)\}$*, since there is no defined symbol at an active position in the right-hand side* $\mathsf{GT}(x,y)$ *of (2). For* $\mathcal{P}_2 = \{(7)\}$*,* $\mathcal{U}^{\blacktriangleright}(\mathcal{P}_2,\mathcal{R},\mu)$ *are the* minus-*,* if-*, and* gt*-rules, since* minus *occurs at an active position in* $\mathsf{D}(\mathsf{minus}(x,y),\mathsf{s}(y))$ *and* minus *depends on* if *and* gt*. For* $\mathcal{P}_3 = \{(5),(11)\text{-}(13),(15)\text{-}(17)\}$*,* $\mathcal{U}^{\blacktriangleright}(\mathcal{P}_3,\mathcal{R},\mu)$ *are the* gt*- and* p*-rules, as* gt *and* p *are the only defined symbols at active positions of right-hand sides in* $\mathcal{P}_3$*.*

In contrast, all $\mathcal{U}^{\triangleright}(\mathcal{P}_i,\mathcal{R},\mu)$ *contain all rules except the* div*-rules, as* minus *and* p *are root symbols of hidden terms and* minus *depends on* if *and* gt*.*

As shown in [4, 20], the direct adaption of the usable rules to the context-sensitive case (i.e., $\mathcal{U}^{\blacktriangleright}(\mathcal{P},\mathcal{R},\mu)$) can only be used for *conservative* CS-TRSs (if $e = \mathbf{i}$) resp. for *strongly conservative* CS-TRSs (if $e = \mathbf{t}$).[16] Let $\mathcal{V}^{\mu}(t)$ (resp. $\mathcal{V}^{\not\mu}(t)$) be all variables occurring at active (resp. inactive) positions of a term t.

Definition 20 (Conservative and Strongly Conservative). *A CS-TRS* $(\mathcal{R},\mu)$ *is* conservative *iff* $\mathcal{V}^{\mu}(r) \subseteq \mathcal{V}^{\mu}(\ell)$ *for all rules* $\ell \to r \in \mathcal{R}$*. It is* strongly conservative *iff it is conservative and moreover,* $\mathcal{V}^{\mu}(\ell) \cap \mathcal{V}^{\not\mu}(\ell) = \varnothing$ *and* $\mathcal{V}^{\mu}(r) \cap \mathcal{V}^{\not\mu}(r) = \varnothing$ *for all rules* $\ell \to r \in \mathcal{R}$*.*

Now we can define the reduction pair processor.

Theorem 21 (CS-Reduction Pair Processor). *Let* $(\succsim,\succ)$ *be a* μ*-reduction pair. For a CS-DP Problem* $d = (\mathcal{P},\mathcal{R},\mu,e)$*, the result of Proc*$(d)$ *is*

- $\{(\mathcal{P} \setminus \succ, \mathcal{R}, \mu, e)\}$*, if* $\mathcal{P} \subseteq (\succ \cup \succsim)$ *and at least one of the following holds:*

(i) $\mathcal{U}^{\blacktriangleright}(\mathcal{P},\mathcal{R},\mu) \subseteq \succsim$*,* $\mathcal{P} \cup \mathcal{U}^{\blacktriangleright}(\mathcal{P},\mathcal{R},\mu)$ *is strongly conservative,* $\succsim$ *is* $\mathcal{C}_\varepsilon$*-compatible*
(ii) $\mathcal{U}^{\blacktriangleright}(\mathcal{P},\mathcal{R},\mu) \subseteq \succsim$*,* $\mathcal{P} \cup \mathcal{U}^{\blacktriangleright}(\mathcal{P},\mathcal{R},\mu)$ *is conservative,* $e = \mathbf{i}$
(iii) $\mathcal{U}^{\triangleright}(\mathcal{P},\mathcal{R},\mu) \subseteq \succsim$*,* $\succsim$ *is* $\mathcal{C}_\varepsilon$*-compatible*
(iv) $\mathcal{R} \subseteq \succsim$

- $\{d\}$*, otherwise.*

Then Proc is sound.

Example 22. As $\mathcal{U}^{\blacktriangleright}(\mathcal{P}_1,\mathcal{R},\mu) = \varnothing$ *and* $\mathcal{P}_1 = \{(2)\}$ *is even strongly conservative, by Thm. 21 (i) or (ii) we only have to orient (2), which already works with the embedding order. So* $(\mathcal{P}_1,\mathcal{R},\mu,\mathbf{i})$ *is transformed to the empty set of DP problems.*

[16] The corresponding counterexamples in [4, 20] show that these restrictions are still necessary for our new notion of CS-DPs. In cases where one cannot use $\mathcal{U}^{\blacktriangleright}$, one can also attempt a termination proof where one drops the replacement map, i.e., where one regards the ordinary TRS $\mathcal{R}$ instead of the CS-TRS $(\mathcal{R},\mu)$. This may be helpful, since $\mathcal{U}^{\triangleright}$ is not necessarily a subset of the non-context-sensitive usable rules, as a function symbol f also $\triangleright$-depends on symbols from *left*-hand sides of f-rules.

