



**HAL**  
open science

# On the (In-)Completeness of Destructive Equality Resolution in the Superposition Calculus

Uwe Waldmann

► **To cite this version:**

Uwe Waldmann. On the (In-)Completeness of Destructive Equality Resolution in the Superposition Calculus. 12th International Joint Conference on Automated Reasoning (IJCAR 2024), Jul 2024, Nancy, France. pp.244-261, 10.1007/978-3-031-63498-7  
15.hal – 04907521

**HAL Id: hal-04907521**

**<https://inria.hal.science/hal-04907521v1>**

Submitted on 23 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# On the (In-)Completeness of Destructive Equality Resolution in the Superposition Calculus

Uwe Waldmann<sup>1</sup>[0000-0002-0676-7195]

MPI for Informatics, Saarland Informatics Campus, Saarbrücken, Germany  
uwe@mpi-inf.mpg.de

**Abstract.** Bachmair’s and Ganzinger’s abstract redundancy concept for the Superposition Calculus justifies almost all operations that are used in superposition provers to delete or simplify clauses, and thus to keep the clause set manageable. Typical examples are tautology deletion, subsumption deletion, and demodulation, and with a more refined definition of redundancy joinability and connectedness can be covered as well. The notable exception is DESTRUCTIVE EQUALITY RESOLUTION, that is, the replacement of a clause  $x \approx t \vee C$  with  $x \notin \text{vars}(t)$  by  $C\{x \mapsto t\}$ . This operation is implemented in state-of-the-art provers, and it is clearly useful in practice, but little is known about how it affects refutational completeness. We demonstrate on the one hand that the naive addition of DESTRUCTIVE EQUALITY RESOLUTION to the standard abstract redundancy concept renders the calculus refutationally incomplete. On the other hand, we present several restricted variants of the Superposition Calculus that are refutationally complete even with DESTRUCTIVE EQUALITY RESOLUTION.

**Keywords:** Automated theorem proving · First-order logic · Superposition calculus

## 1 Introduction

Bachmair’s and Ganzinger’s Superposition Calculus [2] comes with an abstract redundancy concept that describes under which circumstances clauses can be simplified away or deleted during a saturation without destroying the refutational completeness of the calculus. Typical concrete simplification and deletion techniques that are justified by the abstract redundancy concept are tautology deletion, subsumption deletion, and demodulation, and with a more refined definition of redundancy (Duarte and Korovin [4]) joinability and connectedness can be covered as well.

There is one simplification technique left that is not justified by Bachmair’s and Ganzinger’s redundancy criterion, namely DESTRUCTIVE EQUALITY RESOLUTION (DER), that is, the replacement of a clause  $x \approx t \vee C$  with  $x \notin \text{vars}(t)$  by  $C\{x \mapsto t\}$ . This operation is for instance implemented in the E prover (Schulz [6]), and it has been shown to be useful in practice: It increases the number of problems that E can solve and it also reduces E’s runtime per solved problem.

The question how it affects the refutational completeness of the calculus, both in theory and in practice, has been open, though (except for the special case that  $t$  is also a variable, where DER is equivalent to selecting the literal  $x \not\approx t$  so that EQUALITY RESOLUTION becomes the only possible inference with this clause).

In this paper we demonstrate on the one hand that the naive addition of DER to the standard abstract redundancy concept renders the calculus refutationally incomplete. On the other hand, we present several restricted variants of the Superposition Calculus that are refutationally complete even with DER.

By lack of space, some proofs had to be omitted from this version of the paper; they can be found in the technical report [7].

## 2 Preliminaries

**Basic Notions.** We refer to (Baader and Nipkow [1]) for basic notations and results on orderings, multiset operations, and term rewriting.

We use standard set operation symbols like  $\cup$  and  $\in$  and curly braces also for finite multisets. The union  $S \cup S'$  of the multisets  $S$  and  $S'$  over some set  $M$  is defined by  $(S \cup S')(x) = S(x) + S'(x)$  for every  $x \in M$ .

Without loss of generality we assume that all most general unifiers that we consider are idempotent. Note that if  $\sigma$  is an idempotent most general unifier and  $\theta$  is a unifier then  $\theta \circ \sigma = \theta$ .

A clause is a finite multiset of equational literals  $s \approx t$  or  $s \not\approx t$ , written as a disjunction. The empty clause is denoted by  $\perp$ . We call a literal  $L$  in a clause  $C \vee L$  maximal w.r.t. a strict literal ordering, if there is no literal in  $C$  that is larger than  $L$ ; we call it strictly maximal, if there is no literal in  $C$  that is larger than or equal to  $L$ .

We write a rewrite rule as  $u \rightarrow v$ . Semantically, a rule  $u \rightarrow v$  is equivalent to an equation  $u \approx v$ . If  $R$  is a rewrite system, that is, a set of rewrite rules, we write  $s \rightarrow_R t$  to indicate that the term  $s$  can be reduced to the term  $t$  by applying a rule from  $R$ . A rewrite system is called left-reduced, if there is no rule  $u \rightarrow v \in R$  such that  $u$  can be reduced by a rule from  $R \setminus \{u \rightarrow v\}$ .

**The Superposition Calculus.** We summarize the key elements of Bachmair's and Ganzinger's Superposition Calculus [2].

Let  $\succ$  be a reduction ordering that is total on ground terms. We extend  $\succ$  to an ordering on literals, denoted by  $\succ_L$ ,<sup>1</sup> by mapping positive literals  $s \approx t$  to multisets  $\{s, t\}$  and negative literals  $s \not\approx t$  to multisets  $\{s, s, t, t\}$  and by comparing the resulting multisets using the multiset extension of  $\succ$ . We extend the literal ordering  $\succ_L$  to an ordering on clauses, denoted by  $\succ_C$ , by comparing the multisets of literals in these clauses using the multiset extension of  $\succ_L$ .

The inference system of the Superposition Calculus consists of the rules SUPERPOSITION, EQUALITY RESOLUTION, and EQUALITY FACTORING.<sup>2</sup>

<sup>1</sup> There are several equivalent ways to define  $\succ_L$ .

<sup>2</sup> The EQUALITY FACTORING rule can be replaced by the ORDERED FACTORING and the MERGING PARAMODULATION rule. Our results hold also for this variant.

$$\text{SUPERPOSITION: } \frac{D' \vee t \approx t' \quad C' \vee L[u]}{(D' \vee C' \vee L[t'])\sigma}$$

where  $u$  is not a variable;  $\sigma = \text{mgu}(t, u)$ ;  $(C' \vee L[u])\sigma \not\leq_C (D' \vee t \approx t')\sigma$ ;  $(t \approx t')\sigma$  is strictly maximal in  $(D' \vee t \approx t')\sigma$ ; either  $L[u]$  is a positive literal  $s[u] \approx s'$  and  $L[u]\sigma$  is strictly maximal in  $(C' \vee L[u])\sigma$ , or  $L[u]$  is a negative literal  $s[u] \not\approx s'$  and  $L[u]\sigma$  is maximal in  $(C' \vee L[u])\sigma$ ;  $t\sigma \not\leq t'\sigma$ ; and  $s\sigma \not\leq s'\sigma$ .

$$\text{EQUALITY RESOLUTION: } \frac{C' \vee s \not\approx s'}{C'\sigma}$$

where  $\sigma = \text{mgu}(s, s')$  and  $(s \not\approx s')\sigma$  is maximal in  $(C' \vee s \not\approx s')\sigma$ .

$$\text{EQUALITY FACTORING: } \frac{C' \vee r \approx r' \vee s \approx s'}{(C' \vee s' \not\approx r' \vee r \approx r')\sigma}$$

where  $\sigma = \text{mgu}(s, r)$ ;  $s\sigma \not\leq s'\sigma$ ; and  $(s \approx s')\sigma$  is maximal in  $(C' \vee r \approx r' \vee s \approx s')\sigma$ .

The ordering restrictions can be overridden using *selection functions* that determine for each clause a subset of the negative literals that are available for inferences. For simplicity, we leave out this refinement in the rest of this paper. We emphasize, however, that all results that we present here hold also in the presence of selection functions; the required modifications of the proofs are straightforward.

