Superposition Theorem Proving for Commutative Algebraic Theories

Jürgen Stuber

Als Dissertation eingereicht
an der
Technischen Fakultät
der Universität des Saarlandes
im Dezember 1999

Contents

Abstract

We develop special superposition calculi for first-order theorem proving in the theories of abelian groups, commutative rings, and modules and commutative algebras over fields or over the ring of integers, in order to make automated theorem proving in these theories more effective. The calculi are refutationally complete on arbitrary sets of ground clauses, which in particular may contain additional function symbols. The calculi are derived systematically from a representation of the theory as a convergent term rewriting system. Compared to standard superposition they have stronger ordering restrictions so that inferences are applied only to maximal summands, and they contain macro inference rules that use theory axioms in a goal directed fashion. In general we need additional inferences to handle critical peaks between extended clauses. We show that these are not needed for abelian groups and modules, and that for commutative rings and commutative algebras one such inference suffices for any pair of ground clauses.

To facilitate the construction of term orderings for such calculi we introduce theory path orderings.

Zusammenfassung

Wir entwickeln in dieser Arbeit spezielle Superpositionskalküle für die Theorien der abelschen Gruppen, der kommutativen Ringe, und der Moduln und kommutativen Algebren über Körpern und den ganzen Zahlen, mit dem Ziel das automatische Theorembeweisen in Logik erster Stufe für diese Theorien effektiver zu machen. Die Kalküle sind widerlegungsvollständig für beliebige Mengen von Grundklauseln, in denen insbesondere auch beliebige, nicht in der Theorie auftretende, Funktionssymbole vorkommen dürfen. Die Kalküle entwickeln wir systematisch aus einer Darstellung der Theorien als konvergente Termersetzungssysteme. Im Vergleich zu Standardsuperposition haben sie stärkere Ordnungseinschränkungen, so daß Inferenzen nur noch auf maximale Summanden angewendet werden müssen, und sie enthalten Makroinferenzregeln, die Theorieaxiome in zielgerichteter Weise anwenden. Im allgemeinen benötigen wir weiterhin Inferenzen, um kritische Paare zwischen erweiterten Klauseln zu behandeln. Wir zeigen, daß diese für abelsche Gruppen und Moduln nicht nötig sind, und daß für kommutative Ringe und Algebren eine Inferenz für jedes Paar von Grundklauseln genügt.

Um die Konstruktion von Termordnungen für unsere Kalküle zu vereinfachen, führen wir den Begriff der Theoriepfadordnung ein.

Extended Abstract

Automated theorem provers face problems when they are used on theories whose axioms generate large search spaces. Overwhelmed by a huge number of trivial consequences of each fact, they fail to prove even rather simple theorems.

Our goal in this work is to improve the methods for superposition theorem proving (Bachmair and Ganzinger 1994c, Bachmair and Ganzinger 1998a) in the context of algebraic theories. We specifically choose abelian groups, commutative rings, and modules and commutative algebras over the ring of integers or over a field. These theories are important in many applications, for instance various kinds of numbers and also vector spaces fall within their scope. Also, they are well-behaved from an algebraic viewpoint, and they are difficult to handle for automated theorem provers. Their axioms, in particular associativity, commutativity, distributivity and the inverse law, lead to many permuted variants of essentially the same term or equation. Moreover, these theories are among the largest for which refutationally complete calculi have been built.

We develop calculi which are refutationally complete for arbitrary first-order formulas, without restrictions of the logical structure or the set of function symbols. Logically there is no difference between using one of our calculi and using some less specialized refutationally complete calculus together with the axioms that are not integrated in that calculus, because our theories are axiomatizable in first-order logic, except for the base rings of modules or commutative algebras. Computations in these base rings are formalized by constraints, or by usually infinite sets of ground instances.

We achieve calculi that are improved in several respects. First, we strengthen the ordering restrictions so that inferences apply only to a maximal summand. Second, we replace certain direct uses of axioms by macro inferences. Standard superpositions into the inverse law can move a summand from one side of an equation to the other. Instead of doing this in an unrestricted way, we introduce a macro inference called Isolation that isolates the maximal terms on one side. Other cases of standard superposition into theory equations are replaced by introducing semantic matching into the superposition rule. We formalize this by associating to each original equation an extended set of term rewriting rules, called its *symmetrization*, and using these implied rules in the superposition. In this way we avoid to explicitly add extended clauses. Nevertheless we have to consider critical pairs between the sets of extended rules of two symmetrizations. That is, in general we have to add inferences corresponding to these critical pairs to our calculi. By using our knowledge of the form of symmetrizations we show that no such inferences are needed in the cases of abelian groups and modules, and that a single inference for any pair of ground clauses suffices in the cases of commutative rings and algebras. The combination of these stronger ordering restrictions, macro inferences and redundancy criteria promises to be much more

efficient than a more general calculus applied to part of the axioms. For instance, in purely equational reasoning it has been demonstrated that special calculi can improve performance greatly (Zhang 1993, Marché 1996). For the case of ground equations over a finite set of constants as the set of free function symbols our calculi generate essentially the same inferences as the Gröbner base algorithms for the respective theories, for instance the Buchberger algorithm (Buchberger 1970) in the case of a commutative algebra over a field.

To avoid a separate completeness proof for each theory and to gain a better understanding of the general mechanism we have developed a framework that allows to derive superposition calculi systematically from convergent term rewriting systems for the theories. This framework consists of a parameterized superposition calculus, where the parameters are a term ordering, a simplification function and a symmetrization function. We assume certain properties of the parameters which allow to prove refutational completeness of the parameterized calculus. These properties are rather restrictive, hence the construction will not work for every theory presentable by a convergent term rewriting system. For the theories which we consider we define these parameters and show that they satisfy the required properties. In addition we use abelian monoids as a simple running example during the development of the general framework.

The role of a simplification function is to transform equations into a theory-specific normal form, in a way that is compatible with the notion of redundancy for the calculus. The symmetrization function formalizes the interaction between the theory and other equations by assigning to each equation in normal form a set of rewrite rules that has the confluence property of being *symmetrized*, or for our theories even *strongly symmetrized*. Symmetrization is needed in the general framework where it guarantees convergence of peaks with one of the rules from the theory and where it is used for matching in the superposition inference. The property of strong symmetrization is essential for manipulating equational proofs, as it allows to normalize the terms in the proof with respect to the theory. This is necessary in particular for commutative rings and commutative algebras, since due to the critical peaks between extensions transitivity holds only below a certain bound in these cases. This situation requires equality proofs whose terms stay within the given bound. Additional difficulties arise in the proofs that the isolation rules for these theories are compatible with the notion of redundancy, where we need to relax the bound slightly by considering only summands.

For the cases of commutative rings, modules and commutative algebras we need term orderings with properties that cannot be achieved by previously known standard orderings. To simplify the construction of such term orderings we develop the notion of a *theory path ordering* (Stuber 1999) that generalizes the basic idea of the associative path ordering (Bachmair and Plaisted 1985). In our construction we use techniques of Geser (1996) for general path orderings.

Ausführliche Zusammenfassung

Für automatische Theorembeweiser sind viele wichtige Theorien schwierig, da deren Axiomensysteme sehr große Suchräume aufspannen. Überwältigt von einer Vielzahl trivialer Varianten jeder hergeleiteten Formel können sie selbst für einfache Theoreme keinen Beweis finden. Dies betrifft vor allem Theorien aus der Algebra. Das Ziel dieser Arbeit ist es, diese Probleme für das Theorembeweisen mittels Superposition (Bachmair and Ganzinger 1994c, Bachmair and Ganzinger 1998a) zu verringern. Dabei betrachten wir insbesondere die Theorien der abelschen Gruppen, der kommutativen Ringe, und der Moduln und kommutativen Algebren über den ganzen Zahlen oder einem fest vorgegebenen Körper. Diese Theorien haben viele wichtige Anwendungen, unter anderem umfassen sie die meisten Zahlensysteme und die Vektorräume. Sie sind aus mathematischer Sicht recht gutartig, aber für nicht spezialisierte Theorembeweiser schwierig. Ihre Axiome, insbesondere Assoziativität, Kommutativität, Distributivität und das Inversenaxiom, können viele permutierte Varianten einer Gleichung erzeugen. Außerdem sind diese Theorien unter den größten, für die widerspruchsvollständige Kalküle entwickelt wurden.

Wir entwickeln Kalküle, die für beliebige Formeln der Logik erster Stufe widerspruchsvollständig sind, ohne Einschränkung der logischen Struktur oder der Menge der freien Funktionssymbole. Logisch besteht kein Unterschied zwischen der Verwendung unserer Kalküle und der Verwendung eines weniger spezialisierten widerspruchsvollständigen Kalküls zusammen mit den nicht in den Kalkül integrierten Axiomen, da sich unsere Theorien in Logik erster Stufe axiomatisieren lassen, mit Ausnahme der Grundringe der Moduln und Algebren. Operationen in Grundringen werden durch Constraints, oder durch normalerweise unendliche Mengen von Grundinstanzen formalisiert.

Wir erzielen Kalküle, die in verschiedener Hinsicht besser sind. Zum einen verschärfen wir die Ordnungseinschränkungen, so daß Inferenzen nur noch auf maximale Summanden angewandt werden können. Zum anderen ersetzten wir bestimmte direkte Anwendungen der Axiome durch Makroinferenzen. Standardsuperpositionsinferenzen in das Inversenaxiom hinein können Summanden von einer Seite einer Gleichung auf die andere bewegen. Anstatt dies in unkontrollierter Weise zu tun, führen wir eine Makroinferenz Isolation ein, die die maximalen Summanden auf einer Seite der Gleichung isoliert. Andere Standardsuperpositionsinferenzen in Theorieaxiome hinein werden durch semantisches Matching in der Superpositionsregel ersetzt. Formal ordnen wir jeder Gleichung eine Menge von Termersetzungsregeln zu, genannt die *Symmetrisierung* der Gleichung, und benutzen diese implizierten Regeln zur Superposition. Dadurch vermeiden wir das Hinzufügen erweiterter Klauseln. Als Konsequenz müssen wir jedoch kritische Paare zwischen den erweiterten Regeln zweier Symmetrisierungen betrachten, was im allgemeinen zu zusätzlichen Inferenzen in unseren Kalkülen führt. Wir können jedoch unsere Kenntnis der Form der Symmetrisierung ausnutzen. Wir zeigen, daß in den Fällen der abelschen Gruppen und der Moduln keine solchen Inferenzen nötig sind, und daß in den Fällen der kommutativen Ringe und der kommutativen Algebren eine Inferenz für jedes Paar von Grundklauseln genügt. Die Kombination von stärkeren Ordnungseinschränkungen, Makroinferenzen und Redundanzkriterien verspricht eine wesentliche Verbesserung gegenüber einem allgemeinerem Kalkül, der auf einen Teil der Axiome angewendet wird. Im Rahmen von reinem Gleichheitsbeweisern wurde dies bereits an Hand von Testproblemen gezeigt (Zhang 1993, Marché 1996). Für den Fall von reinen Grundgleichungen und einer endlichen Menge von Konstanten als freien Funktionssymbolen erhalten wir in unseren Kalkülen die gleichen Inferenzen, die auch der der Theorie entsprechende Algorithmus zur Berechnung von Gröbner-Basen macht, zum Beispiel der Buchberger-Algorithmus (Buchberger 1970) im Fall einer kommutativen Algebra über einem Körper.

Um einen separaten Vollständigkeitsbeweis für jede einzelne Theorie zu vermeiden, und um ein besseres Verständnis der zugrundeliegenden Mechanismen zu bekommen, haben wir einen allgemeinen Rahmen für die Integration von durch konvergente Termersetzungssysteme repräsentierten Theorien entwickelt. Dieser Rahmen besteht aus einem parametrisierten Superpositionskalkül, der als Parameter eine Terminierungsordnung, eine Simplifikationsfunktion und eine Symmetrisierungsfunktion hat. Wir fordern von den Parametern gewisse Eigenschaften, die einen allgemeinen Beweis der Refutationsvollständigkeit erlauben. Diese Eigenschaften sind recht restriktiv, daher funktioniert diese Konstruktion nicht für alle durch konvergent Termersetzungssysteme gegebenen Theorien. Für die von uns betrachteten Theorien geben wir geeignete Parameter an und zeigen ihre Eigenschaften. Außerdem verwenden wir abelsche Monoide als ein einfaches fortlaufendes Beispiel bei der Entwicklung des allgemeinen Rahmens.

Die Aufgabe der Simplifikationsfunktion ist es, Gleichungen in eine theoriespezifische Normalform zu transformieren, und zwar in einer mit dem Redundanzbegriff des Kalküls verträglichen Weise. Die Symmetrisierungsfunktion formalisiert die Interaktion zwischen Theorieaxiomen und anderen Gleichungen, indem sie jeder Gleichung in Normalform eine Menge von Termersetzungsregeln zuordnet, die *symmetrisiert* oder für unsere Theorien sogar *stark symmetrisiert* sind. Dabei wird Symmetrisierung für die Entwicklung des allgemeinen Rahmens gebraucht, wo sie die Konvergenz von kritischen Paaren mit der Theorie garantiert. Starke Symmetrisierung ist für die Manipulation von Gleichheitsbeweisen wichtig, da sie Normalisieren bezüglich der Theorie erlaubt. Dies ist insbesondere für kommutative Ringe und kommutative Algebren notwendig, da dort wegen der kritischen Paare zwischen Regelerweiterungen die Transitivität nur bis zu einer gewissen Schranke bezüglich der Termordnung angenommen werden kann. Diese Situation erfordert Gleichheitsbeweise, die nur Terme unterhalb der Schranke enthalten. Besonders schwierig sind für diese beiden Theorien die Beweise, daß Isolation mit dem Redundanzbegriff verträglich ist.