For $\mathcal{P}_2 = \{(7)\}, \mathcal{U}^{\blacktriangleright}(\mathcal{P}_2, \mathcal{R}, \mu)$ contains the if*-rules which are not conservative. Hence, we use Thm. 21 (iii) with a reduction pair based on the following max-polynomial interpretation [10]:* $[\mathsf{D}(x,y)] = [\mathsf{minus}(x,y)] = [\mathsf{p}(x)] = x$, $[\mathsf{s}(x)] = x+1$, $[\mathsf{if}(x,y,z)] = \max(y,z)$, $[\mathsf{0}] = [\mathsf{gt}(x,y)] = [\mathsf{true}] = [\mathsf{false}] = 0$. *Then the DP (7) is strictly decreasing and all rules from $\mathcal{U}^{\triangleright}(\mathcal{P}_2, \mathcal{R}, \mu)$ are weakly decreasing. Thus, the processor also transforms $(\mathcal{P}_2, \mathcal{R}, \mu, \mathbf{i})$ to the empty set of DP problems.*

Finally, we regard $\mathcal{P}_3 = \{(5), (11)\text{-}(13), (15)\text{-}(17)\}$ where we use Thm. 21 (iii) with the interpretation $[\mathsf{M}(x,y)] = [\mathsf{minus}(x,y)] = x+y+1$, $[\mathsf{IF}(x,y,z)] = [\mathsf{if}(x,y,z)] = \max(y,z)$, $[\mathsf{U}(x)] = [\mathsf{p}(x)] = [\mathsf{s}(x)] = x$, $[\mathsf{0}] = [\mathsf{gt}(x,y)] = [\mathsf{true}] = [\mathsf{false}] = 0$. *Then the DPs (16) and (17) are strictly decreasing, whereas all other DPs from $\mathcal{P}_3$ and all rules from $\mathcal{U}^{\triangleright}(\mathcal{P}_3, \mathcal{R}, \mu)$ are weakly decreasing. So the processor results in the DP problem $(\{(5), (11)\text{-}(13), (15)\}, \mathcal{R}, \mu, \mathbf{i})$.*

Next we apply $[\mathsf{M}(x,y)] = [\mathsf{minus}(x,y)] = x+1$, $[\mathsf{IF}(x,y,z)] = \max(y,z+1)$, $[\mathsf{if}(x,y,z)] = \max(y,z)$, $[\mathsf{U}(x)] = [\mathsf{p}(x)] = [\mathsf{s}(x)] = x$, $[\mathsf{0}] = [\mathsf{gt}(x,y)] = [\mathsf{true}] = [\mathsf{false}] = 0$. *Now (12) is strictly decreasing and all other remaining DPs and usable rules are weakly decreasing. Removing (12) yields $(\{(5), (11), (13), (15)\}, \mathcal{R}, \mu, \mathbf{i})$.*

Thm. 21 (iii) and (iv) are a significant improvement over previous reduction pair processors [1, 2, 4, 20] for the CS-DPs from Def. 2. The reason is that all previous CS-reduction pair processors require that the context-sensitive subterm relation is contained in $\succsim$ (i.e., $\trianglerighteq_\mu \subseteq \succsim$) whenever there are collapsing DPs. This is a very hard requirement which destroys one of the main advantages of the DP method (i.e., the possibility to filter away arbitrary arguments).[17] With our new non-collapsing CS-DPs, this requirement is no longer needed.