A ground clause  $C$  is called (classically) redundant w.r.t. a set of ground clauses  $N$ , if it follows from clauses in  $N$  that are smaller than  $C$  w.r.t.  $\succ_C$ . A clause is called (classically) redundant w.r.t. a set of clauses  $N$ , if all its ground instances are redundant w.r.t. the set of ground instances of clauses in  $N$ .<sup>3</sup> A ground inference with conclusion  $C_0$  and right (or only) premise  $C$  is called redundant w.r.t. a set of ground clauses  $N$ , if one of its premises is redundant w.r.t.  $N$ , or if  $C_0$  follows from clauses in  $N$  that are smaller than  $C$ . An inference is called redundant w.r.t. a set of clauses  $N$ , if all its ground instances are redundant w.r.t. the set of ground instances of clauses in  $N$ .

Redundancy of clauses and inferences as defined above is a redundancy criterion in the sense of (Waldmann et al. [8]). It justifies typical deletion and simplification techniques such as the deletion of tautological clauses, subsumption deletion (i.e., the deletion of a clause  $C\sigma \vee D$  in the presence of a clause  $C$ ) or demodulation (i.e., the replacement of a clause  $C[s\sigma]$  by  $C[t\sigma]$  in the presence of a unit clause  $s \approx t$ , provided that  $s\sigma \succ t\sigma$ ).

### 3 Incompleteness

There are two special cases where DESTRUCTIVE EQUALITY RESOLUTION (DER) is justified by the classical redundancy criterion: First, if  $t$  is the smallest constant

<sup>3</sup> Note that “redundancy” is called “compositeness” in Bachmair and Ganzinger’s *J. Log. Comput.* article [2]. In later papers the standard terminology has changed.

in the signature, then every ground instance  $(x \not\approx t \vee C)\theta$  follows from the smaller ground instance  $C\{x \mapsto t\}\theta$ . Second, if  $t$  is another variable  $y$ , then every ground instance  $(x \not\approx y \vee C)\theta$  follows from the smaller ground instance  $C\{x \mapsto y\}\{y \mapsto s\}\theta$ , where  $s$  is the smaller of  $x\theta$  and  $y\theta$ .

But it is easy to see that this does not work in general: Let  $\succ$  be a Knuth-Bendix ordering with weights  $w(f) = w(b) = 2$ ,  $w(c) = w(d) = 1$ ,  $w(z) = 1$  for all variables  $z$ , and let  $C$  be the clause  $x \not\approx b \vee f(x) \approx d$ . Then DER applied to  $C$  yields  $D = f(b) \approx d$ . Now consider the substitution  $\theta = \{x \mapsto c\}$ . The ground instance  $C\theta = c \not\approx b \vee f(c) \approx d$  is a logical consequence of  $D$ , but since it is smaller than  $D$  itself,  $D$  makes neither  $C\theta$  nor  $C$  redundant.

Moreover, the following example demonstrates that the Superposition Calculus becomes indeed incomplete, if we add DER as a simplification rule, i.e., if we extend the definition of redundancy in such a way that the conclusion of DER renders the premise redundant.

**Example 1.** Let  $\succ$  be a Knuth-Bendix ordering with weights  $w(f) = 4$ ,  $w(g) = 3$ ,  $w(b) = 4$ ,  $w(b') = 2$ ,  $w(c) = w(c') = w(d) = 1$ ,  $w(z) = 1$  for all variables  $z$ , and let  $N$  be the set of clauses

$$\begin{aligned} C_1 &= \underline{f(x, d)} \approx x \\ C_2 &= \underline{f(x, y)} \not\approx b \vee g(x) \approx d \\ C_3 &= b' \approx c' \vee \underline{b} \approx c \\ C_4 &= \underline{g(b')} \not\approx g(c') \\ C_5 &= \underline{g(c)} \not\approx d \end{aligned}$$

where all the maximal terms in maximal literals are underlined.

At this point, neither demodulation nor subsumption is possible. The only inference that must be performed is SUPERPOSITION between  $C_1$  and  $C_2$ , yielding

$$C_6 = x \not\approx b \vee g(x) \approx d$$

and by using DER,  $C_6$  is replaced by

$$C_7 = g(b) \approx d$$

We could now continue with a SUPERPOSITION between  $C_3$  and  $C_7$ , followed by a SUPERPOSITION with  $C_5$ , followed by EQUALITY RESOLUTION, and obtain

$$C_8 = b' \approx c'$$

from which we can derive the empty clause by SUPERPOSITION with  $C_4$  and once more by EQUALITY RESOLUTION. However, clause  $C_7$  is in fact redundant: The ground clauses  $C_3$  and  $C_4$  imply  $b \approx c$ ; therefore  $C_7$  follows from  $C_3$ ,  $C_4$ , and the ground instances

$$\begin{aligned} C_1\{x \mapsto c\} &= f(c, d) \approx c \\ C_2\{x \mapsto c, y \mapsto d\} &= f(c, d) \not\approx b \vee g(c) \approx d \end{aligned}$$

Because all terms in these clauses are smaller than the maximal term  $g(b)$  of  $C_7$ , all the clauses are smaller than  $C_7$ . Since  $C_7$  is redundant, we are allowed to delete it, and then no further inferences are possible anymore. Therefore the clause set  $N = \{C_1, \dots, C_5\}$  is saturated, even though it is inconsistent and does not contain the empty clause, which implies that the calculus is not refutationally complete anymore.

## 4 Completeness, Part I: The Horn Case

### 4.1 The Idea

On the one hand, Example 1 demonstrates that we cannot simply extend the standard redundancy criterion of the Superposition Calculus with DER without destroying refutational completeness, and that this holds even if we impose a particular strategy on simplification steps (say, that simplifications must be performed eagerly and that demodulation and subsumption have a higher precedence than DER). On the other hand, Example 1 is of course highly unrealistic: Even though clause  $C_7$  is redundant w.r.t. the clauses  $C_1$ ,  $C_2$ ,  $C_3$ , and  $C_4$ , no reasonable superposition prover would ever detect this – in particular, since doing so would require to invent the instance  $C_2\{x \mapsto c, y \mapsto d\}$  of  $C_2$ , which is not in any way syntactically related to  $C_7$ .<sup>4</sup>

This raises the question whether DER still destroys refutational completeness when we restrict the other deletion and simplification techniques to those that are typically implemented in superposition provers, such as tautology detection, demodulation, or subsumption. Are there alternative redundancy criteria that are refutationally complete together with the Superposition Calculus and that justify DER as well as (all/most) commonly implemented deletion and simplification techniques? Given the usual structure of the inductive completeness proofs for saturation calculi, developing such a redundancy criterion would mean in particular to find a suitable clause ordering with respect to which certain clauses have to be smaller than others. The following example illustrates a fundamental problem that we have to deal with:

**Example 2.** Let  $\succ$  be a Knuth-Bendix ordering with weights  $w(f) = w(g) = w(h) = w(c) = 1$ ,  $w(b) = 2$ ,  $w(z) = 1$  for all variables  $z$ . Consider the following set of clauses:

$$\begin{array}{ll} D_1 = h(x) \approx x & C_1 = h(x) \not\approx b \vee f(g(x)) \approx c \\ & C_2 = x \not\approx b \vee f(g(x)) \approx c \\ D_3 = h(c) \not\approx b \vee g(b) \approx g(c) & C_3 = f(g(b)) \approx c \\ & C_4 = h(c) \not\approx b \vee f(g(c)) \approx c \end{array}$$

<sup>4</sup> In fact, a prover might use SMT-style heuristic grounding of non-ground clauses, but then finding the contradiction turns out to be easier than proving the redundancy of  $C_7$ .

Demodulation of  $C_1$  using  $D_1$  yields  $C_2$ , and if we want Demodulation to be a simplification, then every ground instance  $C_1\theta$  should be larger than the corresponding ground instance  $C_2\theta$  in the clause ordering.

DER of  $C_2$  yields  $C_3$ , and if we want DER to be a simplification, then every ground instance  $C_2\theta$  should be larger than  $C_3\theta = C_3$ .

A SUPERPOSITION inference between  $D_3$  and  $C_3$  yields  $C_4$ . The inductive completeness proof for the calculus relies on the fact that the conclusion of an inference is smaller than the largest premise, so  $C_3$  should be larger than  $C_4$ .

By transitivity we obtain that every ground instance  $C_1\theta$  should be larger than  $C_4$  in the clause ordering. The clause  $C_4$ , however, *is* a ground instance of  $C_1$ , which is clearly a contradiction.