Für kommutative Ringe, Moduln und kommutative Algebren benötigen wir Termordnungen mit Eigenschaften, die sich nicht mit bekannten Standardordnungen erreichen lassen. Um die Konstruktion solcher Termordnungen zu vereinfachen entwickeln wir den Begriff der *Theoriepfadordnung* (Stuber 1999), die die Grundidee der assoziativen Pfadordnung (Bachmair and Plaisted 1985) verallgemeinert. Für unsere Konstruktion verwenden wir Beweistechniken von Geser (1996) für allgemeine Pfadordnungen.

Acknowledgments

I thank Harald Ganzinger for his patience and continued support, Uwe Waldmann and Christoph Meyer for their valuable comments on parts of this thesis, Hubert Baumeister, Fritz Eisenbrand, Patrick Maier and Georg Struth for interesting discussions, Sergei Vorobyev for answering my more esoteric questions, and the other members of the Logics and Computation group and all people at the MPI Saarbrücken for their support and for providing an excellent research environment.

1

Introduction

Automated theorem provers face problems when they are used on theories whose axioms generate large search spaces. Overwhelmed by a huge number of trivial consequences of each fact, they fail to prove even rather simple theorems. For instance, this is the case when resolution (Robinson 1965) is applied to equality problems. The problem can be mitigated by combining several applications of axioms into macro steps, and by avoiding redundancies in the search space. An instance of the macro step technique is the paramodulation rule (Robinson and Wos 1969), which replaces certain resolution inferences with equality axioms. Superposition (Bachmair and Ganzinger 1994c, Bachmair and Ganzinger 1998a) places strong ordering restrictions on paramodulation and uses the ordering to obtain powerful techniques for eliminating redundancies. It is the state-of-the-art method for automated first-order equality reasoning.

Our goal in this work is to further improve the methods for superposition theorem proving in the context of algebraic theories. We specifically choose abelian groups, commutative rings, and modules and commutative algebras over the ring of integers or over a field. These theories are important in many applications, for instance various kinds of numbers and also vector spaces fall within their scope. Also, they are well-behaved from an algebraic viewpoint, and they are difficult to handle for automated theorem provers. Their axioms, in particular associativity, commutativity, distributivity and the inverse law, lead to many permuted variants of essentially the same term or equation. They are among the largest theories for which refutationally complete calculi have been built.

We develop refutationally complete calculi for arbitrary first-order formulas, without restrictions of the logical structure or the set of function symbols. Logically there is no difference between using one of our calculi and using some existing refutationally complete calculus together with the axioms of the respective theory, as our theories are axiomatizable in first-order logic. Note however that this does not apply to the integers or the fields used as base rings for modules or commutative algebras. In these cases first-order reasoning is used only for the module or algebra, while we do not allow equations between elements of the base ring. Computations in the base rings are formalized by constraints, or by usually infinite sets of ground instances.

We achieve calculi that are improved in several respects. First, we strengthen the ordering restrictions so that inferences apply only to a maximal summand within the top-level sum. Second, we replace certain direct uses of axioms by macro inferences. Standard superpositions into the inverse law can move a summand from one side of an equation to the other. Instead of doing this in an unrestricted way, we introduce a macro inference called Isolation that isolates the maximal terms on one side. Other cases of standard superposition into theory equations are replaced by introducing semantic matching into the superposition rule. We formalize this by associating to each original equation an extended set of term rewriting rules, called its *symmetrization*. By implicitly using these extensions for semantic matching, we avoid to explicitly add the corresponding extended clauses. Nevertheless, we have to consider critical peaks between extended rules, which lead to inferences between the corresponding clauses. We show that these inferences are not needed in the cases of abelian groups and modules, and that a single inference for any pair of ground clauses suffices in the cases of commutative rings and algebras. The combination of stronger ordering restrictions, macro inferences and redundancy criteria promises to be much more efficient than a more general calculus applied to part of the axioms. For instance, in purely equational reasoning it has been demonstrated that special calculi can improve performance greatly (Zhang 1993, Marché 1996).

For the case of ground equations over a finite set of constants as the set of free function symbols our calculi generate essentially the same inferences as the Gröbner base algorithms for the respective theories, for instance the Buchberger algorithm (Buchberger 1970) in the case of a commutative algebra over a field.

To avoid a separate completeness proof for each theory and to gain a better understanding of the general mechanism we have developed a framework that allows to derive superposition calculi systematically from convergent term rewriting systems for the theories. This framework consists of a parameterized superposition calculus, where the parameters are a term ordering, a simplification function and a symmetrization function. We assume certain properties of the parameters which allow to prove refutational completeness of the parameterized calculus. These properties are rather restrictive, hence the construction will not work for every theory presentable by a convergent term rewriting system. For the theories which we consider we define these parameters and show that they satisfy the required properties. In addition we use abelian monoids as a simple running example during the development of the general framework.

For the cases of commutative rings, modules and commutative algebras we need term orderings that cannot be constructed solely from standard orderings. We develop the notion of a *theory path ordering* to simplify the construction of orderings suitable for our framework (Stuber 1999). Using a technique of Geser (1996), we can prove all properties except compatibility with contexts, which we achieve by making theory symbols minimal in the precedence. This approach is analogous to that of the associative path ordering (Bachmair and Plaisted 1985).

The role of a simplification function is to transform equations into normal form, in a way that is compatible with the notion of redundancy for the calculus. While the general notion is simple, we have to prove that for commutative rings and for commutative algebras the isolation of the maximal term on one side is indeed a simplification.

The symmetrization function captures the interaction between the theory and other equations. We have introduced general notions of symmetrization and strong symmetrization with respect to an arbitrary convergent term rewriting system. Symmetrization is needed in the general framework where it guarantees convergence of peaks with one of the rules from the theory and where it is used for matching in the superposition inference. The property of strong symmetrization is essential for manipulating equational proofs in particular theories. To this end we have shown that strong symmetrization implies semi-compatibility of normalized rewriting and that convergence of a term rewriting system $R \cup T$ modulo E is equivalent to convergence of T-normalized rewriting with R modulo E (Stuber 1997). Finding equational proofs whose terms are small in the term ordering is crucial, because due to critical peaks between extended rules we can only assume that a ground instance of transitivity holds when its middle term is below a certain bound. The bound is closely related to critical peaks between extended rules, i.e., peaks which are not redundant. For the case of abelian groups and modules no such peak exists, which implies that all instances of transitivity are valid. For the case of commutative rings and algebras there is at most one critical peak for any pair of rules. The terms at the top of these critical peaks are single summands. Hence divergence can only occur within summands, and transitivity holds if the greatest summand of the middle term obeys the bound. This allows to handle slightly greater terms, namely sums with more summands than the bound. Only the combination of these techniques allows to prove that the calculi for commutative rings and for commutative algebras over a fixed ring are complete. In particular this applies to the proofs that the isolation rules are compatible with the notion of redundancy of the calculus.

Compared with previous work (Bachmair and Ganzinger 1994a, Bachmair, Ganzinger and Stuber 1995) we have clarified the relation between the limited validity of transitivity, extensions of rules, inferences among them and redundancy of these inferences. We use the presentation of Bachmair and Ganzinger (1998a), where inference systems that reduce any minimal counterexample in a candidate model are shown to be refutationally complete. Originally only clauses derived from the input set can become counterexamples. By allowing instances of transitivity to be counterexamples and appropriately placing them in the clause ordering we get a uniform presentation, with uniform notions of redundancy and of the reduction property.

We also extend the notion of redundancy to be able to refer to the presence of certain rewrite rules in an interpretation. This is needed to exploit the structure of the symmetrization of the rules to prove redundancy of most extension peaks. To this end we introduce a notion of logical consequence based solely on candidate models, which is quite natural and allows to use the standard notion based on all models as a sufficient criterion for redundancy.

The structure of this work is as follows. Chapter 2 contains preliminaries, Chapter 3 presents the theory path ordering, in Chapter 4 we develop the general framework, in Chapters 5 to 8 we apply the framework to the theories from abelian groups to commutative algebras, in Chapter 9 we briefly discuss lifting, and in Chapter 10 we conclude and discuss some limitations and possible extensions of this work.

1.1 Related work

Our work builds on several strands of research, namely automated first-order theorem proving, term rewriting, and the theory of Gröbner bases.

In automated first-order theorem proving there has been a general trend to build larger and larger theories into the calculi. The first step was to build equality into resolution (Robinson 1965), using the paramodulation inference rule (Robinson and Wos 1969). Stickel (1985) introduced theory resolution to provide a general framework for combining theory reasoning with resolution. However, in its general form it allows too many inferences and is not usable without further restriction and refinement. The superposition calculus of Bachmair and Ganzinger (1994c) imposes strong ordering restrictions on paramodulation and allows strong notions of redundancy. Our work can be viewed as a systematic way to derive strong restrictions for equational theories. We use the refined proof technique that proves refutational completeness by showing that a calculus reduces minimal counterexamples (Bachmair and Ganzinger 1998a). Wertz (1992) builds superposition calculi for theorem proving modulo E, and in particular modulo AC. He uses equality interpretations where transitivity holds universally, but has to accept that E holds only below a certain bound. In contrast to this, Bachmair and Ganzinger (1994a) in their AC-superposition calculus sacrifice universal validity of transitivity to get universal validity of AC. In practice this is easier to handle, as AC-matching and AC-unification can be treated as black boxes. Transitivity in the limit is obtained by computing inferences that correspond to critical pairs between extended rules. Rubio (1994), Nieuwenhuis and Rubio (1994, 1997), and Vigneron (1994) consider superposition calculi modulo AC with constraints. Bachmair, Ganzinger and Stuber (1995) develop a calculus for commutative rings with a unit element. They build the calculus on top of the AC-superposition calculus (Bachmair and Ganzinger 1994a), showing that AC-superposition inferences with axioms become redundant if instead some inferences tailored to rings are made. The proof technique was not strong enough to avoid certain shortcomings, namely the explicit representation of the symmetrization and the weaker notion of redundancy. Superposition calculi for cancellative abelian monoids require a notion of rewriting on equations instead of terms, since additive inverses are in general not available (Ganzinger and Waldmann 1996, Waldmann 1997). The special case of divisible torsion-free abelian groups allows to eliminate unshielded variables, which avoids the most prolific inferences (Waldmann 1997, Waldmann 1998). Previously we have shown that our approach is compatible with constraints for the special case of integer modules (Stuber 1996, Stuber 1998a). We have also carried it out for commutative rings in the ground case (Stuber 1998b). Wang (1993) describes a special technique for reasoning in modules over the integers. His approach is restricted to proving Horn clauses, i.e., deducing one equation from a set of equations. He shows completeness only for the case without free function symbols.

Bachmair, Ganzinger and Waldmann (1994) give a superposition calculus for hierarchical specifications. In a hierarchical setting any model must interpret the theory symbols by an interpretation from a given set that formalizes the theory. Thus the interpretation of the free function symbols must be a conservative extension of the base interpretation. That is, free function symbols may not introduce new elements into base sorts ("no junk"), and no new equations between theory terms can become true ("no confusion"). These requirements severely restricts the applicability of this method. Since the interpretation of theory symbols is fixed, it is possible to use black-box decision methods whenever a problem falls entirely into the domain of the built-in theory. Boyer and Moore (1988) discuss the practical implications of such a hierarchical approach. There experiments show that this is too rarely the case to achieve a substantial speed-up. They propose a tighter integration of the theorem prover and the built-in theory. We view our work as a way to achieve this integration.

Term rewriting and Knuth-Bendix-completion (Knuth and Bendix 1970) are techniques for automated first-order theorem proving for the special case of unit equations. Peterson and Stickel (1981) give term rewriting systems modulo AC for various algebraic theories, introduce AC-extensions. We use their convergent term rewriting systems and termination orderings for abelian groups and commutative rings. Jouannaud and Kirchner (1986) give a general theory of term rewriting modulo equational theory. In particular they introduce the notion of coherence and present abstract criteria for a term rewriting system being Church-Rosser modulo an equational theory. Zhang (1993) considers alternative ring problems. He defines special completion procedures for abelian groups and distributivity which compute only a subset of the critical pairs between axioms and other rules, showing that other critical pairs are redundant. He reports that this procedure achieves good results for proving theorems from alternative ring theory. Marché (1996) builds a range of theories from AC to commutative rings into equational completion. For abelian groups what he calls symmetrization is our notion of T-normal form, while the first component of his normalizing pair corresponds to our notion of symmetrization. Symmetrizations are added to the set of rules explicitly. In contrast to our approach redundancy of certain inferences between extensions in the symmetrizations is not proved beforehand and hence not built into the inference system. Marché does not compute inferences below variables; in that case the equation would not be orientable and the completion would fail. Using the Cime system for completion with built-in theories (Contejean and Marché 1996), Marché demonstrates that the special treatment of theories can reduce the number of inferences greatly and can lead to large speedups.