Example 23. If one requires $\trianglerighteq_\mu \subseteq \succsim$, then the reduction pair processor would fail for Ex. 1, since then one cannot make the DP (7) strictly decreasing. The reason is that due to $2 \in \mu(\mathsf{minus})$, $\trianglerighteq_\mu \subseteq \succsim$ implies $\mathsf{minus}(x,y) \succsim y$. So one cannot "filter away" the second argument of minus*. But then a strict decrease of DP (7) together with μ-monotonicity of $\succsim$ implies $\mathsf{D}(\mathsf{s}(x), \mathsf{s}(\mathsf{s}(x))) \succ \mathsf{D}(\mathsf{minus}(x, \mathsf{s}(x)), \mathsf{s}(\mathsf{s}(x))) \succsim \mathsf{D}(\mathsf{s}(x), \mathsf{s}(\mathsf{s}(x)))$, in contradiction to the well-foundedness of $\succ$.*

4.3 Transforming Context-Sensitive Dependency Pairs

To increase the power of the DP method, there exist several processors to transform a DP into new pairs (e.g., *narrowing*, *rewriting*, *instantiating*, or *forward instantiating* DPs [17]). We now adapt the *instantiation* processor to the context-sensitive setting. Similar adaptions can also be done for the other processors.[18]

[17] Moreover, previous CS-reduction pair processors also require $f(x_1, \ldots, x_n) \succsim f^\sharp(x_1, \ldots, x_n)$ for all $f \in \mathcal{D}$ or $f(x_1, \ldots, x_n) \succ f^\sharp(x_1, \ldots, x_n)$ for all $f \in \mathcal{D}$. This requirement also destroys an important feature of the DP method, i.e., that tuple symbols $f^\sharp$ can be treated independently from the original corresponding symbols f. This feature often simplifies the search for suitable reduction pairs considerably.

[18] In the papers on CS-DPs up to now, the only existing adaption of such a processor was the straightforward adaption of the *narrowing* processor in the case $e = \mathbf{t}$, cf. [2]. However, this processor would not help for the TRS of Ex. 1.

The idea of this processor is the following. For a DP $s \to t$, we investigate which DPs $v \to w$ can occur before $s \to t$ in chains. To this end, we use the same estimation as for dependency graphs in Sect. 4.1, i.e., we check whether there is an mgu θ of $\text{REN}^\mu(\text{CAP}^\mu(w))$ and s if $e = \mathbf{t}$ and analogously for $e = \mathbf{i}$.[19] Then we replace $s \to t$ by the new DPs $s\theta \to t\theta$ for all such mgu's θ. This is sound since in any chain $\ldots, v \to w, s \to t, \ldots$ where an instantiation of w reduces to an instantiation of s, one could use the new DP $s\theta \to t\theta$ instead.

Theorem 24 (CS-Instantiation Processor). *Let $\mathcal{P}' = \mathcal{P} \uplus \{s \to t\}$. For $d = (\mathcal{P}', \mathcal{R}, \mu, e)$, let the result of Proc(d) be $(\mathcal{P} \cup \overline{\mathcal{P}}, \mathcal{R}, \mu, e)$ where*

- $\overline{\mathcal{P}} = \{s\theta \to t\theta \mid \theta = mgu(\text{REN}^\mu(\text{CAP}^\mu(w)), s), v \to w \in \mathcal{P}'\}$, *if* $e = \boldsymbol{t}$
- $\overline{\mathcal{P}} = \{s\theta \to t\theta \mid \theta = mgu(\text{REN}^\mu_v(\text{CAP}^\mu_v(w)), s), v \to w \in \mathcal{P}', s\theta, v\theta \text{ normal}\}$, *if* $e = \boldsymbol{i}$

Then Proc is sound.

Example 25. For the TRS of Ex. 1, we still had to solve the problem $(\{(5), (11), (13), (15)\}, \mathcal{R}, \mu, \mathbf{i})$, *cf. Ex. 22. DP (11) has the variable-renamed left-hand side* $\mathsf{IF}(\mathsf{true}, x', y')$. *So the only DP that can occur before (11) in chains is* (5) *with the right-hand side* $\mathsf{IF}(\mathsf{gt}(y, \mathsf{0}), \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y)), x)$. *Recall* $\text{REN}^\mu(\text{CAP}^\mu(\mathsf{IF}(\mathsf{gt}(y, \mathsf{0}), \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y)), x))) = \mathsf{IF}(z', \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y)), x)$, *cf. Sect. 4.1. So the mgu is* $\theta = [z'/\mathsf{true},\ x'/\mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y)),\ y'/x]$. *Hence, we can replace (11) by*

$$\mathsf{IF}(\mathsf{true}, \mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y)), x) \to \mathsf{U}(\mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y))) \tag{20}$$

Here the CS variant of the instantiation processor is advantageous over the non-CS one which uses CAP *instead of* CAP^μ, *where* CAP *replaces all subterms with defined root (e.g.,* $\mathsf{minus}(\mathsf{p}(x), \mathsf{p}(y))$*) by fresh variables. So the non-CS processor would not help here as it only generates a variable-renamed copy of (11).*

When re-computing the dependency graph, there is no arc from (20) to (15) as $\mu(\mathsf{U}) = \varnothing$. *So the DP problem is decomposed into* $(\{(15)\}, \mathcal{R}, \mu, \mathbf{i})$ *(which is easily solved by the reduction pair processor) and* $(\{(5), (20), (13)\}, \mathcal{R}, \mu, \mathbf{i})$.