On the other hand, a closer inspection reveals that, depending on the limit rewrite system  $R_*$  that is produced in the completeness proof for the Superposition Calculus, the SUPERPOSITION inference between  $D_3$  and  $C_3$  is only needed, when  $D_3$  produces the rewrite rule  $g(b) \rightarrow g(c) \in R_*$ , and that the only critical case for DER is the one where  $b$  can be reduced by some rule in  $R_*$ . Since the limit rewrite system  $R_*$  is by construction left-reduced, these two conditions are mutually exclusive. This observation indicates that we might be able to find a suitable clause ordering if we choose it depending on  $R_*$ .

## 4.2 Ground Case

**The Normalization Closure Ordering.** Let  $\succ$  be a reduction ordering that is total on ground terms. Let  $R$  be a left-reduced ground rewrite system contained in  $\succ$ .

For technical reasons that will become clear later, we design our ground superposition calculus in such a way that it operates on ground closures  $(C \cdot \theta)$ . Logically, a ground closure  $(C \cdot \theta)$  is equivalent to a ground instance  $C\theta$ , but an ordering may treat two closures that represent the same ground instance in different ways. We consider closures up to  $\alpha$ -renaming and ignore the behavior of  $\theta$  on variables that do not occur in  $C$ , that is, we treat closures  $(C_1 \cdot \theta_1)$  and  $(C_2 \cdot \theta_2)$  as equal whenever  $C_1$  and  $C_2$  are equal up to bijective variable renaming and  $C_1\theta_1 = C_2\theta_2$ . We also identify  $(\perp \cdot \theta)$  and  $\perp$ .

Intuitively, in order to compare ground closures  $C \cdot \theta$ , we normalize all terms occurring in  $C\theta$  with  $R$ , we compute the multiset of all the redexes occurring during the normalization and all the resulting normal forms, and we compare these multisets using the multiset extension of  $\succ$ . Since we would like to give redexes and normal forms in negative literals a slightly larger weight than redexes and normal forms in positive literals, and redexes in positive literals below the top a slightly larger weight than redexes at the top, we combine each of these terms with a label (0 for positive at the top, 1 for positive below the top, 2 for negative). Moreover, whenever some term  $t$  occurs several times in  $C$  as a subterm, we want to count the redexes resulting from the normalization of  $t\theta$  only once (with the maximum of the labels). The reason for this is that DER can produce several copies of the same term  $t$  in a clause if the variable to be

eliminated occurs several times in the clause; by counting all redexes stemming from  $t$  only once, we ensure that this does not increase the total number of redexes. Formally, we first compute the set (not multiset!) of all subterms  $t$  of  $C$ , so that duplicates are deleted, and then compute the multiset of redexes for all terms  $t\theta$  (and analogously for terms occurring at the top of a literal).

**Definition 3.** We define the subterm sets  $ss_{>\epsilon}^+(C)$  and  $ss^-(C)$  and the topterm sets  $ts^+(C)$  and  $ts^-(C)$  of a clause  $C$  by

$$\begin{aligned} ss^-(C) &= \{t \mid C = C' \vee s[t]_p \not\approx s'\} \\ ss_{>\epsilon}^+(C) &= \{t \mid C = C' \vee s[t]_p \approx s', p > \epsilon\} \\ ts^-(C) &= \{t \mid C = C' \vee t \not\approx t'\} \\ ts^+(C) &= \{t \mid C = C' \vee t \approx t'\}. \end{aligned}$$

We define the labeled subterm set  $lss(C)$  and the labeled topterm set  $lts(C)$  of a clause  $C$  by

$$\begin{aligned} lss(C) &= \{(t, 2) \mid t \in ss^-(C)\} \\ &\quad \cup \{(t, 1) \mid t \in ss_{>\epsilon}^+(C) \setminus ss^-(C)\} \\ &\quad \cup \{(t, 0) \mid t \in ts^+(C) \setminus (ss_{>\epsilon}^+(C) \cup ss^-(C))\} \\ lts(C) &= \{(t, 2) \mid t \in ts^-(C)\} \cup \{(t, 0) \mid t \in ts^+(C) \setminus ts^-(C)\}. \end{aligned}$$

We define the  $R$ -redex multiset  $rm_R(t, m)$  of a labeled ground term  $(t, m)$  with  $m \in \{0, 1, 2\}$  by

$$\begin{aligned} rm_R(t, m) &= \emptyset \text{ if } t \text{ is } R\text{-irreducible;} \\ rm_R(t, m) &= \{(u, m)\} \cup rm_R(t', m) \text{ if } t \rightarrow_R t' \text{ using the rule } u \rightarrow v \in R \\ &\quad \text{at position } p \text{ and } p = \epsilon \text{ or } m > 0; \\ rm_R(t, m) &= \{(u, 1)\} \cup rm_R(t', m) \text{ if } t \rightarrow_R t' \text{ using the rule } u \rightarrow v \in R \\ &\quad \text{at position } p \text{ and } p > \epsilon \text{ and } m = 0. \end{aligned}$$

**Lemma 4.** For every left-reduced ground rewrite system  $R$  contained in  $\succ$ ,  $rm_R(t, m)$  is well-defined.

**Definition 5.** We define the  $R$ -normalization multiset  $nm_R(C \cdot \theta)$  of a ground closure  $(C \cdot \theta)$  by

$$\begin{aligned} nm_R(C \cdot \theta) &= \bigcup_{(f(t_1, \dots, t_n), m) \in lss(C)} rm_R(f(t_1 \theta \downarrow_R, \dots, t_n \theta \downarrow_R), m) \\ &\quad \cup \bigcup_{(x, m) \in lss(C)} rm_R(x\theta, m) \\ &\quad \cup \bigcup_{(t, m) \in lts(C)} \{(t\theta \downarrow_R, m)\} \end{aligned}$$

**Example 6.** Let  $C = h(g(g(x))) \approx f(f(b))$ ; let  $\theta = \{x \mapsto b\}$ . Then  $lss(C) = \{(h(g(g(x))), 0), (g(g(x)), 1), (g(x), 1), (x, 1), (f(f(b)), 0), (f(b), 1), (b, 1)\}$  and  $lts(C) = \{(h(g(g(x))), 0), (f(f(b)), 0)\}$ .



Let  $R = \{f(b) \rightarrow b, g(g(b)) \rightarrow b\}$ . Then  $\text{nm}_R(C \cdot \theta) = \{(g(g(b)), 1), (f(b), 1), (f(b), 0), (h(b), 0), (b, 0)\}$ , where the first element is a redex from the normalization of  $g(g(x))\theta$ , the second from the normalization of  $f(b)\theta$ , the third from the normalization of  $f(f(b))\theta$ . The remaining elements are the normal forms of  $h(g(g(x)))\theta$  and  $f(f(b))\theta$ .

The  $R$ -normalization closure ordering  $\succ_R$  compares ground closures  $(C \cdot \theta_1)$  and  $(D \cdot \theta_2)$  using a lexicographic combination of three orderings:

- first, the multiset extension  $((\succ, >)_{\text{lex}})_{\text{mul}}$  of the lexicographic combination of the reduction ordering  $\succ$  and the ordering  $>$  on natural numbers applied to the multisets  $\text{nm}_R(C \cdot \theta_1)$  and  $\text{nm}_R(D \cdot \theta_2)$ ,
- second, the traditional clause ordering  $\succ_c$  applied to  $C\theta_1$  and  $D\theta_2$ ,
- and third, an arbitrary well-founded ordering  $\succ_{\text{cio}}$  on ground closures that is total on ground closures  $(C \cdot \theta_1)$  and  $(D \cdot \theta_2)$  with  $C\theta_1 = D\theta_2$  and that has the property that  $(C \cdot \theta_1) \succ_{\text{cio}} (D \cdot \theta_2)$  whenever  $C\theta_1 = D\theta_2$  and  $D$  is an instance of  $C$  but not vice versa.

**Lemma 7.** *If  $(C \cdot \theta)$  and  $(C\sigma \cdot \theta')$  are ground closures, such that  $C\theta = C\sigma\theta'$ , and  $C$  and  $C\sigma$  are not equal up to bijective renaming, then  $(C \cdot \theta) \not\succeq_R (C\sigma \cdot \theta')$ .*

**Example 8.** Let  $C = h(f(x)) \approx f(y)$ , let  $\theta' = \{x \mapsto b\}$ ; let  $\theta = \{x \mapsto b, y \mapsto b\}$ ; let  $\sigma = \{y \mapsto x\}$ . Let  $R = \{f(b) \rightarrow b\}$ .