The notion of symmetrization originates from string rewriting systems for finitely presented groups. There a group is represented as a finite set of generators and a finite set of relations of the form $w = 1$ where w is a word over the generators and their inverses. A *symmetrized* presentation of a relation consists of all cyclic permutations of the relation or its inverse. The notion is already present in the work of Dehn (1911), where it is used to simplify presentations. Dehn shows how to decide the word problem and isomorphism problem for some finitely presented groups related to topological problems. Greendlinger (1960a, 1960b) defines symmetrized set of relations explicitly and uses it to extend the results of Dehn to groups whose symmetrized sets of relations have only small overlaps. Le Chenadec (1986) generalizes this result to various other theories and shows that symmetrized presentations can be derived from the canonical term rewriting systems for the theories.

Another strand of research leading to this work is concerned with Gröbner or standard bases for polynomial simplification. The first algorithm to compute Gröbner bases is the Buchberger algorithm (Buchberger 1970, Buchberger 1984, Buchberger 1987). Originally only for multivariate polynomials over a field, it has been generalized to polynomials over other rings, for instance Euclidean rings by Kandri-Rody and Kapur (1988). For an overview see the book by Becker and Weispfenning (1993). More recently, Bachmair and Tiwari (1997) have even covered the case of commutative noetherian rings. To abstract from specific base rings Buchberger (1984) introduced the notion of a reduction ring and showed that standard bases can be computed in reduction rings and in the ring of multivariate polynomials over a reduction ring. Stifter (1987) has generalized this to rings with zero divisors and shown that various constructions preserve reduction rings (Stifter 1991). She has also considered modules over reduction rings (Stifter 1993). The computation of Gröbner bases generalizes both the Euclidean algorithm and Gaussian elimination (Buchberger 1987). The Euclidean algorithm also shows up explicitly in our work for the theories of modules and commutative algebras over the integers (see Section 7.6). Compared to Gröbner bases algorithms we are more restrictive with respect to the underlying rings, where we allow only fields and the ring of integers. The axioms of Euclidean rings or reduction rings are not strong enough to for our method. For general Euclidean rings it is not even clear whether a symmetrization exists. Reduction rings are almost strong enough, but fail to show that isolation is a simplification rule, because the axiom for transforming proofs (A5) does not apply due to its very restrictive preconditions. We also restrict ourselves to integral domains, because strong symmetrization cannot be achieved in the presence of zero divisors. Logically the techniques for finitely presented groups and polynomial rings correspond to the case of ground unit equations where the free function symbols are a finite set of constants, with various underlying theories. In this respect our approach generalizes them.

The relation between completion for term rewriting systems, which is the basis of our calculus, and Gröbner basis algorithms has already been noticed by Buchberger and Loos (1983) and Buchberger (1987). They remark that both decide equivalence with respect to a canonical term rewriting system or a Gröbner base by normalizing and comparing normal forms for equality, but apart from this their discussion of analogies is informal. Bündgen (1991, 1996) formalizes this by encoding Gröbner basis computation, including the computation in the base rings, in term rewriting systems. He introduces the notion of semi-compatibility and uses a similar technique to manipulate his proofs, but he does not make it explicit. Marché (1994, 1996) has used a similar approach based on his notion of normalized rewriting. Bachmair and Ganzinger (1994b) use constraints to express computations in the base ring, which separates them from computation in the base polynomial ring and abstracts from the details of the particular base ring. We use the same term rewriting system to represent the theory of commutative algebras.

Refutationally complete superposition calculi for algebraic theories require simplification orderings which are total on ground terms, and which obey additional restrictions imposed by the algebraic theories. In particular, theorem proving modulo some equational theory E requires that the term ordering is E-compatible. The most important theory in practice is AC (Bachmair and Plaisted 1985, Delor and Puel 1993, Rubio and Nieuwenhuis 1995, Kapur and Sivakumar 1997, Baader 1997, Rubio 1999). Additionally, presenting the theory by a term rewriting system modulo E (Bachmair, Ganzinger and Stuber 1995, Stuber 1998a, Stuber 1998b) requires that the ordering orients the rules in this system in the right direction.

Geser (1996) introduces the general path ordering to construct semantic path orderings. A general path ordering is built recursively from a status function as its only ingredient. The properties of the ordering are derived from corresponding properties of the status function. In its original form it is not suitable to derive E-compatible orderings, as the property of the status function needed for compatibility with contexts does not hold in the presence of flattening. Nevertheless, the other properties like transitivity and the subterm property carry over, and the method can easily be extended to other properties, like totality and E-antisymmetry. To present a such a status function we use a technique that is very similar to that of Baader (1997).

2

Preliminaries

2.1 Mathematical structures

Sets

We write tuples with angled brackets, like $\langle x_1, \ldots, x_n \rangle$.

Functions

Given a function $f : A \to B$ and $a \in A,\ \ b \in B$, we define $f[a/b]$ as the function which maps a to b and acts like f on $A \setminus \{a\}$.

Binary relations

A *binary relation* on a set M is a subset of $M \times M$. We write $x\ R\ y$ for $\langle x, y \rangle \in R$. The *composition* of two binary relations R and S, written $R \cdot S$, is defined as $\{\langle x, z \rangle \mid x\ R\ y \text{ and } y\ R\ z\}$. The *inverse* of R is $R^{-1} = \{\langle y, x \rangle \mid x\ R\ y\}$. For relation symbols such as $>$, $\succ$ and $\Rightarrow$ we use their mirror image to denote the inverse, i.e., $<$, $\prec$ and $\Leftarrow$, respectively. A binary relation R is called *reflexive* if $(=) \subseteq R$, *irreflexive* if $R \cap (=) = \emptyset$, *symmetric* if $R = R^{-1}$, *antisymmetric* if $R \cap R^{-1} = (=)$, and *transitive* if $R \cdot R \subseteq R$. The *transitive closure* R^+ is the smallest subset of M that contains R and is transitive. Analogously we define the *reflexive-transitive closure* R^* and the *symmetric closure* $R^{\Leftrightarrow}$. For arrow symbols such as $\to$ and $\Rightarrow$ the symmetric closure is denoted by $\leftrightarrow$ and $\Leftrightarrow$. respectively. Let S be a symmetric binary relation. R is called S*-compatible* if $S \cdot R \cdot S \subseteq R$.

We define the *product* $R_1 \times R_2$ of two binary relations R_1 and R_2 on M_1 and M_2, respectively, by $\langle x_1, x_2 \rangle\ (R_1 \times R_2)\ \langle y_1, y_2 \rangle$ if and only if $x_1\ R_1\ y_1$ and $x_2\ R_2\ y_2$. This is easily seen to be associative, hence the n-fold product $R_1 \times \cdots \times R_n$ is well defined.

2.2 Orderings

A *quasi-ordering* is a binary relation that is reflexive and transitive. A *(partial) ordering* is a quasi-ordering that is antisymmetric. A *strict (partial) ordering* is a binary relation that is irreflexive and transitive. An *equivalence* is a reflexive, symmetric, and transitive binary relation.

Any strict ordering $<$ can be extended to a nonstrict ordering $(\leq) = (<) \cup (=)$. We will usually not mention whether an ordering is strict or not, as this will be clear from the context and the relation symbol used. Each quasi-ordering $\geq$ can be split into its *strict part* $(>) = (\geq) \setminus (\leq)$ and its *equivalence kernel* $(\sim) = (\geq) \cap (\leq) = (\geq) \setminus (>)$. Then $\geq$

and $>$ obey the strict transitivity laws

$$x > y \text{ and } y \geq z \text{ implies } x > z, \text{ and}$$
$$x \geq y \text{ and } y > z \text{ implies } x > z.$$

On the other hand, if $\sim$ is an equivalence relation and $>$ is an $\sim$-compatible strict partial ordering then $(\geq) = (\sim) \cup (>)$ is a quasi-ordering. An equivalence relation S is contained in the equivalence kernel of any S-compatible quasi-ordering $\succeq$, because $x\ S\ y$ implies $x\ S\ y \succeq y\ S\ y$.

In the context of some quasi-ordering $\succeq$ we always use $\succ$ for its strict part and $\sim$ for its equivalence kernel. If we want to be more formal, we will write $\succeq(\geq)$, $\succ(\geq)$ and $\sim(\geq)$ for $\geq$ itself, its strict part and its equivalence kernel, respectively. $\succeq(\geq)$ allows to use infix notation for quasi-orderings constructed as the product or intersection of other quasi-orderings.

Moreover, a quasi-ordering $\succeq$ corresponds to a partial ordering $\geq$ on $\sim$-equivalence classes, where $x \succeq y$ if and only if $[x]_\sim \geq [y]_\sim$. For $\geq$ we observe that it is well-defined by transitivity of $\succeq$, that it is reflexive and transitive because $\succeq$ is, and that antisymmetry follows because $[x]_\sim = [y]_\sim$ whenever $x \sim y$. If, on the other hand, $\geq$ and some equivalence relation $\sim$ are used to define $\succeq$, then $\succeq$ is reflexive and transitive because $\geq$ is.

A binary relation R on M is called *total* if $R \cup R^{-1} = M^2$. A strict ordering $<$ is called *total* if its reflexive closure $(<) \cup (=)$ is total. A binary relation R is called *S-antisymmetric* if $R \cap R^{-1} \subseteq S$ and *total up to S* if $R \cup R^{-1} \cup S = M^2$. Usually R is an ordering and S is an equivalence. Note that if $S_1 \subseteq S_2$ then S_1-antisymmetry implies S_2-antisymmetry, and that a binary relation on M is always M^2-antisymmetric. x is a *minimal* element of a binary relation $>$ if there exists no y in M such that $x > y$. We call x the *smallest* element of $>$ if $y \geq x$ for all y in M. A binary relation $>$ is called *terminating* if there is no infinite descending chain $x_1 > x_2 > \cdots$ of elements in M. A binary relation $>$ is *well-founded* if each nonempty subset of M has a minimal element with respect to $>$. Using the axiom of choice one can show the equivalence of termination and well-foundedness. A *well-ordering* is a total ordering whose strict part is well-founded. In a well-ordering every nonempty subset of M has a smallest element.

A subset S of M is *downward-closed* with respect to $>$ if whenever x is in S and $x > y$ then also y is in S.

Lexicographic product

Let $\succeq_1$ and $\succeq_2$ be quasi-orderings on M_1 and M_2, respectively. The *lexicographic product* $\succeq_1 \times_{lex} \succeq_2$ of $\succeq_1$ and $\succeq_2$ is the binary relation on $M_1 \times M_2$ that is defined by

$$\langle x_1, x_2 \rangle \succeq (\succeq_1 \times_{lex} \succeq_2)\ \langle y_1, y_2 \rangle$$

if either $x_1 \succ_1 y_1$, or $x_1 \sim_1 y_1$ and $x_2 \succeq_2 y_2$. It is easily seen that $\times_{lex}$ is associative, hence the lexicographic product $\succeq_1 \times_{lex} \cdots \times_{lex} \succeq_n$ of n quasi-orderings $\succeq_1, \ldots, \succeq_n$ is well-defined.

Proposition 2.1 *Let $\succeq_1, \ldots, \succeq_n$ be quasi-orderings and let $\succeq_{lex} = \succeq_1 \times_{lex} \cdots \times_{lex} \succeq_n$.*

1. *$\succeq_{lex}$ is a quasi-ordering.*

2. *If there exists an infinite descending chain*

$$\langle x_{11}, \ldots, x_{1n} \rangle \succ_{lex} \langle x_{21}, \ldots, x_{2n} \rangle \succ_{lex} \cdots$$

then there exists an infinite descending subchain

$$x_{j_1 i} \succ_i x_{j_2 i} \succ_i \dots$$

for some $i = 1, \dots, n$ *and* $1 \le j_1 < j_2 < \dots$.

3. *If* $\succ_1, \dots, \succ_n$ *are well-founded then* $\succ_{lex}$ *is well-founded.*
4. *If* $\succeq_1, \dots, \succeq_n$ *are total then* $\succeq_{lex}$ *is total.*
5. *Let* $S_1, \dots, S_n$ *be equivalence relations and* $S = S_1 \times \cdots \times S_n$. *If* $\succeq_i$ *is* S_i*-compatible for* $i = 1, \dots, n$ *then* $\succeq_{lex}$ *is* S*-compatible.*
6. *Let* $S_1, \dots, S_n$ *be equivalence relations and* $S = S_1 \times \cdots \times S_n$. *If* $\succeq_i$ *is* S_i*-antisymmetric for* $i = 1, \dots, n$ *then* $\succeq_{lex}$ *is* S*-antisymmetric.*

Proof: We only prove part of these properties for $n = 2$. For $n > 2$ the n-fold lexicographic product can be considered as the iterated product of two quasi-orderings. Hence properties which are preserved for $n = 2$ are also preserved for $n > 2$.

(2) Suppose there exists an infinite descending chain

$$\langle x_{11}, x_{21} \rangle \succ_{lex} \langle x_{12}, x_{22} \rangle \succ_{lex} \dots .$$

Either there exists an infinite descending chain $x_{1j_1} \succ_1 x_{1j_2} \succ_1 \dots$ in the first component. Or $x_j \sim_1 x_{j+1} \sim_1 \dots$ from some $j \ge 0$ on, and there exists an infinite descending chain $x_{2j_1} \succ_2 x_{2j_2} \succ_2 \dots$ in the second component, for some $j \le j_1 < j_2 < \dots$.

(3) This is the contrapositive of (2).