Now we apply the reduction pair processor again with the following rational *polynomial interpretation [11]:* $[\mathsf{M}(x, y)] = \frac{3}{2}x + \frac{1}{2}y$, $[\mathsf{minus}(x, y)] = 2x + \frac{1}{2}y$, $[\mathsf{IF}(x, y, z)] = \frac{1}{2}x + y + \frac{1}{2}z$, $[\mathsf{if}(x, y, z)] = \frac{1}{2}x + y + z$, $[\mathsf{U}(x)] = x$, $[\mathsf{p}(x)] = [\mathsf{gt}(x, y)] = \frac{1}{2}x$, $[\mathsf{s}(x)] = 2x + 2$, $[\mathsf{true}] = 1$, $[\mathsf{false}] = [\mathsf{0}] = 0$. *Then (20) is strictly decreasing and can be removed, whereas all other remaining DPs and usable rules are weakly decreasing. A last application of the dependency graph processor then detects that there is no cycle anymore and thus, it returns the empty set of DP problems. Hence, termination of the TRS from Ex. 1 is proved. As shown in our experiments in Sect. 5, this proof can easily be performed automatically.*

5 Experiments and Conclusion

We have developed a new notion of context-sensitive dependency pairs which improves significantly over previous notions. There are two main advantages:

[19] The counterexample of [4, Ex. 8] in Footnote 14 again illustrates why REN^μ_v is also needed in the innermost case (whereas this is unnecessary for non-CS rewriting).

(1) **Easier adaption of termination techniques to CS rewriting**
Now CS-DPs are very similar to DPs for ordinary rewriting and consequently, the existing powerful termination techniques from the DP framework can easily be adapted to context-sensitive rewriting. We have demonstrated this with some of the most popular DP processors in Sect. 4. Our adaptions subsume the existing earlier adaptions of the dependency graph [2], of the usable rules [20], and of the modifications for innermost rewriting [4], which were previously developed for the notion of CS-DPs from [1].

(2) **More powerful termination analysis for CS rewriting**
Due to the absence of collapsing CS-DPs, one does not have to impose extra restrictions anymore when extending the DP processors to CS rewriting, cf. Ex. 23. Hence, the power of termination proving is increased substantially.

To substantiate Claim (2), we performed extensive experiments. We implemented our new non-collapsing CS-DPs and all DP processors from this paper in the termination prover AProVE [16].[20] In contrast, the prover MU-TERM [3] uses the collapsing CS-DPs. Moreover, the processors for these CS-DPs are not formulated within the DP framework and thus, they cannot be applied in the same flexible and modular way. While MU-TERM was the most powerful tool for termination analysis of context-sensitive rewriting up to now (as demonstrated by the *International Competition of Termination Tools* 2007 [27]), due to our new notion of CS-DPs, now AProVE is substantially more powerful. For instance, AProVE easily proves termination of our leading example from Ex. 1, whereas MU-TERM fails. Moreover, we tested the tools on all 90 context-sensitive TRSs from the *Termination Problem Data Base* that was used in the competition. We used a time limit of 120 seconds for each example. Then MU-TERM can prove termination of 68 examples, whereas the new version of AProVE proves termination of 78 examples (including all 68 TRSs where MU-TERM is successful).[21] Since 4 examples are known to be non-terminating, at most 8 more of the 90 examples could potentially be detected as terminating. So due to the results of this paper, termination proving of context-sensitive rewriting has now become very powerful. To experiment with our implementation and for details, we refer to `http://aprove.informatik.rwth-aachen.de/eval/CS-DPs/`.

References

1. B. Alarcón, R. Gutiérrez, and S. Lucas. Context-sensitive dependency pairs. In *Proc. FSTTCS'06*, LNCS 4337, pages 297-308, 2006.
2. B. Alarcón, R. Gutiérrez, and S. Lucas. Improving the context-sensitive dependency graph. In *Proc. PROLE'06*, ENTCS 188, pages 91-103, 2007.
3. B. Alarcón, R. Gutiérrez, J. Iborra, S. Lucas. Proving termination of context-sensitive rewriting with MU-TERM. *Pr. PROLE'06*, ENTCS 188, p. 105-115, 2007.

[20] We also used the subterm criterion and forward instantiation processors, cf. Sect. 4.