Then  $\text{nm}_R(C \cdot \theta) = \{(f(b), 1), (f(b), 0), (h(b), 0), (b, 0)\}$  and  $\text{nm}_R(C\sigma \cdot \theta') = \{(f(b), 1), (h(b), 0), (b, 0)\}$ , and therefore  $(C \cdot \theta) \not\succeq_R (C\sigma \cdot \theta')$ . The subterm  $f(x)$  occurs twice in  $C\sigma$  (with labels 0 and 1), but only once in  $\text{lss}(C\sigma)$  (with the larger of the two labels), and the same holds for the redex  $f(b)$  stemming from  $f(x)\theta'$  in  $\text{nm}_R(C\sigma \cdot \theta')$ .

**Parallel Superposition.** In the normalization closure ordering, redexes and normal forms stemming from several occurrences of the same term  $u$  in a closure  $(C \cdot \theta)$  are counted only once. When we perform a SUPERPOSITION inference, this fact leads to a small problem: Consider a closure  $(C[u, u] \cdot \theta)$ . In the  $R$ -normalization multiset of this closure, the redexes stemming from the two copies of  $u\theta$  are counted only once. Now suppose that one of the two copies of  $u$  is replaced by a smaller term  $v$  in a SUPERPOSITION inference. The resulting closure  $(C[v, u] \cdot \theta)$  should be smaller than the original one, but it isn't: The redexes stemming from  $u\theta$  are still counted once, and additionally, the  $R$ -normalization multiset now contains the redexes stemming from  $v\theta$ .

There is an easy fix for this problem, though: We have to replace the ordinary SUPERPOSITION rule by a PARALLEL SUPERPOSITION rule, in which *all* copies of a term  $u$  in a clause  $C$  are replaced whenever one copy occurs in a maximal side of a maximal literal. Note that this is a well-known optimization that superposition provers implement (or should implement) anyhow.

We need one further modification of the inference rule: The side conditions of the superposition calculus use the traditional clause ordering  $\succ_c$ , but our completeness proof and redundancy criterion will be based on the orderings  $\succ_R$ .

The difference between these orderings becomes relevant in particular when we consider (PARALLEL) SUPERPOSITION inferences where the clauses overlap at the top of a positive literal. In this case, the  $\succ_R$ -smaller of the two premises may actually be the  $\succ_C$ -larger one. Therefore, the usual condition that the left premise of a (PARALLEL) SUPERPOSITION inference has to be  $\succ_C$ -minimal has to be dropped for these inferences.

$$\text{PARALLEL SUPERPOSITION: } \frac{D' \vee t \approx t' \quad C[u, \dots, u]_{p_1, \dots, p_k}}{(D' \vee C[t', \dots, t']_{p_1, \dots, p_k})\sigma}$$

where  $u$  is not a variable;  $\sigma = \text{mgu}(t, u)$ ;  $p_1, \dots, p_k$  are all the occurrences of  $u$  in  $C$ ; if one of the occurrences of  $u$  in  $C$  is in a negative literal or below the top in a positive literal then  $C\sigma \not\prec_C (D' \vee t \approx t')\sigma$ ;  $(t \approx t')\sigma$  is strictly maximal in  $(D' \vee t \approx t')\sigma$ ; either one of the occurrences of  $u$  in  $C$  is in a positive literal  $L[u] = s[u] \approx s'$  and  $L[u]\sigma$  is strictly maximal in  $C\sigma$ , or one of the occurrences of  $u$  in  $C$  is in a negative literal  $L[u] = s[u] \not\approx s'$  and  $L[u]\sigma$  is maximal in  $(C' \vee L[u])\sigma$ ;  $t\sigma \not\prec t'\sigma$ ; and  $s\sigma \not\prec s'\sigma$ .

**Ground Closure Horn Superposition.** We will show that our calculus is refutationally complete for Horn clauses by lifting a similar result for ground closure Horn superposition. We emphasize that our calculus is not a basic or constraint calculus such as (Bachmair et al. [3]) or (Nieuwenhuis and Rubio [5]). Even though the ground version that we present here operates on closures, it is essentially a rephrased version of the standard ground Superposition Calculus. This explains why we also have to consider superpositions below variable positions.

The ground closure calculus uses the following three inference rules. We assume that in binary inferences the variables in the premises  $(D \cdot \theta_2)$  and  $(C \cdot \theta_1)$  are renamed in such a way that  $C$  and  $D$  do not share variables. We can then assume without loss of generality that the substitutions  $\theta_2$  and  $\theta_1$  agree.

$$\text{PARALLEL SUPERPOSITION I: } \frac{(D' \vee t \approx t' \cdot \theta) \quad (C[u, \dots, u]_{p_1, \dots, p_k} \cdot \theta)}{((D' \vee C[t', \dots, t']_{p_1, \dots, p_k})\sigma \cdot \theta)}$$

where  $u$  is not a variable;  $t\theta = u\theta$ ;  $\sigma = \text{mgu}(t, u)$ ;  $p_1, \dots, p_k$  are all the occurrences of  $u$  in  $C$ ; if one of the occurrences of  $u$  in  $C$  is in a negative literal or below the top in a positive literal then  $(D' \vee t \approx t')\theta \prec_C C\theta$ ; one of the occurrences of  $u$  in  $C$  is either in a positive literal  $s[u] \approx s'$  such that  $(s[u] \approx s')\theta$  is strictly maximal in  $C\theta$  or in a negative literal  $s[u] \not\approx s'$  such that  $(s[u] \not\approx s')\theta$  is maximal in  $C\theta$ ;  $s[u]\theta \succ s'\theta$ ;  $(t \approx t')\theta$  is strictly maximal in  $(D' \vee t \approx t')\theta$ ; and  $t\theta \succ t'\theta$ .

$$\text{PARALLEL SUPERPOSITION II: } \frac{(D' \vee t \approx t' \cdot \theta) \quad (C \cdot \theta)}{(D' \vee C \cdot \theta[x \mapsto u[t\theta]])}$$

where  $x$  is a variable of  $C$ ;  $x\theta = u[t\theta]$ ; if one of the occurrences of  $x$  in  $C$  is in a negative literal or below the top in a positive literal then  $(D' \vee t \approx t')\theta \prec_C C\theta$ ;

one of the occurrences of  $x$  in  $C$  is either in a positive literal  $s[x] \approx s'$  such that  $(s[x] \approx s')\theta$  is strictly maximal in  $C\theta$  or in a negative literal  $s[x] \not\approx s'$  such that  $(s[x] \approx s')\theta$  is maximal in  $C\theta$ ;  $s[x]\theta \succ s'\theta$ ;  $(t \approx t')\theta$  is strictly maximal in  $(D' \vee t \approx t')\theta$ ; and  $t\theta \succ t'\theta$ .

$$\text{EQUALITY RESOLUTION: } \frac{(C' \vee s \not\approx s' \cdot \theta)}{(C' \sigma \cdot \theta)}$$

where  $s\theta = s'\theta$ ;  $\sigma = \text{mgu}(s, s')$ ; and  $(s \not\approx s')\theta$  is maximal in  $(C' \vee s \not\approx s')\theta$ .

The following lemmas compare the conclusion  $\text{concl}(\iota)$  of an inference  $\iota$  with its right or only premise:

**Lemma 9.** *Let  $\iota$  be a ground EQUALITY RESOLUTION inference. Then  $\text{concl}(\iota)$  is  $\succ_R$ -smaller than its premise.*

**Lemma 10.** *Let  $\iota$  be a ground PARALLEL SUPERPOSITION inference*

$$\frac{(D' \vee t \approx t' \cdot \theta) \quad (C[u, \dots, u]_{p_1, \dots, p_k} \cdot \theta)}{((D' \vee C[t', \dots, t']_{p_1, \dots, p_k}) \sigma \cdot \theta)}$$

with  $t\theta = u\theta$  and  $\sigma = \text{mgu}(t, u)$  or

$$\frac{(D' \vee t \approx t' \cdot \theta) \quad (C \cdot \theta)}{(D' \vee C \cdot \theta[x \mapsto u[t'\theta]])}$$

with  $x\theta = u[t'\theta]$ . If  $(t\theta \rightarrow t'\theta) \in R$ , then  $\text{concl}(\iota)$  is  $\succ_R$ -smaller than  $(C \cdot \theta)$ .