(5) Let S_1, S_2 be equivalence relations and $S = S_1 \times S_2$, and suppose that $\succeq_i$ is S_i-compatible for $i = 1, 2$. Consider tuples $\langle x_1, x_2 \rangle$, $\langle x_1', x_2' \rangle$, $\langle y_1, y_2 \rangle$ and $\langle y_1', y_2' \rangle$ such that

$$\langle x_1, x_2 \rangle \; S \; \langle x_1', x_2' \rangle \succeq_{lex} \langle y_1', y_2' \rangle \; S \; \langle y_1, y_2 \rangle.$$

Then $x_1 \; S_1 \; x_1' \succeq_1 y_1' \; S_1 \; y_1$, and by S_1-compatibility of $\succeq_1$ also $x_1 \succeq_1 y_1$. If $x_1 \succ_1 y_1$ then also $\langle x_1, x_2 \rangle \succeq_{lex} \langle y_1, y_2 \rangle$. Otherwise $x_1 \sim_1 y_1$. By S_1-compatibility of $\succeq_1$ and symmetry of S_1 this implies $x_1' \sim_1 y_1'$ and hence $x_2' \succeq_2 y_2'$. By S_2-compatibility of $\succeq_2$ this implies $x_2 \succeq_2 y_2$ and in turn $\langle x_1, x_2 \rangle \succeq_{lex} \langle y_1, y_2 \rangle$.

(6) $\langle x_1, x_2 \rangle \sim_{lex} \langle y_1, y_2 \rangle$ implies $x_i \sim_i y_i$ for $i = 1, 2$. Since $\succeq_i$ is S_i-antisymmetric, $x_i \sim_i y_i$ implies $x_i \; S_i \; y_i$ for $i = 1, 2$. Hence $\langle x_1, x_2 \rangle \; S_1 \times S_2 \; \langle y_1, y_2 \rangle$. □

Lexicographic combination

Let $\succeq_1, \dots, \succeq_n$ be quasi-orderings on M. Their *lexicographic combination* $\succeq$ is a binary relation on M defined by $s \succeq t$ if and only if

$$\langle s, \dots, s \rangle \succeq (\succeq_1 \times_{lex} \dots \times_{lex} \succeq_n) \; \langle t, \dots, t \rangle.$$

Proposition 2.2 *Let* $\succeq_1, \dots, \succeq_n$ *be quasi-orderings and let* $\succeq$ *be the lexicographic combination of* $\succeq_1, \dots, \succeq_n$.

1. $\succeq$ *is a quasi-ordering.*

2. *If there exists an infinite descending chain*

$$x_1 \succ x_2 \succ \dots$$

then there exists an infinite descending subchain

$$x_{j_1} \succ_i x_{j_2} \succ_i \dots$$

for some $1 \leq j_1 < j_2 < \dots$.

3. *If* $\succ_1, \dots, \succ_n$ *are well-founded then* $\succ$ *is well-founded.*

4. *If* $\succeq_1, \dots, \succeq_n$ *are total then* $\succeq$ *is total.*

5. *Let* $S_1 \supseteq \dots \supseteq S_n$ *be equivalence relations. If* $\succeq_i$ *is* S_i*-compatible for* $i = 1, \dots, n$ *then* $\succeq$ *is* S_n*-compatible.*

6. *Let* $S_1, \dots, S_n$ *be equivalence relations and* $S = \bigcap_{i=1}^n S_i$. *If* $\succeq_i$ *is* S_i*-antisymmetric for* $i = 1, \dots, n$ *then* $\succeq$ *is* S*-antisymmetric.*

Proof: These properties follow from the properties of the lexicographic product. □

Note that S-antisymmetry of the lexicographic combination holds if the last component is S-antisymmetric, since the other components are always M^2-antisymmetric.

The length-lexicographic extension

The set of *tuples* (or words) over a set M is $M^* = \bigcup_{n \geq 0} M^n$. We write $\bar{x}$ for a tuple $\langle x_1, \dots, x_n \rangle$. The length $|\langle x_1, \dots, x_n \rangle|$ of $\langle x_1, \dots, x_n \rangle$ is n. The *length ordering* $\succeq_{len}$ on tuples is defined by $\bar{x} \succeq_{len} \bar{y}$ if and only if $|\bar{x}| \geq |\bar{y}|$.

Let $\succeq$ be a quasi-ordering on M. The *lexicographic extension* $\succeq_{lex}^n$ of $\succeq$ to n-tuples is the n-fold lexicographic product $\succeq \times_{lex} \cdots \times_{lex} \succeq$. The *lexicographic extension* $\succeq_{lex}(\succeq)$ of $\succeq$ to M^* is $\bigcup_{n \geq 0} \succeq_{lex}^n$. Note that this makes the lexicographic extension partial, as tuples of different length are not comparable. We prefer this definition, because the natural total lexicographic extension does not preserve well-foundedness. Also, in most applications only tuples of the same length are compared, e.g. in the lexicographic path ordering. The *length-lexicographic extension* of $\succeq$ to M^*, written $\succeq_{llex}(\succeq)$, is defined as the lexicographic combination of the length ordering $\succeq_{len}$ on M^* and the lexicographic extension of $\succeq$ to M^*. We denote the *extension* of a binary relation S on M to tuples of length n over M by $S^{\langle n \rangle}$, and the extension to tuples of arbitrary length by $S^{\langle * \rangle} = \bigcup_{n \geq 0} R^{\langle n \rangle}$.

Proposition 2.3 *Let* $\succeq$ *be a quasi-ordering on* M.

1. $\succeq_{llex}(\succeq)$ *is a quasi-ordering.*

2. *If there exists an infinite descending chain*

$$\bar{x}_1 \succ_{llex}(\succeq) \bar{x}_2 \succ_{llex}(\succeq) \dots$$

then there exists an infinite descending subchain

$$x_{ij_1} \succ x_{ij_2} \succ \dots$$

for some $1 \leq j_1 < j_2 < \dots$.

3. *If $\succ$ is well-founded then $\succ_{llex}(\succeq)$ is well-founded.*

4. *If $\succeq$ is total then $\succeq_{llex}(\succeq)$ is total.*

5. *Let S be an equivalence relation. If $\succeq$ is S-compatible then $\succeq_{llex}(\succeq)$ is $S^{\langle * \rangle}$-compatible.*

6. *Let S be an equivalence relation. If $\succeq$ is S-antisymmetric then $\succeq_{llex}(\succeq)$ is $S^{\langle * \rangle}$-antisymmetric.*

Proof: (5) The length ordering is $S^{\langle * \rangle}$-compatible, since only tuples of the same length can be related in $S^{\langle * \rangle}$. If $\succeq$ is S-compatible then the lexicographic extension is $S^{\langle * \rangle}$-compatible, and lexicographic combination preserves $S^{\langle * \rangle}$-compatibility. Hence the result is $S^{\langle * \rangle}$-compatible.

(6) $\bar{x} \sim_{llex} \bar{y}$ implies that $\bar{x}$ and $\bar{y}$ have the same length and are equivalent in the lexicographic extension. Then $x_i \sim y_i$ for $i = 1, \ldots, n$, which implies $x_i \; S \; y_i$ by S-antisymmetry of S, and hence $\bar{x} \; S^{\langle * \rangle} \; \bar{y}$. □

The multiset extension

The multiset extension was introduced by Dershowitz and Manna (1979). Jouannaud and Lescanne (1982) show that other natural definitions are equivalent.

A (finite) *multiset* M over a set S is a function from S into the natural numbers such that $M(x) > 0$ only for finitely many x in S. We denote the set of multisets over S by $\mathbb{N}^S_{\mathrm{fin}}$. For each x in S, $M(x)$ denotes the number of occurrences of x in M. Multisets can be written by enumerating their elements. E.g., we write $\{1, 0, 0\}$ for the multiset M over $\mathbb{N}$ with $M(0) = 2$, $M(1) = 1$ and $M(x) = 0$ for $x > 1$. We say that x is an element of M if $M(x) > 0$. The *union* and *difference* of multisets M and N are defined by $(M \cup N)(x) = M(x) + N(x)$ and $(M \setminus N)(x) = \max(M(x) - N(x), 0)$.

Proposition 2.4 *$\{x_1, \ldots, x_m\} = \{y_1, \ldots, y_n\}$ if and only if $m = n$ and there exists a permutation π such that $x_i = y_{\pi(i)}$ for $i = 1, \ldots, m$.*

The *multiset extension* $\succ_{mul}$ of a strict partial ordering $\succ$ is the strict partial ordering on multisets over S that is defined by $M \succ_{mul} N$ if and only if $M \neq N$ and for all x in S such that $N(x) > M(x)$ there exists an y in S such that $y \succ x$ and $M(y) > N(y)$.

Proposition 2.5 *Let $\succ$ be a strict partial ordering on S.*

1. *$\succ_{mul}$ is a strict partial ordering.*

2. *If there exists an infinite descending chain*

$$M_1 \succ_{mul} M_2 \succ_{mul} \cdots$$

of multisets then there exists an infinite descending chain

$$x_1 \succ x_2 \succ \cdots$$

in S and indices $1 \leq j_1 < j_2 < \ldots$ such that $x_i \in M_{j_i}$ for $i \geq 1$.

3. *If $\succ$ is well-founded then $\succ_{mul}$ is well-founded.*

4. *If $\succeq$ is total then $\succeq_{mul}$ is total.*

5. *Let M_1, M_2 and N be multisets over S. Then $M_1 \succ_{mul} M_2$ if and only if $N \cup M_1 \succ_{mul} N \cup M_2$.*

6. *If $(\succ') \supseteq (\succ)$ is a strict partial ordering on S then $(\succ'_{mul}) \supseteq (\succ_{mul})$.*

Proof: (1) was shown by

(2) We extract this slightly stronger result from the proof of well-foundedness of Dershowitz and Manna (1979). Consider some infinite descending chain

$$M_1 \succ_{mul} M_2 \succ_{mul} \cdots$$

of multisets over S. We extend S by a new minimal element $\perp$ and use the descending chain to build an infinite tree whose nodes are labeled with elements from $S \cup \{\perp\}$. We label the root arbitrarily by $\perp$ and add a child for each element of M_1. Since $M_i \succ_{mul} M_{i+1}$, the set $N = M_i \setminus M_{i+1}$ is not empty, and for each y in $N' = M_{i+1} \setminus M_i$ there exists some x in N such that $x \succ y$. We partition N' into a family $(N'_x)_{x \in N}$ such that $x \succ y$ for all $y \in N'_x$. We add the elements of N'_x as children of x to the tree for each $x \in N$. If N'_x is empty we add the single child $\perp$. Since at each step at least one element is added, the tree is infinite. The multisets are finite, hence the tree has finite degree, and by König's Lemma it has an infinite branch. By our construction an infinite descending chain

$$x_1 \succ x_2 \succ \ldots$$

begins at a child of the root node, and there exist $1 \leq j_1 < j_2 < \ldots$ such that $x_i \in M_{j_i}$ for $i \geq 1$.

(3) follows from (2).

(4) Suppose $\succeq$ is total, and let M and N be distinct multisets. Let y be the greatest element of the symmetric difference $(M \setminus N) \cup (N \setminus M)$. Suppose $M(y) > N(y)$. Since $y \succ x$ for any element x such that $N(x) > M(x)$, we conclude $M \succ_{mul} N$. Otherwise $N(y) > M(y)$, and by the same argument $N \succ_{mul} M$.

(5) This follows since addition on natural numbers is monotonic.

(6) Since $y \succ x$ implies $y \succ' x$ for $x, y \in S$, $M \succ_{mul} N$ implies $M \succ'_{mul} N$ for M and N multisets over S. □

Next we consider the multiset extension of a quasi-ordering. Let $\succeq$ be a quasi-ordering on S. We extend $\sim$ to an equivalence relation $\sim_{mul}$ on multisets over S by $M \sim_{mul} N$ if and only if $\sum_{y \sim x} M(y) = \sum_{y \sim x} N(y)$ for any x in S. This justifies setting $[M](x) = \sum_{y \sim x} M(y)$. We may also consider $[M]$ as a multiset over $S/\sim$ by letting $[M]([x]) = [M](x)$. Furthermore, $\succeq$ induces a partial ordering $\succeq_{S/\sim}$ on $S/\sim$. We apply the multiset extension to the strict part $\succ_{S/\sim}$ of $\succeq_{S/\sim}$, and obtain a strict partial ordering on multisets of equivalence classes. Via the canonical homomorphism this induces a strict partial ordering $\succ_{mul}$ on multisets over S that is compatible with $\sim_{mul}$. By combining $\succ_{mul}$ and $\sim_{mul}$ we obtain the desired quasi-ordering $(\succeq_{mul}) = (\succ_{mul}) \cup (\sim_{mul})$ on multisets over S.

Proposition 2.6 *Let $\succeq$ be a quasi-ordering on S.*

1. *$\succeq_{mul}$ is a quasi-ordering.*

2. *If there exists an infinite descending chain*

$$M_1 \succ_{mul} M_2 \succ_{mul} \dots$$

then there exists an infinite descending subchain

$$x_1 \succ x_2 \succ \dots$$

for some $1 \leq j_1 < j_2 < \dots$ *such that* $x_i \in M_{j_i}$ *for* $i \geq 1$.

3. *If* $\succ$ *is well-founded then* $\succ_{mul}$ *is well-founded.*

4. *If* $\succeq$ *is total then* $\succeq_{mul}$ *is total.*

5. *Let* M_1, M_2 *and* N *be multisets over* S. *Then* $M_1 \succeq_{mul} M_2$ *if and only if* $N \cup M_1 \succeq_{mul} N \cup M_2$.

Analogous to the length-lexicographic extension there is also a size-multiset extension, where the size of the multiset is combined lexicographically with the multiset extension.