[21] If AProVE is restricted to use exactly the same processors as MU-TERM, then it still succeeds on 74 examples. So its superiority is indeed mainly due to the new CS-DPs which enable an easy adaption of the DP framework to the CS setting.

4. B. Alarcón and S. Lucas. Termination of innermost context-sensitive rewriting using dependency pairs. In *Proc. FroCoS'07*, LNAI 4720, pages 73-87, 2007.
5. B. Alarcón, F. Emmes, C. Fuhs, J. Giesl, R. Gutiérrez, S. Lucas, P. Schneider-Kamp, and R. Thiemann. Improving context-sensitive dependency pairs. Technical Report AIB-2008-13, 2008. `http://aib.informatik.rwth-aachen.de/`.
6. T. Arts and J. Giesl. Termination of term rewriting using dependency pairs. *Theoretical Computer Science*, 236:133-178, 2000.
7. F. Baader and T. Nipkow. *Term Rewriting and All That*. Cambridge, 1998.
8. C. Borralleras, S. Lucas, and A. Rubio. Recursive path orderings can be context-sensitive. In *Proc. CADE'02*, LNAI 2392, pages 314-331, 2002.
9. N. Dershowitz. Termination by abstraction. *ICLP'04*, LNCS 3132, p. 1-18, 2004.
10. C. Fuhs, J. Giesl, A. Middeldorp, P. Schneider-Kamp, R. Thiemann, and H. Zankl. Maximal termination. In *Proc. RTA'08*, LNCS 5117, pages 110-125, 2008.
11. C. Fuhs, R. Navarro-Marset, C. Otto, J. Giesl, S. Lucas, and P. Schneider-Kamp. Search techniques for rational polynomial orders. In *Proc. AISC'08*, LNAI 5144, pages 109-124, 2008.
12. J. Giesl and A. Middeldorp. Innermost termination of context-sensitive rewriting. In *Proc. DLT'02*, LNCS 2450, pages 231-244, 2003.
13. J. Giesl and A. Middeldorp. Transformation techniques for context-sensitive rewrite systems. *Journal of Functional Programming*, 14(4):379-427, 2004.
14. J. Giesl, R. Thiemann, P. Schneider-Kamp. The DP framework: combining techniques for automated termination proofs. In *LPAR'04*, LNAI 3452, 301-331, 2005.
15. J. Giesl, R. Thiemann, P. Schneider-Kamp. Proving and disproving termination of higher-order functions. In *Proc. FroCoS'05*, LNAI 3717, pages 216-231, 2005.
16. J. Giesl, P. Schneider-Kamp, R. Thiemann. AProVE 1.2: Automatic termination proofs in the DP framework. In *Proc. IJCAR'06*, LNAI 4130, pages 281-286, 2006.
17. J. Giesl, R. Thiemann, P. Schneider-Kamp, and S. Falke. Mechanizing and improving dependency pairs. *Journal of Automatic Reasoning*, 37(3):155-203, 2006.
18. B. Gramlich. Generalized sufficient conditions for modular termination of rewriting. *Appl. Algebra in Engineering, Comm. and Computing*, 5:131-151, 1994.
19. B. Gramlich and S. Lucas. Simple termination of context-sensitive rewriting. In *Proc. RULE'02*, ACM Press, pages 29-41, 2002.
20. R. Gutiérrez, S. Lucas, and X. Urbain. Usable rules for context-sensitive rewrite systems. In *Proc. RTA'08*, LNCS 5117, pages 126-141, 2008.
21. N. Hirokawa and A. Middeldorp. Automating the dependency pair method. *Information and Computation*, 199(1,2):172-199, 2005.
22. N. Hirokawa and A. Middeldorp. Tyrolean Termination Tool: techniques and features. *Information and Computation*, 205(4):474-511, 2007.
23. S. Lucas. Context-sensitive computations in functional and functional logic programs. *Journal of Functional and Logic Programming*, 1998(1):1-61, 1998.
24. S. Lucas. Context-sensitive rewriting strategies. *Inf. Comp.*, 178(1):293-343, 2002.
25. S. Lucas. Polynomials for proving termination of context-sensitive rewriting. In *Proc. FOSSACS'04*, LNCS 2987, pages 318-332, 2004.
26. S. Lucas. Proving termination of context-sensitive rewriting by transformation. *Information and Computation*, 204(12):1782-1846, 2006.
27. C. Marché and H. Zantema. The termination competition. In *Proc. RTA'07*, LNCS 4533, pages 303-313, 2007.
28. X. Urbain. Modular & incremental automated termination proofs. *Journal of Automated Reasoning*, 32(4):315-355, 2004.