*Proof.* Since  $t\theta$  is replaced by  $t'\theta$  at all occurrences of  $u$  or at or below all occurrences of  $x$  in  $C$ , one copy of the redex  $t\theta$  is removed from  $\text{nm}_R(C \cdot \theta)$ . Moreover all terms in  $D'\theta$  are smaller than  $t\theta$ , and consequently all redexes stemming from  $D'\theta$  are smaller than  $t\theta$ . Therefore  $\text{nm}_R(C \cdot \theta)$  is larger than  $\text{nm}_R(D' \vee C[t', \dots, t']_{p_1, \dots, p_k} \cdot \theta)$  or  $\text{nm}_R(D' \vee C \cdot \theta[x \mapsto u[t'\theta]])$ . In the second case, this implies  $(C \cdot \theta) \succ_R \text{concl}(\iota)$  immediately. In the first case, it implies  $(C \cdot \theta) \succ_R (D' \vee C[t', \dots, t']_{p_1, \dots, p_k} \cdot \theta)$  and  $(C \cdot \theta) \succ_R \text{concl}(\iota)$  follows using Lemma 7.  $\square$

**Redundancy.** We will now construct a redundancy criterion for ground closure Horn superposition that is based on the ordering(s)  $\succ_R$ .

**Definition 11.** Let  $N$  be a set of ground closures. A ground closure  $(C \cdot \theta)$  is called redundant w.r.t.  $N$ , if for every left-reduced ground rewrite system  $R$  contained in  $\succ$  we have (i)  $R \models (C \cdot \theta)$  or (ii) there exists a ground closure  $(D \cdot \theta) \in N$  such that  $(D \cdot \theta) \prec_R (C \cdot \theta)$  and  $R \not\models (D \cdot \theta)$ .

**Definition 12.** Let  $N$  be a set of ground closures. A ground inference  $\iota$  with right or only premise  $(C \cdot \theta)$  is called redundant w.r.t.  $N$ , if for every left-reduced ground rewrite system  $R$  contained in  $\succ$  we have (i)  $R \models \text{concl}(\iota)$ , or (ii) there exists a ground closure  $(C' \cdot \theta) \in N$  such that  $(C' \cdot \theta) \prec_R (C \cdot \theta)$  and  $R \not\models (C' \cdot \theta)$ , or (iii)  $\iota$  is a SUPERPOSITION inference with left premise  $(D' \vee t \approx t' \cdot \theta)$  where  $t\theta \succ t'\theta$ , and  $(t\theta \rightarrow t'\theta) \notin R$ , or (iv)  $\iota$  is a SUPERPOSITION inference where the left premise is not the  $\succ_R$ -minimal premise.

Intuitively, a redundant closure cannot be a minimal counterexample, i.e., a minimal closure that is false in  $R$ . A redundant inference is either irrelevant for the completeness proof (cases (iii) and (iv)), or its conclusion (and thus its right or only premise) is true in  $R$ , provided that all closures that are  $\succ_R$ -smaller than the right or only premise are true in  $R$  (cases (i) and (ii)) – which means that the inference can be used to show that the right or only premise cannot be a minimal counterexample.

We denote the set of redundant closures w.r.t.  $N$  by  $Red_C(N)$  and the set of redundant inferences by  $Red_I(N)$ .

**Example 13.** Let  $\succ$  be a KBO where all symbols have weight 1. Let  $C = g(b) \not\approx c \vee f(c) \not\approx d$  and  $C' = f(g(b)) \not\approx d$ . Then the closure  $(C \cdot \emptyset)$  is redundant w.r.t.  $\{(C' \cdot \emptyset)\}$ : Let  $R$  be a left-reduced ground rewrite system contained in  $\succ$ . Assume that  $C$  is false in  $R$ . Then  $g(b)$  and  $c$  have the same  $R$ -normal form. Consequently, every redex or normal form in  $nm_R(C' \cdot \emptyset)$  was already present in  $nm_R(C \cdot \emptyset)$ . Moreover, the labeled normal form  $(c \downarrow_R, 2)$  that is present in  $nm_R(C \cdot \emptyset)$  is missing in  $nm_R(C' \cdot \emptyset)$ . Therefore  $(C \cdot \emptyset) \succ_R (C' \cdot \emptyset)$ . Besides, if  $(C \cdot \emptyset)$  is false in  $R$ , then  $(C' \cdot \emptyset)$  is false as well.

Note that  $C \prec_C C'$ , therefore  $C$  is not classically redundant w.r.t.  $\{C'\}$ .

**Lemma 14.** ( $Red_I, Red_C$ ) is a redundancy criterion in the sense of (Waldmann et al. [8]), that is, (1) if  $N \models \perp$ , then  $N \setminus Red_C(N) \models \perp$ ; (2) if  $N \subseteq N'$ , then  $Red_C(N) \subseteq Red_C(N')$  and  $Red_I(N) \subseteq Red_I(N')$ ; (3) if  $N' \subseteq Red_C(N)$ , then  $Red_C(N) \subseteq Red_C(N \setminus N')$  and  $Red_I(N) \subseteq Red_I(N \setminus N')$ ; and (4) if  $\iota$  is an inference with conclusion in  $N$ , then  $\iota \in Red_I(N)$ .

*Proof.* (1) Suppose that  $N \setminus Red_C(N) \not\models \perp$ . Then there exists a left-reduced ground rewrite system  $R$  contained in  $\succ$  such that  $R \models N \setminus Red_C(N)$ . We show that  $R \models N$  (which implies  $N \not\models \perp$ ). Assume that  $R \not\models N$ . Then there exists a closure  $(C \cdot \theta) \in N \cap Red_C(N)$  such that  $R \not\models (C \cdot \theta)$ . By well-foundedness of  $\succ_R$  there exists a  $\succ_R$ -minimal closure  $(C \cdot \theta)$  with this property. By definition of  $Red_C(N)$ , there must be a ground closure  $(D \cdot \theta) \in N$  such that  $(D \cdot \theta) \prec_R (C \cdot \theta)$  and  $R \not\models (D \cdot \theta)$ . By minimality of  $(C \cdot \theta)$ , we get  $(D \cdot \theta) \in N \setminus Red_C(N)$ , contradicting the initial assumption.

(2) Obvious.

(3) Let  $N' \subseteq Red_C(N)$  and let  $(C \cdot \theta) \in Red_C(N)$ . We show that  $(C \cdot \theta) \in Red_C(N \setminus N')$ . Choose  $R$  arbitrarily. If  $R \models (C \cdot \theta)$ , we are done. Otherwise there exists a ground closure  $(D \cdot \theta) \in N$  such that  $(D \cdot \theta) \prec_R (C \cdot \theta)$  and  $R \not\models (D \cdot \theta)$ . By well-foundedness of  $\succ_R$  there exists a  $\succ_R$ -minimal closure  $(D \cdot \theta)$  with this property. If  $(D \cdot \theta)$  were contained in  $N'$  and hence in  $Red_C(N)$ , there would exist a ground closure  $(D' \cdot \theta) \in N$  such that  $(D' \cdot \theta) \prec_R (D \cdot \theta)$  and  $R \not\models (D' \cdot \theta)$ , contradicting minimality. Therefore  $(D \cdot \theta) \in N \setminus N'$  as required. The second part of (3) is proved analogously.

(4) Let  $\iota$  be an inference with  $\text{concl}(\iota) \in N$ . Choose  $R$  arbitrarily. We have to show that  $\iota$  satisfies part (i), (ii), (iii), or (iv) of Def. 12. Assume that (i), (iii), and (iv) do not hold. Then  $R \not\models \text{concl}(\iota)$ , and by Lemmas 9 and 10,  $\text{concl}(\iota)$  is

$\succ_R$ -smaller than the right or only premise of  $\iota$ , therefore part (ii) is satisfied if we take  $\text{concl}(\iota)$  as  $(C' \cdot \theta)$ .  $\square$

**Constructing a Candidate Interpretation.** In usual completeness proofs for superposition-like calculi, one constructs a candidate interpretation (a set of ground rewrite rules) for a saturated set of ground clauses by induction over the *clause ordering*. In our case, this is impossible since the limit *closure ordering* depends on the generated set of rewrite rules itself. We can still construct the candidate interpretation by induction over the *term ordering*, though: Instead of inspecting ground closures one by one as in the classical construction, we inspect all ground closures  $(C \cdot \theta)$  for which  $C\theta$  contains the maximal term  $s$  simultaneously, and if for at least one of them the usual conditions for productivity are satisfied, we choose the  $\succ_{R_s}$ -smallest one of these to extend  $R_s$ .