The size $|M|$ of a multiset M over S is $\sum_{s \in S} M(s)$. The *size ordering* $\succeq_{size}$ on multisets is defined by $M \succeq_{size} N$ if and only if $|M| \geq |N|$. The the *size-multiset extension* $\succeq_{smul}$ of a quasi-ordering $\succeq$ is the lexicographic combination of the size ordering and the multiset extension of $\succeq$.

Proposition 2.7 *Let* $\succeq$ *be a quasi-ordering on* S.

1. $\succeq_{smul}$ *is a quasi-ordering.*

2. *If there exists an infinite descending chain*

$$M_1 \succ_{smul} M_2 \succ_{smul} \dots$$

then there exists an infinite descending subchain

$$x_1 \succ x_2 \succ \dots$$

for some $1 \leq j_1 < j_2 < \dots$ *such that* $x_i \in M_{j_i}$ *for* $i \geq 1$.

3. *If* $\succ$ *is well-founded then* $\succ_{smul}$ *is well-founded.*

4. *If* $\succeq$ *is total then* $\succeq_{smul}$ *is total.*

5. *Let* M_1, M_2 *and* N *be multisets over* S. *Then* $M_1 \succeq_{smul} M_2$ *if and only if* $N \cup M_1 \succeq_{smul} N \cup M_2$.

2.3 Strictly monotonic functions

Here we show that certain uses of the constructions of the previous section preserve monotonicity. This will be useful when we prove compatibility with contexts of term orderings which are based on these constructions.

Let $\succeq_i$ be a quasi-ordering on S_i for $i = 1, 2$, and let f be a function from S_1 to S_2. The function f is called *monotonic* if $x \succeq_1 y$ implies $f(x) \succeq_2 f(y)$, and *strictly monotonic* if it is monotonic and $x \succ_1 y$ implies $f(x) \succ_2 f(y)$.

The extension of a function $f : S_1 \to S_2$ to a function $f_{mul} : \mathbb{N}_{\text{fin}}^{S_1} \to \mathbb{N}_{\text{fin}}^{S_2}$ is defined by $f_{mul}(M)(y) = \sum_{f(x)=y} M(x)$ for all x in S_1 and y in S_2. For functions $f, g : S \to \mathbb{N}_{\text{fin}}^S$ the function $f \cup g : S \to \mathbb{N}_{\text{fin}}^S$ is defined by $(f \cup g)(x) = f(x) \cup g(x)$.

Proposition 2.8 *1. Let $\succeq$ be a quasi-ordering on S. Then the identity on S is strictly monotonic.*

2. *The composition of strictly monotonic functions is strictly monotonic.*

3. *For any multiset N over S and any multiset extension $\succeq_{mul}$ the function $M \mapsto M \cup N$ from $\succeq_{mul}$ to $\succeq_{mul}$ is strictly monotonic.*

4. *If f is a strictly monotonic function from $\succeq$ to $\succeq'$ then the extension f_{mul} of f to multisets is a strictly monotonic function from $\succeq_{mul}$ to $\succeq'_{mul}$.*

5. *Let f and g be strictly monotonic functions from $\succeq$ to $\succeq'_{mul}$. Then $f \cup g$ is a strictly monotonic function from $\succeq$ to $\succeq'_{mul}$.*

6. *Let f_i be a strictly monotonic function from $\succeq_i$ to $\succeq'_i$ for $i = 1, \ldots, n$. Then*

$$\langle x_1, \ldots, x_n \rangle \mapsto \langle f_1(x_1), \ldots, f_n(x_n) \rangle$$

is a strictly monotonic function from $\succeq_1 \times_{lex} \cdots \times_{lex} \succeq_n$ to $\succeq'_1 \times_{lex} \cdots \times_{lex} \succeq'_n$.

7. *Let $f_{\langle x_1,\ldots,x_{i-1}\rangle}$ be a strictly monotonic function from $\succeq_i$ to $\succeq'_i$ for $i = 1, \ldots, n$ and $\langle x_1, \ldots, x_{i-1} \rangle$ in $S_1 \times \cdots \times S_{i-1}$ such that $x_1 \sim_1 y_1, \ldots, x_i \sim_i y_i$ implies*

$$f_{\langle x_1,\ldots,x_{i-1}\rangle}(x_i) \sim'_i f_{\langle y_1,\ldots,y_{i-1}\rangle}(y_i)$$

for any tuples $\langle x_1, \ldots, x_i \rangle$ and $\langle y_1, \ldots, y_i \rangle$ in $S_1 \times \cdots \times S_i$. Then

$$\langle x_1, \ldots, x_n \rangle \mapsto \langle f_{\langle\rangle}(x_1), f_{\langle x_1\rangle}(x_2), \ldots, f_{\langle x_1,\ldots,x_{n-1}\rangle}(x_n) \rangle$$

is a strictly monotonic function from $\succeq_1 \times_{lex} \cdots \times_{lex} \succeq_n$ to $\succeq'_1 \times_{lex} \cdots \times_{lex} \succeq'_n$.

Proof: (5) Suppose $x \succeq y$. Then by assumption $f(x) \succeq'_{mul} f(y)$ and $g(x) \succeq'_{mul} g(y)$, and

$$(f \cup g)(x) = f(x) \cup g(x) \succeq'_{mul} f(y) \cup g(x) \succeq'_{mul} f(y) \cup g(y)$$

by the context property of the multiset extension. Analogously for the strict part.

(6) is the special case of (7) where for each i the functions $f_{\langle x_1,\ldots,x_{i-1}\rangle}$ are the same for all values of $x_1, \ldots, x_{i-1}$.

(7) It suffices to consider the case $n = 2$. Let $f = \langle x_1, x_2 \rangle \mapsto \langle f_{\langle\rangle}(x_1), f_{\langle x_1\rangle}(x_2) \rangle$, $\succeq = \succeq_1 \times_{lex} \succeq_2$ and $\succeq' = \succeq'_1 \times_{lex} \succeq'_2$.

(7.1) For monotonicity suppose $\langle x_1, x_2 \rangle \succeq \langle y_1, y_2 \rangle$. Then either $x_1 \succ_1 y_1$, or $x_1 \sim_1 y_1$ and $x_2 \succeq_2 y_2$. Then $f_{\langle\rangle}(x_1) \succ_1 f_{\langle\rangle}(y_1)$, or $f_{\langle\rangle}(x_1) \sim'_1 f_{\langle\rangle}(y_1)$ and $f_{\langle x_1\rangle}(x_2) \succeq'_2 f_{\langle x_1\rangle}(y_2) \sim'_2 f_{\langle y_1\rangle}(y_2)$. Hence $f(\langle x_1, x_2 \rangle) \succeq f(\langle y_1, y_2 \rangle)$.

(7.2) For strict monotonicity replace $\succeq_2$ by $\succ_2$. □

2.4 Algebra

See also the books by Lang (1993) or Scheja and Storch (1994).

Let $f : S \to S$ be a binary operator. Then f is called *associative* if

$$f(x, f(y, z)) = f(f(x, y), z) \tag{2.1}$$

for any $x, y, z \in S$. It is called *commutative* if

$$f(x, y) = f(y, x) \tag{2.2}$$

for any $x, y \in S$. A *semigroup* $\langle S, \cdot \rangle$ consists of a set S and an associative operator $\cdot$ on M. A *monoid* $\langle M, \cdot, 1 \rangle$ is a semigroup $\langle M, \cdot \rangle$ together with a *unit element* 1 of M such that

$$1 \cdot x = x \cdot 1 = x \tag{2.3}$$

for any x in M. A *group* $\langle G, \cdot, \_^{-1}, 1 \rangle$ is a monoid $\langle M, \cdot, 1 \rangle$ together with a function $\_^{-1} : G \to G$ such that

$$x \cdot x^{-1} = x^{-1} \cdot x = 1 \tag{2.4}$$

for any x in G. A semigroup is called *commutative* or *abelian* if its operator is commutative. Commutative semigroups, monoids and groups are often written additively as $\langle S, + \rangle$, $\langle M, +, 0 \rangle$ and $\langle M, +, -, 0 \rangle$, respectively. A ring $\langle R, +, \cdot, -, 0, 1 \rangle$ consists of an abelian group $\langle R, +, 0 \rangle$ and a monoid $\langle R, \cdot, 1 \rangle$ such that the distributivity laws

$$a \cdot (b_1 + b_2) = a \cdot b_1 + a \cdot b_2 \tag{2.5}$$
$$(a_1 + a_2) \cdot b = a_1 \cdot b + a_2 \cdot b \tag{2.6}$$

hold. A ring is called *commutative* if its multiplication is commutative. The element 1 is called the *unit element* of R. We will consider only commutative rings with a unit element. A *zero divisor* is an element $a \neq 0$ such that $ab = 0$ for some $b \neq 0$. A commutative ring without zero divisors is called an *integral domain*. In an integral domain multiplication obeys the cancellation law. That is, $ab = ac$ implies $b = c$ for $a \neq 0$. A *unit* of a ring is an element that has a multiplicative inverse. We say that a is *divisible* by b, written $a \mid b$, if there exists some $c \in R$ such that $ac = b$. Two elements a and b in R are *associated* if $a \mid b$ and $b \mid a$. In an integral domain this is equivalent to the existence of a unit c such that $ac = b$.

2.5 First-order predicate logic

For an introduction to first-order logic see for instance the book by Fitting (1996).

Syntax

We consider first-order languages over the single predicate symbol $\approx$ for equality and some set F of function symbols. Additionally there is a set X of variables. A function symbol f has the *arity* $\alpha(f)$. Function symbols with arity 0 are called *constants*.

The set of *terms* over F and X is the smallest set $\mathcal{T}_F(X)$ such that $X \subseteq \mathcal{T}_F(X)$, and if $f \in F$, $\alpha(f) = n$ and $t_1, \ldots, t_n \in \mathcal{T}_F(X)$ then $f(t_1, \ldots, t_n) \in \mathcal{T}_F(X)$.

The set of *equations* over F and X is the smallest set $\mathcal{E}_F(X)$ such that if t_1 and t_2 are terms in $\mathcal{T}_F(X)$ then $t_1 \approx t_2$ is in $\mathcal{E}_F(X)$. The set of *formulas* over F and X is defined as the smallest set $\mathcal{F}_F(X)$ such that:

1. $\mathcal{E}_F(X) \subseteq \mathcal{F}_F(X)$.
2. $\top$ and $\bot$ are formulas in $\mathcal{F}_F(X)$.
3. If ϕ is in $\mathcal{F}_F(X)$ then $\neg\phi$ is a in $\mathcal{F}_F(X)$.

4. If ϕ and ψ are formulas in $\mathcal{F}_F(X)$ then $\phi \wedge \psi$, $\phi \vee \psi$, $\phi \rightarrow \psi$ and $\phi \leftrightarrow \psi$ are formulas in $\mathcal{F}_F(X)$.

5. If ϕ is a formula in $\mathcal{F}_F(X)$ and x is a variable then $\forall x\, \phi$ and $\exists x\, \phi$ are in $\mathcal{F}_F(X)$.

We drop brackets according to the standard conventions: $\neg$, $\forall x$ and $\exists x$ take precedence over $\wedge$ and $\vee$, which in turn take precedence over $\rightarrow$ and $\leftrightarrow$. A variable x is a *free variable* of the formula ϕ if it occurs in a atomic subformula of ϕ which is not below a quantifier $\forall x$ or $\exists x$. All other variables in ϕ are called *bound*. Terms and formulas which contain no variables at all are called *ground*.

A *substitution* σ is a mapping from the set of variables X to the set of terms $\mathcal{T}_F(X)$, where $\sigma(x) \neq X$ for only finitely many $x \in X$. It can be extended in a unique way to a homomorphism $\hat{sigma}$ from $\mathcal{T}_F(X)$ to $\mathcal{T}_F(X)$ by

$$\begin{aligned} \hat{\sigma}(x) &= \sigma(x) && x \in X \\ \hat{\sigma}(f(t_1, \ldots, t_n)) &= f(\hat{\sigma}(t_1), \ldots, \hat{\sigma}(t_n)) && f \in F \end{aligned}$$

We will write $t\sigma$ instead of $\hat{\sigma}(t)$ from now on. The *domain* $\mathcal{D}om(\sigma)$ of a substitution σ is the set of variables $\{x \mid \sigma(x) \neq x\}$, its *range* $\mathcal{R}an(\sigma)$ is the set of terms $\{x\sigma \mid x \in \mathcal{D}om(\sigma)\}$. A substitution is *ground* if its range contains only ground terms. A term s is an *instance* of t if $s = t\sigma$ for some substitution σ. A *ground instance* of t is an instance of t that is ground. We denote the set of ground instances of t by $gnd(t)$.

A *context* u is a term that contains a single occurrence of the special variable $[]$. We write $u[]$ to indicate that u is a context. $[]$ itself is the *empty context*. Then $u[t]$ denotes the term $u\{[]/t\}$, where $[]$ is substituted by t.

We write $f(\ldots, t, \ldots)$ as an abbreviation for $f(u_1, \ldots, u_i, t, u_{i+1}, \ldots, u_n)$. If we write $f(\ldots, s, \ldots)$ and $f(\ldots, t, \ldots)$ in the same context then s and t are in the same argument position, and the terms represented by the dots are equal. Let F be a set of function symbols. An F-*context* is a context whose function symbols are from F.