Let  $N$  be a set of ground closures. For every ground term  $s$  we define  $R_s = \bigcup_{t \prec_s} E_t$ . Furthermore we define  $E_s = \{s \rightarrow s'\}$ , if  $(C \cdot \theta)$  is the  $\succ_{R_s}$ -smallest closure in  $N$  such that  $C = C' \vee u \approx u'$ ,  $s = u\theta$  is a strictly maximal term in  $C\theta$ , occurs only in a positive literal of  $C\theta$ , and is irreducible w.r.t.  $R_s$ ,  $s' = u'\theta$ ,  $C\theta$  is false in  $R_s$ , and  $s \succ s'$ , provided that such a closure  $(C \cdot \theta)$  exists. We say that  $(C \cdot \theta)$  produces  $s \rightarrow s'$ . If no such closure exists, we define  $E_s = \emptyset$ . Finally, we define  $R_* = \bigcup_t E_t$ .

The following two lemmas are proved as usual:

**Lemma 15.** *Let  $s$  be a ground term, let  $(C \cdot \theta)$  be a closure. If every term that occurs in negative literals of  $C\theta$  is smaller than  $s$  and every term that occurs in positive literals of  $C\theta$  is smaller than or equal to  $s$ , and if  $R_s \models (C \cdot \theta)$ , then  $R_* \models (C \cdot \theta)$ .*

**Lemma 16.** *If a closure  $(C' \vee u \approx u' \cdot \theta)$  produces  $u\theta \rightarrow u'\theta$ , then  $R_* \models (C' \vee u \approx u' \cdot \theta)$  and  $R_* \not\models (C' \cdot \theta)$ .*

**Lemma 17.** *Let  $(C_1 \cdot \theta)$  and  $(C_2 \cdot \theta)$  be two closures. If  $s$  is a strictly maximal term and occurs only positively in both  $C_1\theta$  and  $C_2\theta$ , then  $(C_1 \cdot \theta) \succ_{R_s} (C_2 \cdot \theta)$  if and only if  $(C_1 \cdot \theta) \succ_{R_*} (C_2 \cdot \theta)$ .*

**Lemma 18.** *Let  $(D \cdot \theta) = (D' \vee t \approx t' \cdot \theta)$  and  $(C \cdot \theta)$  be two closures in  $N$ . If  $(D \cdot \theta)$  produces  $t\theta \rightarrow t'\theta$  in  $R_*$ , and  $t\theta$  occurs in  $C\theta$  in a negative literal or below the top a term in a positive literal, then  $(D \cdot \theta) \prec_{R_*} (C \cdot \theta)$  and  $D\theta \prec_C C\theta$ .*

**Lemma 19.** *Let  $(D \cdot \theta) = (D' \vee t \approx t' \cdot \theta)$  and  $(C \cdot \theta)$  be two closures in  $N$ . If  $(D \cdot \theta)$  produces  $t\theta \rightarrow t'\theta$  in  $R_*$ ,  $t\theta$  occurs in  $C\theta$  at the top of the strictly maximal side of a positive maximal literal, and  $R_* \not\models (C \cdot \theta)$ , then  $(D \cdot \theta) \prec_{R_*} (C \cdot \theta)$ .*

We can now show that the Ground Closure Horn Superposition Calculus is refutationally complete:

**Theorem 20.** *Let  $N$  be a saturated set of ground closures that does not contain  $(\perp \cdot \theta)$ . Then  $R_* \models N$ .*

*Proof.* Suppose that  $R_* \not\models N$ . Let  $(C \cdot \theta)$  be the  $\succ_{R_*}$ -smallest closure in  $N$  such that  $R_* \not\models (C \cdot \theta)$ .

Case 1:  $C = C' \vee s \not\approx s'$  and  $s\theta \not\approx s'\theta$  is maximal in  $C\theta$ . By assumption,  $R_* \not\models s\theta \not\approx s'\theta$ , hence  $s\theta \downarrow_{R_*} = s'\theta \downarrow_{R_*}$ .

Case 1.1:  $s\theta = s'\theta$ . Then there is an EQUALITY RESOLUTION inference from  $(C \cdot \theta)$  with conclusion  $(C'\sigma \cdot \theta)$ , where  $\theta \circ \sigma = \theta$ . By saturation the inference is redundant, and by minimality of  $(C \cdot \theta)$  w.r.t.  $\succ_{R_*}$  this implies  $R_* \models (C'\sigma \cdot \theta)$ . But then  $R_* \models (C \cdot \theta)$ , contradicting the assumption.

Case 1.2:  $s\theta \neq s'\theta$ . W.l.o.g. let  $s\theta \succ s'\theta$ . Then  $s\theta$  must be reducible by a rule  $t\theta \rightarrow t'\theta \in R_*$ , which has been produced by a closure  $(D \cdot \theta) = (D' \vee t \approx t' \cdot \theta)$  in  $N$ . By Lemma 18,  $(D \cdot \theta) \prec_{R_*} (C \cdot \theta)$  and  $D\theta \prec_C C\theta$ . If  $s\theta$  and  $t\theta$  overlap at a non-variable position of  $s$ , there is a PARALLEL SUPERPOSITION I inference  $\iota$  between  $(D \cdot \theta)$  and  $(C \cdot \theta)$ ; otherwise they overlap at or below a variable position of  $s$  and there is a PARALLEL SUPERPOSITION II inference  $\iota$  with premises  $(D \cdot \theta)$  and  $(C \cdot \theta)$ . By Lemma 16,  $R_* \not\models (D' \cdot \theta)$ . By saturation the inference is redundant, and by minimality of  $(C \cdot \theta)$  w.r.t.  $\succ_{R_*}$  we know that  $R_* \models \text{concl}(\iota)$ . Since  $R_* \not\models (D' \cdot \theta)$  this implies  $R_* \models (C \cdot \theta)$ , contradicting the assumption.

Case 2:  $C\theta = C'\theta \vee s\theta \approx s'\theta$  and  $s\theta \approx s'\theta$  is maximal in  $C\theta$ . By assumption,  $R_* \not\models s\theta \approx s'\theta$ , hence  $s\theta \downarrow_{R_*} \neq s'\theta \downarrow_{R_*}$ . W.l.o.g. let  $s\theta \succ s'\theta$ .

Case 2.1:  $s\theta$  is reducible by a rule  $t\theta \rightarrow t'\theta \in R_*$ , which has been produced by a closure  $(D \cdot \theta) = (D' \vee t \approx t' \cdot \theta)$  in  $N$ . By Lemmas 18 and 19, we obtain  $(D \cdot \theta) \prec_{R_*} (C \cdot \theta)$ , and, provided that  $t\theta$  occurs in  $s\theta$  below the top, also  $D\theta \prec_C C\theta$ . Therefore there is a PARALLEL SUPERPOSITION (I or II) inference  $\iota$  with left premise  $(D \cdot \theta)$  and right premise  $(C \cdot \theta)$ , and we can derive a contradiction analogously to Case 1.2.

Case 2.2: It remains to consider the case that  $s\theta$  is irreducible by  $R_*$ . Then  $s\theta$  is also irreducible by  $R_{s\theta}$ . Furthermore, by Lemma 17,  $\succ_{R_{s\theta}}$  and  $\succ_{R_*}$  agree on all closures in which  $s\theta$  is a strictly maximal term and occurs only positively. Therefore  $(C \cdot \theta)$  satisfies all conditions for productivity, hence  $R_* \models (C \cdot \theta)$ , contradicting the assumption.  $\square$

### 4.3 Lifting

It remains to lift the refutational completeness result for ground closure Horn superposition to the non-ground case.

If  $C$  is a general clause, we call every ground closure  $(C \cdot \theta)$  a ground instance of  $C$ . If

$$\frac{C_n \dots C_1}{C_0}$$

is a general inference and

$$\frac{(C_n \cdot \theta) \dots (C_1 \cdot \theta)}{(C_0 \cdot \theta)}$$

is a ground inference, we call the latter a ground instance of the former. The function  $\mathcal{G}$  maps every general clause  $C$  and every general inference  $\iota$  to the set of its ground instances. We extend  $\mathcal{G}$  to sets of clauses  $N$  or sets of inferences  $I$  by defining  $\mathcal{G}(N) := \bigcup_{C \in N} \mathcal{G}(C)$  and  $\mathcal{G}(I) := \bigcup_{\iota \in I} \mathcal{G}(\iota)$ .

**Lemma 21.**  $\mathcal{G}$  is a grounding function, that is, (1)  $\mathcal{G}(\perp) = \{\perp\}$ ; (2) if  $\perp \in \mathcal{G}(C)$  then  $C = \perp$ ; and (3) for every inference  $\iota$ ,  $\mathcal{G}(\iota) \subseteq \text{Red}_1(\mathcal{G}(\text{concl}(\iota)))$ .

The grounding function  $\mathcal{G}$  induces a lifted redundancy criterion  $(\text{Red}_1^{\mathcal{G}}, \text{Red}_C^{\mathcal{G}})$  where  $\iota \in \text{Red}_1^{\mathcal{G}}(N)$  if and only if  $\mathcal{G}(\iota) \subseteq \text{Red}_1(\mathcal{G}(N))$  and  $C \in \text{Red}_C^{\mathcal{G}}(N)$  if and only if  $\mathcal{G}(C) \subseteq \text{Red}_C(\mathcal{G}(N))$ .

**Lemma 22.** Every ground inference from closures in  $\mathcal{G}(N)$  is a ground instance of an inference from  $N$  or contained in  $\text{Red}_1(\mathcal{G}(N))$ .

*Proof.* Let  $\iota$  be a ground inference from closures in  $\mathcal{G}(N)$ . If  $\iota$  is a PARALLEL SUPERPOSITION I or EQUALITY RESOLUTION inference, then it is a ground instance of a PARALLEL SUPERPOSITION or EQUALITY RESOLUTION with premises in  $N$ . It remains to consider PARALLEL SUPERPOSITION II inferences

$$\frac{(D' \vee t \approx t' \cdot \theta) \quad (C \cdot \theta)}{(D' \vee C \cdot \theta[x \mapsto u[t'\theta]])}$$

with  $x\theta = u[t\theta]$ . Let  $R$  be a left-reduced ground rewrite system contained in  $\succ$ . If  $(t\theta \rightarrow t'\theta) \notin R$ , then  $\iota$  satisfies case (iii) of Def. 12. Otherwise  $(C \cdot \theta[x \mapsto u[t'\theta]])$  is a ground instance of the clause  $C \in N$  and  $\succ_R$ -smaller than the premise  $(C \cdot \theta)$ . If  $R \models (C \cdot \theta[x \mapsto u[t'\theta]])$ , then  $R \models \text{concl}(\iota)$ , so  $\iota$  satisfies case (i) of Def. 12; otherwise it satisfies case (ii) of Def. 12.  $\square$

**Theorem 23.** The Horn Superposition Calculus (using PARALLEL SUPERPOSITION) together with the lifted redundancy criterion  $(\text{Red}_1^{\mathcal{G}}, \text{Red}_C^{\mathcal{G}})$  is refutationally complete.

*Proof.* This follows immediately from Lemma 22 and Thm. 32 from (Waldmann et al. [8]).  $\square$

#### 4.4 Deletion and Simplification

It turns out that DER, as well as most concrete deletion and simplification techniques that are implemented in state-of-the-art superposition provers are in fact covered by our abstract redundancy criterion. There are some unfortunate exceptions, however.

**DER.** DESTRUCTIVE EQUALITY RESOLUTION, that is, the replacement of a clause  $x \not\approx t \vee C$  with  $x \notin \text{vars}(t)$  by  $C\{x \mapsto t\}$  is covered by the redundancy criterion. To see this, consider an arbitrary ground instance  $(x \not\approx t \vee C \cdot \theta)$  of  $x \not\approx t \vee C$ . Let  $R$  be a left-reduced ground rewrite system contained in  $\succ$ . Assume that the instance is false in  $R$ . Then  $x\theta$  and  $t\theta$  have the same  $R$ -normal form. Consequently, any redex or normal form in  $\text{nm}_R(C\{x \mapsto t\} \cdot \theta)$  was already present in  $\text{nm}_R(x \not\approx t \vee C \cdot \theta)$  (possibly with a larger label, if  $x$  occurs only positively in  $C$ ). Moreover, the labeled normal form  $(x\theta \downarrow_R, 2)$  that is present in  $\text{nm}_R(x \not\approx t \vee C \cdot \theta)$  is missing in  $\text{nm}_R(C\{x \mapsto t\} \cdot \theta)$ . Therefore  $(x \not\approx t \vee C \cdot \theta) \succ_R (C\{x \mapsto t\} \cdot \theta)$ . Besides, both closures have clearly the same truth value in  $R$ , that is, false.

**Subsumption.** Propositional subsumption, that is, the deletion of a clause  $C \vee D$  with nonempty  $D$  in the presence of a clause  $C$  is covered by the redundancy criterion. This follows from the fact that every ground instance  $((C \vee D) \cdot \theta)$  of the deleted clause is entailed by a smaller ground instance  $(C \cdot \theta)$  of the subsuming clause. This extends to all simplifications that replace a clause by a subsuming clause in the presence of certain other clauses, for instance the replacement of a clause  $t\sigma \approx t'\sigma \vee C$  by  $C$  in the presence of a clause  $t \not\approx t'$ , or the replacement of a clause  $u[t\sigma] \not\approx u[t'\sigma] \vee C$  by  $C$  in the presence of a clause  $t \approx t'$ .

First-order subsumption, that is, the deletion of a clause  $C\sigma \vee D$  in the presence of a clause  $C$  is not covered, however. This is due to the fact that  $\succ_R$  makes the instance  $C\sigma$  smaller than  $C$ , rather than larger (see Lemma 7).

**Tautology Deletion.** The deletion of (semantic or syntactic) tautologies is obviously covered by the redundancy criterion.

**Parallel Rewriting with Condition Literals.** Parallel rewriting with condition literals, that is, the replacement of a clause  $t \not\approx t' \vee C[t, \dots, t]_{p_1, \dots, p_k}$ , where  $t \succ t'$  and  $p_1, \dots, p_k$  are all the occurrences of  $t$  in  $C$  by  $t \not\approx t' \vee C[t', \dots, t]_{p_1, \dots, p_k}$  is covered by the redundancy criterion. This can be shown analogously as for DER.

**Demodulation.** Parallel demodulation is the replacement of a clause  $C[t\sigma, \dots, t\sigma]_{p_1, \dots, p_k}$  by  $C[t'\sigma, \dots, t'\sigma]_{p_1, \dots, p_k}$  in the presence of another clause  $t \approx t'$  where  $t\sigma \succ t'\sigma$ . In general, this is *not* covered by our redundancy criterion. For instance, if  $\succ$  is a KBO where all symbols have weight 1 and if  $R = \{f(b) \rightarrow b, g(g(b)) \rightarrow b\}$ , then replacing  $f(f(f(b)))$  by  $g(g(b))$  in some clause  $f(f(f(b))) \not\approx c$  yields a clause with a larger  $R$ -normalization multiset, since the labeled redexes  $\{(f(b), 2), (f(b), 2), (f(b), 2)\}$  are replaced by  $\{(g(g(b)), 2)\}$  and  $g(g(b)) \succ f(b)$ .

A special case is supported, though: If  $t'\sigma$  is a proper subterm of  $t\sigma$ , then the  $R$ -normalization multiset either remains the same or becomes smaller, since every redex in the normalization of  $t'\sigma$  occurs also in the normalization of  $t\sigma$ .