A *position* in a term is a sequence of natural numbers. The composition of positions is denoted by a decimal dot. The set of positions of a term is defined recursively by $Pos(x) = \{\varepsilon\}$ for variables x and $Pos(f(t_1, \ldots, t_n)) = \{\varepsilon\} \cup \{\pi.i \mid 1 \leq i \leq n,\ \pi \in Pos(t_i)\}$. The empty sequence is called the *root position* and is denoted by ε.

Semantics

An F-*structure* (or F-*interpretation*) I consists of (i) a nonempty set U_I, called the *universe* of I, (ii) for each function symbol f in F a function $f_I : U_I^{\alpha(f)} \rightarrow U_I$, and (iii) a binary relation $\approx_I$. Note that we have not required any congruence properties for the interpretation of equality, as we will need to consider interpretations which are not transitive. A *Herbrand interpretation* is an interpretation I where $U_I = \mathcal{T}_F$. An *assignment* for I is a function $\alpha : X \rightarrow U_I$, which can be extended in a unique way to a function $\hat{\alpha} : \mathcal{T}_F(X) \rightarrow U_I$. A formula ϕ is *true* in I with respect to an assignment α, written $I, \alpha \models \phi$, if either

$$
\begin{aligned}
&\phi = \top, && && (2.7)\\
&\phi = s \approx t, && \text{and } \hat{\alpha}(s) \approx_I \hat{\alpha}(t), && (2.8)\\
&\phi = \neg\psi, && \text{and } I, \alpha \not\models \psi, && (2.9)\\
&\phi = \psi_1 \wedge \psi_2, && \text{and } I, \alpha \models \psi_1 \text{ and } I, \alpha \models \psi_2, && (2.10)\\
&\phi = \psi_1 \vee \psi_2, && \text{and } I, \alpha \models \psi_1 \text{ or } I, \alpha \models \psi_2, && (2.11)\\
&\phi = \psi_1 \rightarrow \psi_2, && \text{and } I, \alpha \models \psi_1 \text{ implies } I, \alpha \models \psi_2, && (2.12)\\
&\phi = \psi_1 \rightarrow \psi_2, && \text{and } I, \alpha \models \psi_1 \text{ if and only if } I, \alpha \models \psi_2, && (2.13)\\
&\phi = \forall x\, \psi, && \text{and } I, \alpha[x/a] \models \psi \text{ for all } a \in U_I, \text{ or} && (2.14)\\
&\phi = \exists x\, \phi && \text{and } I, \alpha[x/a] \models \phi \text{ for some } a \in U_I. && (2.15)
\end{aligned}
$$

Let ϕ be a formula and I an interpretation. We say that ϕ is *true* in I or I is a *model* of ϕ, written $I \models \phi$, if $I, \alpha \models \phi$ for all assignments α. I is a *model* of a set of formulas Φ, if $I \models \phi$ for all $\phi \in \Phi$. By $Mod(\Phi)$ we denote the set of all models of Φ. We say that Φ is *valid* if it is true in all interpretations. Let $\mathcal{I}$ be a set of interpretations. Then we write $\mathcal{I} \models \Phi$ if $I \models \Phi$ for all $I \in \mathcal{I}$. We say that ψ is a *logical consequence* of Φ and write $\Phi \models \psi$ if $Mod(\Phi) \models \psi$. Two formulas ϕ and ψ are said to be *logically equivalent*, written $\phi \equiv \psi$, if ϕ is a logical consequence of ψ and vice-versa. A set Φ of formulas is called *consistent* or *satisfiable* if Φ has a model, and *inconsistent* or *unsatisfiable* otherwise. It is well-known that any satisfiable set of formulas has a Herbrand model. In some contexts we will consider only a subclass $\mathcal{M}$ of all models and write $\Phi \models_{\mathcal{M}} \psi$ if for all models in $\mathcal{M}$ that satisfy Φ also ψ holds. A special case is when the subclass is determined by a set of formulas T. In this context we use the following specialized notions: A formula ψ is a *T-consequence* of ϕ if $T \cup \{\phi\} \models \psi$. Two formulas ϕ and ψ are said to be *T-equivalent* if ψ is a T-consequence of ϕ and vice-versa. A set of formulas Φ is called *T-consistent* if $T \cup \Phi$ has a model, and *T-inconsistent* otherwise.

Clause form

A *literal* is a formula of the form A or $\neg A$ where A is an atomic formula. A *clause* is a formula of the form $L_1 \vee \ldots \vee L_k$ where $L_1, \ldots, L_k$ are literals. A quantifier-free formula is in *clause form* if it is a conjunction of clauses. We may equivalently consider such a formula as a set of clauses. It is well-known that any formula can be transformed into a clause set that is satisfiable if and only if the original formula is satisfiable.

Equality

Let F be a set of function symbols. An interpretation I is called an *equality interpretation* (with respect to F) if $\approx_I$ is a congruence (with respect to F). Formally, an equality interpretation has to satisfy the following sets of axioms:

Reflexivity

$$\text{Refl} = \{x \approx x\}$$

Symmetry

$$\text{Symm} = \{x \approx y \rightarrow y \approx x\}$$

Transitivity

$$\mathrm{Trans} = \{(x \approx y \ \wedge \ y \approx z) \rightarrow x \approx z\}$$

Monotonicity

$$\mathrm{Mon} = \{(x_1 \approx y_1 \ \wedge \ \ldots \ \wedge \ x_n \approx y_n) \rightarrow f(x_1, \ldots, x_n) \approx f(y_1, \ldots, y_n) \mid f \in F\}$$

We let

$$\mathrm{Eq} = \mathrm{Refl} \cup \mathrm{Symm} \cup \mathrm{Trans} \cup \mathrm{Mon}.$$

Let E be a set of equations. Two terms s and t are *E-equivalent*, written $s =_E t$, if $\mathrm{Eq} \cup E \models s \approx t$.

Axioms for algebraic structures

We will need the following algebraic axioms:

Associativity

$$\mathrm{A}(+) = \{(x + y) + z) \approx x + (y + z)\}$$

Commutativity

$$\mathrm{C}(+) = \{x + y \approx y + x\}$$

Distributivity

$$\mathrm{D}(+, \cdot) = \{x \cdot (y + z) \approx x \cdot y + x \cdot z\}$$

Unit law

$$\mathrm{U}(+, 0) = \{x + 0 \approx x\}$$

Idempotency

$$\mathrm{I}(+) = \{x + x \approx x\}$$

Inverse law

$$\mathrm{Inv}(+, -, 0) = \{x + (-x) \approx 0\}$$

We write $\mathrm{AC}(f)$ for $\mathrm{A}(f) \cup \mathrm{C}(f)$ or just AC when the associative-commutative function symbols are known from the context. Furthermore, we let $\mathrm{ACU}(+, 0) = \mathrm{AC}(+) \cup \mathrm{U}(+, 0)$, $\mathrm{ACD}(+, \cdot) = \mathrm{AC}(+) \cup \mathrm{AC}(\cdot) \cup \mathrm{D}(+, \cdot)$ and $\mathrm{ACI}(+) = \mathrm{AC}(+) \cup \mathrm{I}(+)$, with their respective abbreviations ACU, ACD and ACI when the function symbols are known from the context. An equation is called *collapse-free* if neither side is a variable. A set of equations is called *collapse-free* if all its equations are collapse-free.

Automated theorem proving

Given some first-order formula ϕ, the goal of automated theorem proving is to determine whether ϕ is valid. Usually this takes place in the context of some theory T, and in this case we want to know whether $T \models \phi$. This can be done by transforming $\neg\phi$ into clause form and testing for inconsistency. An inference is a pair $\langle\langle C_1, \ldots, C_k\rangle, D\rangle$ of a tuple of clauses $C_1, \ldots, C_k$, called the *premises* of the inference, and a clause D, called the *conclusion*. The inference is *sound* with respect to T if the conclusion is a T-consequence of the premises. We say that the conclusion can be derived from the premises by the inference. A *calculus* or *inference system* is a set of inferences. A *refutation* of a clause set N in a calculus Calc is a sequence $C_1, \ldots, C_n = \perp$ of clauses with $n \geq k$ such that each clause C_i is either from N or can be derived from clauses earlier in the sequence by an inference in Calc. A calculus is *refutationally complete* for a theory T if for any T-inconsistent set N of clauses there exists a refutation of N.

2.6 Constraints

A *constraint* is a logical formula Γ over a language of predicate and function symbols. A *constraint system* is a set of constraints together with a satisfaction relation $\models$ that specifies which ground substitutions σ satisfy a constraint Γ. A substitution σ *satisfies* (or *solves*) a constraint if $\sigma \models \Gamma$. A constraint denotes the set of ground substitutions which satisfy the constraint. A *constrained formula* is a logical formula with a constraint, written $\phi\ [\Gamma]$. The set of ground instances $gnd(\phi\ [\Gamma])$ of a constrained formula is $\{\phi\sigma \mid \sigma \models \Gamma\}$.

2.7 Term orderings

An overview of term orderings is given by Dershowitz (1987). A binary relation R on terms is called *compatible with contexts* if $s\ R\ t$ implies $u[s]\ R\ u[t]$ for any context u. It is called *closed under substitutions* if $s\ R\ t$ implies $s\sigma\ R\ t\sigma$ for every substitutions σ. We say that R has the *subterm property* if $u[t]\ R\ t$ for every nonempty context u. A binary relation R on terms is called *E-compatible* if it is $=_E$-compatible, i.e., $s =_E s'\ R\ t' =_E t$ implies $s\ R\ t$ for all terms s, t, s' and t'. It is called *E-antisymmetric* if it is $=_E$-antisymmetric, i.e., $s\ R\ t$ and $t\ S\ s$ implies $s =_E t$. Note that if a quasi-ordering $\succeq$ on terms is E-compatible and E-antisymmetric then $\sim\ =\ =_E$.

A *reduction ordering* is a strict ordering on terms that is well-founded, compatible with contexts and stable under substitutions. If in addition it has the subterm property, it is called a *simplification ordering*. We say that $\succ$ orients the rewrite rule $l \Rightarrow r$ from left to right if $l \succ r$.

We say that a quasi-ordering $\succeq$ *strictly* has some property if both $\succeq$ and its strict part $\succ$ have the property. A *reduction quasi-ordering* is a quasi-ordering on terms that is well-founded, strictly compatible with contexts and strictly closed under substitutions. If in addition $\succ$ has the subterm property, it is called a *simplification quasi-ordering*. The strict part of a simplification quasi-ordering is a simplification ordering.

Lemma 2.9 *The lexicographic combination of simplification quasi-orderings is a simplification quasi-ordering.*

If $(=_E) \subseteq (\sim)$ then transitivity of $\succeq$ implies E-compatibility. Thus if $R \subseteq (\succ)$ and $E \subseteq (\sim)$ for a term rewriting system R modulo E and a reduction quasi-ordering then R is terminating modulo E.

The *subterm ordering* $\trianglerighteq$ is defined by $s \trianglerighteq t$ if and only if t is a subterm of s. The subterm ordering is the smallest simplification ordering.

Polynomial interpretation

A *polynomial interpretation* is a function p that maps terms $t[x_1, \ldots, x_k]$ to multivariate polynomials in $\mathbb{Z}[x_1, \ldots, x_k]$. The interpretation is defined inductively from polynomials p_f in $\mathbb{Z}[y_1, \ldots, y_{\alpha(f)}]$ for each function symbol f, by

$$p(f(t_1, \ldots, t_n)) = p_f(p(t_1), \ldots, p(t_n)) \tag{2.16}$$

$$p(x_i) = x_i. \tag{2.17}$$

To compare polynomials we use an ordering $>_p$ defined by $p[x_1, \ldots, x_k] >_p q[x_1, \ldots, x_k]$ if $p[n_1, \ldots, n_k] > q[n_1, \ldots, n_k]$ for all $n_1, \ldots, n_k \in \mathbb{N}^{\geq 2}$. Finally, we define $\succeq_p$ such that $s \succeq_p t$ if and only if $p(s) \geq_p p(t)$.

Proposition 2.10 *Let p be a polynomial interpretation such that for each polynomial $p_f(x_1, \ldots, x_n)$ all coefficients are nonnegative and each variable occurs at least once in the polynomial, if f is a constant then $p_f \geq 2$, and if f is unary then $p_f \neq x$. Then $\succeq_p$ is a simplification quasi-ordering that is total on ground terms.*

Lemma 2.11 *(Ben Cherifa and Lescanne 1987) If $p_f(x, y)$ has the form $axy + b(x+y) + c$ where $ac + b - b^2 = 0$ then $\succeq_p$ is* AC(f)*-compatible.*

The AC-RPO

Let $\succeq_{acrpo}$ be the associative commutative recursive path ordering (AC-RPO) of Rubio and Nieuwenhuis (1995)[1], with respect to some arbitrary precedence.

Lemma 2.12 (Rubio-Nieuwenhuis) *$\succ_{acrpo}$ is a simplification ordering that is* AC*-compatible and total up to* AC *on ground terms.*

2.8 Term rewriting

For a general introduction to term rewriting systems we refer the reader to the book of Baader and Nipkow (1998). The survey by Dershowitz and Jouannaud (1990) contains more material on the equational case, which is treated in-depth by Jouannaud and Kirchner (1986).

We state the properties of term rewriting systems modulo a set of equations E. The standard case can be obtained by letting $E = \emptyset$. A *rewrite rule* consists of two terms, the *left-hand side* l and the *right-hand side* r, written $l \Rightarrow r$, such that $\mathcal{V}ar(r) \subseteq \mathcal{V}ar(l)$. In the context of a term ordering $\succ$ such that $l \succ r$ we identify the equation $l \approx r$ with the rewrite rule $l \Rightarrow r$. We say the equation $l \approx r$ is *oriented* from left to right by $\succ$ if $l \succ r$. A *term rewriting system(TRS)* is a set R of rewrite rules. A term rewriting system R *rewrites* a term s to t, written $s \Rightarrow_R t$, if there exists a rule $l \Rightarrow r \in R$, a substitution σ and a context $u[\,]$ such that $s = u[l\sigma]$ and $t = u[r\sigma]$. The subterm $l\sigma$ is called the *redex* of the rewrite step. We say that a term t is *irreducible* or in *normal form* with respect to R if there exists no term t' such that $t \Rightarrow_R t'$. If $s \stackrel{*}{\Rightarrow}_R t$ and t is irreducible we say that t is a *normal form* of s.

[1] This ordering has recently been improved by Rubio (1999).

We let $i(R)$ denote the term rewriting system consisting of all instances of rules in R, and $gnd(R)$ the set of all ground instances. We write $E\backslash R$ for the term rewriting system $\{l' \Rightarrow r \mid l \Rightarrow r \in i(R),\ l' =_E l\}$. We say that R rewrites s to t with E-matching if $s \Rightarrow_{E\backslash R} t$. This is equivalent to the standard definition, where $s \Rightarrow_{E\backslash R} t$ if there exists a rule $l \Rightarrow r \in R$, a substitution σ and a context $u[]$ such that $s = u[l']$, $l' =_{\mathrm{AC}} l\sigma$ and $t = u[r\sigma]$. Note that for a ground term rewriting system R no further instantiation is possible and

$$E\backslash R = \{l' \Rightarrow r \mid l \Rightarrow r \in R,\ l' =_E l\}.$$

An *equational proof* of $s \approx t$ is a sequence

$$s = t_0 \Leftrightarrow_{E\cup R} t_1 \Leftrightarrow_{E\cup R} \ldots \Leftrightarrow_{E\cup R} t_n = t$$

where each step is either the application of an equation in E or a rewriting step with R. A proof of the form $t_1 \Leftarrow_R s \Rightarrow_R t_2$ is called a *peak*, and a proof of the form $t_1 \Leftrightarrow_E s \Rightarrow_R t_2$ is called a *cliff*. We write $s \Downarrow_{R;E} t$ for a *valley proof* $s \overset{*}{\Rightarrow}_R s' \overset{*}{\Leftrightarrow}_E t' \overset{*}{\Leftarrow}_R t$ and say that s and t *converge* in R modulo E. Note that by not mentioning E in rewrite steps we assume that R includes any E-steps needed for matching. In most cases E is clear from the context and we write $s \Downarrow_R t$ for $s \Downarrow_{R;E} t$. If even R is clear we simply write $s \Downarrow t$. By $R^{\Downarrow}$ we denote the set of equations provable by a rewrite proof, that is, $\{s \approx t \mid s \Downarrow_R t\}$.

The term rewriting system R is *Church-Rosser modulo* E if $s \overset{*}{\Leftrightarrow}_{E\cup R} t$ implies $s \Downarrow_{R;E} t$. Let $\Rightarrow_{R/E} = \overset{*}{\Leftrightarrow}_E \cdot \Rightarrow_R \cdot \overset{*}{\Leftrightarrow}_E$. Then R is *terminating modulo* E if $\Rightarrow_{R/E}$ is terminating, i.e., there is no infinite sequence

$$t_1 \Rightarrow_{R/E} t_2 \Rightarrow_{R/E} \cdots .$$

If R is both Church-Rosser and terminating modulo E it is called *convergent modulo* E. In this case the normal form of every term is unique up to E and we denote the normal form with respect to R of a term t by $R(t)$.

Given termination, it suffices to test $s \Downarrow_{R;E} t$ for all peaks $t_1 \Leftarrow_R t \Rightarrow_R t_2$ and cliffs $t_1 \Leftrightarrow_E t \Rightarrow_R t_2$, in order to obtain convergence of R modulo E (Jouannaud and Kirchner 1986).

For the case $E = \mathrm{AC}$ convergence of cliffs is ensured by adding AC-*extensions*. For a rule $l \Rightarrow r$ in R with $l = f(s,t)$ where $f \in F_{\mathrm{AC}}$, its AC-extension is $f(x,l) \Rightarrow f(x,r)$, where x is a new variable (Peterson and Stickel 1981). An AC-extension $f(x,l) \Rightarrow f(x,r)$ is needed only if the cliff

$$f(f(x,s),t) \Leftrightarrow_{\mathrm{AC}} f(x,f(s,t)) \Rightarrow_R f(x,r)$$

does not already converge without the extension. In the presence of AC-extensions it is not necessary to rewrite at a position with an AC-symbol f that also occurs immediately above. It suffices to consider redexes at the root of an f-context.

We will need versions of the Church-Rosser properties which hold only up to some bound with respect to a given term ordering $\succ$. We assume a downward-closed set $\mathcal{T}$ of ground terms, which allows to formalize bounds of the forms $\prec s$ and $\preceq s$ as well as the absence of a bound by the sets $\{t \mid \prec s\}$, $\{t \mid \preceq s\}$ and the set of all terms, respectively. An *equational proof* of $t_1 \approx t_n$ on $\mathcal{T}$ is a sequence

$$t_1 \Leftrightarrow_{R\cup E} t_2 \Leftrightarrow_{R\cup E} \ldots \Leftrightarrow_{R\cup E} t_n$$

where $t_i \in \mathcal{T}$ for all $i = 1, \ldots, n$. Then the term rewriting system $R \subseteq (\succ)$ is *Church-Rosser modulo E on $\mathcal{T}$* if $s \overset{*}{\Leftrightarrow}{}^{\mathcal{T}}_{E \cup R}\, t$ implies $s \Downarrow_{R;E} t$. That is, we require convergence only for equations that can be proved by an equational proof entirely within $\mathcal{T}$. We let the set $\mathrm{Trans}^{\mathcal{T}}$ of *transitivity instances over $\mathcal{T}$* consist of all ground instances

$$(x \not\approx y \vee y \not\approx z \vee x \approx z)\sigma$$

such that $y\sigma \in \mathcal{T}$. The following lemma goes back to Bachmair and Ganzinger (1994a):

Lemma 2.13 *Let R be a term rewriting system that is terminating modulo E, let $(\succ) = (\overset{+}{\Rightarrow}_{R/E})$, and let $\mathcal{T}$ be a set of ground terms that is downward-closed with respect to $\succ$. Then R is Church-Rosser modulo E on $\mathcal{T}$ if and only if $R^{\Downarrow} \models \mathrm{Trans}^{\mathcal{T}}$.*

Proof: For the only-if direction consider an instance

$$t_1 \not\approx t \vee t \not\approx t_2 \vee t_1 \approx t_2$$

of transitivity in $\mathrm{Trans}^{\mathcal{T}}$. Then $t \in \mathcal{T}$. Let us assume that $t_1 \approx t$ and $t \approx t_2$ hold in $R^{\Downarrow}$, which implies the existence of an equational proof

$$t_1 \overset{*}{\Rightarrow}_R t_1' \overset{*}{\Leftrightarrow}_E t_1'' \overset{*}{\Leftarrow}_R t \overset{*}{\Rightarrow}_R t_2'' \overset{*}{\Leftrightarrow}_E t_2' \overset{*}{\Leftarrow}_R t_2.$$

Since all terms in the subproof $t_1' \overset{*}{\Leftrightarrow}_{E \cup R} t_2'$ are less or equal to t they must be in $\mathcal{T}$, and we can apply the Church-Rosser property to obtain $t_1' \Downarrow_R t_2'$, and consequently $t_1 \Downarrow_R t_2$.

For the if-direction suppose that R is not Church-Rosser modulo E on $\mathcal{T}$. Then there exist either a peak $t_1 \Leftarrow_R t \Rightarrow_R t_2$ or a cliff $t_1 \Leftrightarrow_E t \Rightarrow_R t_2$ with $t \in \mathcal{T}$ such that the corresponding instance of transitivity in $\mathrm{Trans}^{\mathcal{T}}$ does not hold. □

3

Theory Path Orderings

3.1 General path orderings

We use the general path ordering of Geser (1996) as a starting point for constructing quasi-orderings modulo E on ground terms. Geser's method of proving compatibility with contexts from a status being "prepared for contexts" cannot cope with flattening, hence it is not applicable in this setting. Consequently, we weaken the notion of a status to a that of a prestatus, which need not be prepared for contexts. By inspection of Geser's proofs one sees that his proofs use only the properties of a prestatus, with the exception of the proofs of reflexivity and of compatibility with contexts. To repair reflexivity we provide another proof. Compatibility with contexts will be the main topic of Section 3.2. Beyond the work of Geser we show that natural conditions on a prestatus imply that the induced GPO is total and E-antisymmetric.

We write $\rhd_{mul}$ for $\succ_{mul}(\unrhd)$. That is, $\rhd_{mul}$ is the strict part of the multiset extension of $\unrhd$. Since $\rhd$ is well-founded, $\rhd_{mul}$ is also well-founded. Given two terms s and t we let $\mathrm{fin}(s,t)$ be the set $\{\langle s',t'\rangle \mid \{s,t\} \rhd_{mul} \{s',t'\}\}$. That is, a pair $\langle s',t'\rangle$ is in $\mathrm{fin}(s,t)$ if either both terms are proper subterms of s or t, or if one term is equal to s or t and the other is a proper subterm of the other. For instance,

$$\mathrm{fin}(f(a),g(b)) = \{\langle f(a),b\rangle, \langle b,f(a)\rangle, \langle a,g(b)\rangle, \langle g(b),a\rangle, \langle a,a\rangle, \langle a,b\rangle, \langle b,a\rangle, \langle b,b\rangle\}.$$

A *quasi-ordering functional* is a function $\succeq^{st}$ which maps any quasi-ordering $\succeq$ on ground terms to a quasi-ordering $\succeq^{st}(\succeq)$ on ground terms. A quasi-ordering functional $\succeq^{st}$ is *subterm founded* on a set of pairs of terms S if $s \succeq^{st}(\succeq)\ t$ is equivalent to $s \succeq^{st}(\succeq \cap\, \mathrm{fin}(s,t))\ t$ for any quasi-ordering $\succeq$ and any pair $\langle s,t\rangle$ in S. We say that $\succeq^{st}$ is *subterm founded* if it is subterm founded on the set of all pairs of ground terms. A quasi-ordering functional $\succeq^{st}$ *decreases infinite derivations* if for every infinite derivation

$$s_1 \succ^{st}(\succeq)\ s_2 \succ^{st}(\succeq)\ \ldots$$

there exists an infinite derivation $t_1 \succ t_2 \succ \ldots$ such that $s_i \rhd t_1$ for some $i \geq 1$. A quasi-ordering functional $\succeq^{st}$ is called a *prestatus* (on S) if (i) $\succeq^{st}$ is subterm founded (on S), and (ii) $\succeq^{st}$ decreases infinite derivations. The *general path ordering* $\succeq_{gpo}(\succeq^{st})$ induced by a prestatus $\succeq^{st}$ is the smallest quasi-ordering such that $s = f(s_1,\ldots,s_m) \succeq_{gpo}(\succeq^{st})\ g(t_1,\ldots,t_n) = t$ if

1. $s_i \succeq_{gpo}(\succeq^{st})\ t$ for some $i = 1,\ldots,m$, or
2. $s \succ_{gpo}(\succeq^{st})\ t_j$ for each $j = 1,\ldots,n$ and $s \succ^{st}(\succeq_{gpo}(\succeq^{st}))\ t$.

Where $\succeq^{st}$ is understood from the context we will write $\succeq_{gpo}$ for $\succeq_{gpo}(\succeq^{st})$.

Lemma 3.1 *(Geser 1996) Let $\succeq^{st}$ be a prestatus.*

1. *If $s \succeq_{gpo} t$ and $t \rhd t'$ then $s \succ_{gpo} t'$.*
2. *If $s \rhd s'$ and $s' \succeq_{gpo} t$ then $s \succ_{gpo} t$.*
3. *$\succeq_{gpo}$ is transitive and $\succ_{gpo}$ is well-founded.*

The proofs below follow a general schema for proving that $\succeq_{gpo}$ has some property P, using subterm foundedness and preservation of P by $\succeq^{st}$.

- Consider some instance $P[t_1, \ldots, t_n]$ of P. Prove $P[t_1, \ldots, t_n]$ for the case where some atom $t_i \succeq t_j$ in P becomes true by case 1 of the definition of $\succeq_{gpo}$. It remains to consider only case 2, where the prestatus is used.
- Let *Fin* be the union of all sets $\text{fin}(t_i, t_j)$ where $t_i \succeq t_j$ is an atom in P.
- Restrict $\succeq_{gpo}$ to Fin, that is, consider $(\succeq_{gpo}) \cap \text{Fin}$.
- Extend $(\succeq_{gpo}) \cap \text{Fin}$ to some $\succeq_P$ that satisfies P, and that coincides with $\succeq_{gpo}$ on Fin. That is, $(\succeq_{gpo}) \cap \text{Fin} = (\succeq_P) \cap \text{Fin}$.
- Then by subterm foundedness

$$\begin{array}{lll} t_i \succeq^{st}(\succeq_{gpo})\ t_j & \text{if and only if} & t_i \succeq^{st}(\succeq_{gpo} \cap \text{Fin})\ t_j \\ & \text{if and only if} & t_i \succeq^{st}(\succeq_P \cap \text{Fin})\ t_j \\ & \text{if and only if} & t_i \succeq^{st}(\succeq_P)\ t_j. \end{array}$$

- Use that $\succeq^{st}$ preserves P to conclude that $P[t_1, \ldots, t_n]$ holds for $\succeq^{st}(\succeq_{gpo})$ and hence for $\succeq_{gpo}$.