## 5 Completeness, Part II: The Non-Horn Case

In the non-Horn case, the construction that we have seen in the previous section fails for (PARALLEL) SUPERPOSITION inferences at the top of positive literals. Take an LPO with precedence  $f \succ c_6 \succ c_5 \succ c_4 \succ c_3 \succ c_2 \succ c_1 \succ b$  and consider the ground closures  $(f(x_1) \approx c_1 \vee f(x_2) \approx c_2 \vee f(x_3) \approx c_3 \cdot \theta)$  and  $(f(x_4) \approx c_4 \vee f(x_5) \approx c_5 \vee f(x_6) \approx c_6 \cdot \theta)$ , where  $\theta$  maps all variables to the same constant  $b$ . Assume that the first closure produces the rewrite rule  $(f(b) \approx c_3) \in R_*$ . The  $R_*$ -normalization multisets of both closures are dominated by three occurrences of the labeled redex  $(f(b), 0)$ . However, a SUPERPOSITION inference between the closures yields  $(f(x_1) \approx c_1 \vee f(x_2) \approx c_2 \vee f(x_4) \approx c_4 \vee f(x_5) \approx c_5 \vee c_3 \approx c_6 \cdot \theta)$ ,



whose  $R_*$ -normalization multiset contains four occurrences of the labeled redex  $(f(b), 0)$ , hence the conclusion of the inference is larger than both premises. If we want to change this, we must ensure that the weight of positive literals depends primarily on their larger sides, and if the larger sides are equal, on their smaller sides. That means that in the non-Horn case, the clause ordering must treat positive literals as the traditional clause ordering  $\succ_C$ . But that has two important consequences: First, DER may no longer be used to eliminate variables that occur also in positive literals (since DER might now increase the weight of these literals). On the other hand, unrestricted demodulation becomes possible for positive literals.

We sketch the key differences between the non-Horn and the Horn case; for the details, we refer to the technical report [7].

We define the subterm set  $ss^-(C)$  and the topterm set  $ts^-(C)$  of a clause  $C$  as in the Horn case:

$$\begin{aligned} ss^-(C) &= \{t \mid C = C' \vee s[t]_p \not\approx s'\} \\ ts^-(C) &= \{t \mid C = C' \vee t \not\approx t'\} \end{aligned}$$

We do not need labels anymore. Instead, for every redex or normal form  $u$  that appears in negative literals we include the two-element multiset  $\{u, u\}$  in the  $R$ -normalization multiset to ensure that a redex or normal form  $u$  in negative literals has a larger weight than a positive literal  $u \approx v$  with  $u \succ v$ . We define the  $R$ -redex multiset  $rm_R(t)$  of a ground term  $t$  by

$$\begin{aligned} rm_R(t) &= \emptyset \text{ if } t \text{ is } R\text{-irreducible;} \\ rm_R(t) &= \{\{u, u\}\} \cup rm_R(t') \text{ if } t \rightarrow_R t' \text{ using the rule } u \rightarrow v \in R. \end{aligned}$$

The  $R$ -normalization multiset  $nm_R(C \cdot \theta)$  of a ground closure  $(C \cdot \theta)$  is

$$\begin{aligned} nm_R(C \cdot \theta) &= \bigcup_{f(t_1, \dots, t_n) \in ss^-(C)} rm_R(f(t_1 \theta \downarrow_R, \dots, t_n \theta \downarrow_R)) \\ &\quad \cup \bigcup_{x \in ss^-(C)} rm_R(x\theta) \\ &\quad \cup \bigcup_{t \in ts^-(C)} \{\{t\theta \downarrow_R, t\theta \downarrow_R\}\} \\ &\quad \cup \bigcup_{(s \approx s') \in C} \{\{s\theta, s'\theta\}\} \end{aligned}$$

Once more, the  $R$ -normalization closure ordering  $\succ_R$  compares ground closures  $(C \cdot \theta_1)$  and  $(D \cdot \theta_2)$  using a lexicographic combination of three orderings:

- first, the twofold multiset extension  $(\succ_{mul})_{mul}$  of the reduction ordering  $\succ$  applied to the multisets  $nm_R(C \cdot \theta_1)$  and  $nm_R(D \cdot \theta_2)$ ,
- second, the traditional clause ordering  $\succ_C$  applied to  $C\theta_1$  and  $D\theta_2$ ,
- and third, the same closure ordering  $\succ_{clo}$  as in the Horn case.

With this ordering, we can again prove Lemmas 9 and 10 and their analogue for EQUALITY FACTORING, which implies Lemma 14. In the construction of a candidate interpretation, we define again  $R_s = \bigcup_{t \succ_s} E_t$  for every ground term  $s$  and  $R_* = \bigcup_t E_t$ . We define  $E_s = \{s \rightarrow s'\}$ , if  $(C \cdot \theta)$  is the  $\succ_{R_s}$ -smallest closure

in  $N$  such that  $C = C' \vee u \approx u'$ ,  $u\theta \approx u'\theta$  is a strictly maximal literal in  $C\theta$ ,  $s = u\theta$ ,  $s' = u'\theta$ ,  $s \succ s'$ ,  $s$  is irreducible w.r.t.  $R_s$ ,  $C\theta$  is false in  $R_s$ , and  $C'\theta$  is false in  $R_s \cup \{s \rightarrow s'\}$ , provided that such a closure  $(C \cdot \theta)$  exists. If no such closure exists, we define  $E_s = \emptyset$ .

We can then reprove Thm. 20 for the non-Horn case. The only difference in the proof is one additional subcase before Case 2.1: If  $s\theta \approx s'\theta$  is maximal, but not strictly maximal in  $C\theta$ , or if  $C'\theta$  is true in  $R_{s\theta} \cup \{s\theta \rightarrow s'\theta\}$ , then there is an EQUALITY FACTORING inference with the premise  $(C \cdot \theta)$ . This inference must be redundant, which yields again a contradiction.

The lifting to non-ground clauses works as in Sect. 4.3.

## 6 Discussion

We have demonstrated that the naive addition of DESTRUCTIVE EQUALITY RESOLUTION (DER) to the standard abstract redundancy concept destroys the refutational completeness of the calculus, but that there exist restricted variants of the Superposition Calculus that are refutationally complete even with DER (restricted to negative literals in the non-Horn case). The key tool for the completeness proofs is a closure ordering that is structurally very different from the classical ones – it is not a multiset extension of some literal ordering – but that still has the property that the redundancy criterion induced by it is compatible with the Superposition Calculus.

Of course there is a big gap between the negative result and the positive results. The new redundancy criterion justifies DER as well as most deletion and simplification techniques found in realistic saturation provers, but only propositional subsumption and only a very restricted variant of demodulation. The question whether the Superposition Calculus is refutationally complete together with a redundancy criterion that justifies both DER (in full generality even in the non-Horn case) and *all* deletion and simplification techniques found in realistic saturation provers (including unrestricted demodulation and first-order subsumption) is still open. Our work is an intermediate step towards a solution to this problem. There may exist a more refined closure ordering that allows us to prove the completeness of such a calculus. On the other hand, if the combination is really incomplete, a counterexample must make use of those operations that our proof fails to handle, that is, DER in positive literals in non-Horn problems, first-order subsumption, or demodulation with unit equations that are contained in the usual term ordering  $\succ$  but yield closures that are larger w.r.t.  $\succ_{R_*}$ .

**Acknowledgments.** I thank the anonymous IJCAR reviewers for their helpful comments and Stephan Schulz for drawing my attention to the problem at the CASC dinner in 2013.

**Disclosure of Interests.** The author has no competing interests to declare that are relevant to the content of this article.

## References

1. Baader, F., Nipkow, T.: Term Rewriting and All That. Cambridge University Press, Cambridge, UK (1998)
2. Bachmair, L., Ganzinger, H.: Rewrite-based equational theorem proving with selection and simplification. *J. Log. Comput.* **4**(3), 217–247 (1994)
3. Bachmair, L., Ganzinger, H., Lynch, C., Snyder, W.: Basic paramodulation. *Information and Computation* **121**(2), 172–192 (Sep 1995)
4. Duarte, A., Korovin, K.: Ground joinability and connectedness in the superposition calculus. In: Blanchette, J., Kovács, L., Pattinson, D. (eds.) *Automated Reasoning, IJCAR 2022*. pp. 169–187. Springer (2022)
5. Nieuwenhuis, R., Rubio, A.: Theorem proving with ordering and equality constrained clauses. *J. Symb. Comput.* **19**(4), 321–351 (1995)
6. Schulz, S.: E—a brainiac theorem prover. *AI Commun.* **15**(2–3), 111–126 (2002)
7. Waldmann, U.: On the (in-)completeness of destructive equality resolution in the superposition calculus. Tech. rep., Max Planck Institute for Informatics (2024), <https://people.mpi-inf.mpg.de/~uwe/paper/Waldmann-DER-TR-2024.pdf>
8. Waldmann, U., Tourret, S., Robillard, S., Blanchette, J.C.: A comprehensive framework for saturation theorem proving. *Journal of Automated Reasoning* **66**, 499–539 (2022). <https://doi.org/10.1007/s10817-022-09621-7>