Geser (1996) uses this schema to prove transitivity of $\succeq_{gpo}$.

Lemma 3.2 *Let $\succeq^{st}$ be a prestatus. Then $\succeq_{gpo}$ is reflexive.*

Proof: We use induction on terms with respect to $\rhd$. Let $s = f(s_1, \ldots, s_m)$ and suppose $\succeq_{gpo}$ is reflexive on all subterms of s. In particular, $s_i \succeq_{gpo} s_i$ for $i = 1, \ldots, m$. By Lemma 3.1 we get $s \succ_{gpo} s_i$ for $i = 1, \ldots, m$. It remains to show that $s \succeq^{st}(\succ_{gpo})\ s$. Let $\succeq_1$ be the reflexive closure of $(\succeq_{gpo}) \cap \text{fin}(s, s)$. By induction hypothesis $s' \succeq_{gpo} s'$ for all s' such that $\{s, s\} \rhd_{mul} \{s', s'\}$, so $(\succeq_{gpo}) \cap \text{fin}(s, s) = (\succeq_1) \cap \text{fin}(s, s)$. Since $\succeq_1$ is reflexive, and since $\succeq^{st}$ preserves reflexivity, $s \succeq^{st}(\succeq_1)\ s$. Then

$$\begin{array}{lll} s \succeq^{st}(\succeq_1)\ s & \text{if and only if} & s \succeq^{st}(\succeq_1 \cap \text{fin}(s, s))\ s \\ & \text{if and only if} & s \succeq^{st}(\succeq_{gpo} \cap \text{fin}(s, s))\ s \\ & \text{if and only if} & s \succeq^{st}(\succeq_{gpo})\ s \end{array}$$

and we conclude $s \succeq_{gpo} s$. □

Lemma 3.3 *Let $\succeq^{st}$ be a prestatus that preserves totality. Then $\succeq_{gpo}$ is total.*

Proof: By induction on pairs of terms with respect to the multiset extension of the subterm ordering. Consider two terms $s = f(s_1, \ldots, s_m)$ and $t = g(t_1, \ldots, t_n)$.

(1) Suppose $s_i \succeq_{gpo} t$ for some $i = 1, \ldots, m$. Then $s \succeq_{gpo} t$.

(2) Suppose $t_j \succeq_{gpo} s$ for some $j = 1, \ldots, n$. Then $t \succeq_{gpo} s$.

(3) Otherwise by induction hypothesis $s \succ_{gpo} t_j$ for each $j = 1, \ldots, n$ and $t \succ_{gpo} s_i$ for each $i = 1, \ldots, n$. It remains to show that $s \succeq^{st}(\succeq_{gpo})\ t$ or $t \succeq^{st}(\succeq_{gpo})\ s$. Consider some total ordering $\succeq_t$ on terms. We let $s' \succeq_{total} t'$ if either $\langle s', t' \rangle \in \mathrm{fin}(s,t)$ and $s' \succeq_{gpo} t'$, or $\langle s', t' \rangle \notin \mathrm{fin}(s,t)$ and $s' \succeq_t t'$. Then $(\succeq_{total}) \cap \mathrm{fin}(s,t) = (\succeq_{gpo}) \cap \mathrm{fin}(s,t)$, so $s \succeq^{st}(\succeq_{total})\ t$ if and only if $s \succeq^{st}(\succeq_{gpo})\ t$ by subterm foundedness. Clearly $\succeq_{total}$ is total, so $\succeq^{st}(\succeq_{total})$ is total since $\succeq^{st}$ preserves totality. We conclude that $s \succeq^{st}(\succeq_{gpo})\ t$ or $t \succeq^{st}(\succeq_{gpo})\ s$. □

Lemma 3.4 *Let $\succeq^{st}$ be a prestatus that preserves E-antisymmetry. Then $\succeq_{gpo}$ is E-antisymmetric.*

Proof: Let s and t be terms such that $s \succeq_{gpo} t$ and $t \succeq_{gpo} s$ and suppose that $\succeq_{gpo}$ is E-antisymmetric for all pairs of terms s' and t' such that $\{s, t\} \rhd_{mul} \{s', t'\}$. We have to show $s =_E t$.

(1) Suppose $s \succeq_{gpo} t$ by case (i) of the definition of $\succeq_{gpo}$. Then $s_i \succeq_{gpo} t$ for some $i = 1, \ldots, m$. By Lemma 3.1 we get $s \succ_{gpo} t$, a contradiction to $t \succeq_{gpo} s$.

(2) Analogously, $t \succeq_{gpo} s$ by case (i) leads to a contradiction.

(3) So $s \succeq_{gpo} t$ and $t \succeq_{gpo} s$ by case (ii), which implies $s \succeq^{st}(\succeq_{gpo})\ t$ and $t \succeq^{st}(\succeq_{gpo})\ s$. Let $(\succeq_{as}) = (\succeq_{gpo}) \cap \mathrm{fin}(s,t)$. By induction hypothesis $\succeq_{as}$ is E-antisymmetric. Since $\succeq^{st}$ preserves E-antisymmetry and is subterm founded, $s \succeq^{st}(\succeq_{gpo})\ t$ and $t \succeq^{st}(\succeq_{gpo})\ s$ imply $s =_E t$. □

We say that $\succeq^{st}$ is *prepared for E-compatibility* if $s\sigma \succeq^{st}(\succeq)\ t\sigma$ for any ground instance $s\sigma \approx t\sigma$ of an equation $s \approx t$ in E and for any quasi-ordering $\succeq$.

Lemma 3.5 *Let E be a set of equations, let $\succeq^{st}$ be a prestatus that is prepared for E-compatibility, and suppose that $\succeq_{gpo}$ is compatible with contexts. Then $\succeq_{gpo}$ is E-compatible.*

Proof: By induction on $\langle s, t \rangle$ with respect to $\rhd_{mul}$. Let $s = f(s_1, \ldots, s_m)$ and $t = g(t_1, \ldots, t_n)$. We have to show $s \succeq_{gpo} t$ for any E-step $s \Leftrightarrow_E t$. Together with reflexivity and transitivity of $\succeq_{gpo}$ this implies $(=_E) \subseteq (\succeq_{gpo})$.

(1) Suppose the E-step is not at the root of s and t. Then $s_i \Leftrightarrow_E t_i$ for some $i = 1, \ldots, n$ and $s_j = t_j$ for all $j \neq i$ in $1, \ldots, n$. By using the induction hypothesis we get $s_i \succeq_{gpo} t_i$ and $s \succeq_{gpo} t$ by compatibility with contexts.

(2) It remains to consider an E-step at the root position. We can write s as $s'\sigma$ and t as $t'\sigma$ where $s' \approx t'$ is the equation in E that is used. Since $\succeq^{st}$ is prepared for E-compatibility, $s \succeq^{st}(\succeq_{gpo})\ t$ follows. □

3.2 Theory path orderings

Theory path orderings generalize the idea underlying the associative path ordering (APO) of Bachmair and Plaisted (1985), that compatibility with contexts can be achieved for path orderings if interpreted function symbols are minimal in the precedence. It combines a lexicographic path ordering on nontheory function symbols with a special treatment of symbols in the theory, which is formalized by a status function.

We let F denote the set of all function symbols, F_E the function symbols in E and F_T the function symbols in the theory T. Since in general not all function symbols in F_T need to be treated specially by the ordering, we select a set F_I of *interpreted function symbols*interpreted function symbol such that $F_E \subseteq F_I \subseteq F_T$. Function symbols not in F_I are called *free*. We let $\mathcal{I}$ be the set of terms over F with an interpreted function symbol at the root, and $\mathcal{A}$ the set of terms with a free function symbol at the root. Terms in $\mathcal{A}$ are called *atomic*. A *precedence* $\succeq_p$ is a quasi-ordering on function symbols whose strict part is well-founded. A precedence $\succeq_p$ is called *TPO-admissible* for F_I if $f \sim_p g$ for any pair of function symbols from F_I, $f \succ_p g$ whenever $f \notin F_I$ and $g \in F_I$, and $f = g$ whenever $f \sim_p g$ for f and g not in F_I. A partial ordering $\succeq$ on $F \setminus F_I$ can be extended to a TPO-admissible quasi-ordering on F by letting $f \succeq g$ whenever either (i) f and g not in F_I and $f \succeq g$, or (ii) g in F_I. If $\succeq$ is total this is the only TPO-admissible extension. E.g., for $F_I = \{+, 0\}$ and free function symbols $\{a, f\}$ with a given precedence $f \succ_p a$ the TPO-admissible extension is $f \succ_p a \succ_p + \sim_p 0$. A quasi-ordering functional $\succeq^{st}$ is *strictly internally prepared for contexts* with respect to F_I if

$$s \succeq^{st}(\succeq)\ t \text{ implies } f(\ldots, s, \ldots) \succeq^{st}(\succeq)\ f(\ldots, t, \ldots) \qquad \text{and}$$
$$s \succ^{st}(\succeq)\ t \text{ implies } f(\ldots, s, \ldots) \succ^{st}(\succeq)\ f(\ldots, t, \ldots)$$

for any f in F_I. We say that a quasi-ordering functional $\succeq^{st}$ has the *multiset properties* for F_I if it satisfies

$$s \succeq t \iff s \succeq^{st}(\succeq)\ t \qquad \text{for } s \in \mathcal{A} \text{ and } t \in \mathcal{A} \tag{M0}$$
$$s_1 \succeq^{st}(\succeq)\ t \vee \ldots \vee s_m \succeq^{st}(\succeq)\ t \impliedby f(s_1, \ldots, s_m) \succeq^{st}(\succeq)\ t \quad \text{for } t \in \mathcal{A} \tag{M1}$$
$$s_1 \succeq^{st}(\succeq)\ t \vee \ldots \vee s_m \succeq^{st}(\succeq)\ t \implies f(s_1, \ldots, s_m) \succeq^{st}(\succeq)\ t \tag{M2}$$
$$s \succ^{st}(\succeq)\ t_1 \wedge \ldots \wedge s \succ^{st}(\succeq)\ t_n \implies s \succeq^{st}(\succeq)\ f(t_1, \ldots, t_n) \quad \text{for } s \in \mathcal{A} \tag{M3}$$
$$s \succ^{st}(\succeq)\ t_1 \wedge \ldots \wedge s \succ^{st}(\succeq)\ t_n \impliedby s \succeq^{st}(\succeq)\ f(t_1, \ldots, t_n) \tag{M4}$$

for any f in F_I. A quasi-ordering functional $\succeq^{st}$ is called a *TPO-status* for F_I if it is subterm founded on $\mathcal{I}^2$, decreases infinite derivations in $\mathcal{I}$, is strictly internally prepared for contexts with respect to F_I, and has the multiset properties for F_I.

The theory path ordering $\succeq_{tpo}(\succeq_p, \succeq^{st})$ induced by a TPO-admissible precedence $\succeq_p$ and a TPO-status $\succeq^{st}$ is defined as the smallest binary relation such that

$$s = g(s_1, \ldots, s_m) \succeq_{tpo}(\succeq_p, \succeq^{st})\ h(t_1, \ldots, t_n) = t$$

if

1. $s_i \succeq_{tpo}(\succeq_p, \succeq^{st})\ t$ for some $i = 1, \ldots, m$, or
2. $s \succ_{tpo}(\succeq_p, \succeq^{st})\ t_j$ for each $j = 1, \ldots, n$ and either
   (a) $g \succ_p h$,
   (b) $g \sim_p h \notin F_I$ and $\langle s_1, \ldots, s_m\rangle \succeq_{lex}(\succeq_{tpo}(\succeq_p, \succeq^{st}))\ \langle t_1, \ldots, t_n\rangle$, or
   (c) $g \sim_p h \in F_I$ and $s \succeq^{st}(\succeq_{tpo}(\succeq_p, \succeq^{st}))\ t$.

We assume that each function symbol has a fixed arity, hence $m = n$ in case (2b). Where $\succeq_p$ and $\succeq^{st}$ are understood we write $\succeq_{tpo}$ for $\succeq_{tpo}(\succeq_p, \succeq^{st})$. To view $\succeq_{tpo}(\succeq_p, \succeq^{st})$ as a general path ordering we have to define a suitable quasi-ordering functional. We define $\succeq^{st}_T(\succeq_p, \succeq^{st})$ by $s = g(s_1, \ldots, s_m) \succeq^{st}_T(\succeq_p, \succeq^{st})(\succeq)\ h(t_1, \ldots, t_n) = t$ if and only if

(a) $g \succ_p h$,

(b) $g \sim_p h \notin F_I$ and $\langle s_1, \ldots, s_m \rangle \succeq_{lex}(\succeq) \langle t_1, \ldots, t_n \rangle$, or

(c) $g \sim_p h \in F_I$ and $s \succeq^{st}(\succeq)\ t$.

Then clearly $\succeq_{tpo}(\succeq_p, \succeq^{st}) = \succeq_{gpo}(\succeq_T^{st}(\succeq_p, \succeq^{st}))$.

Lemma 3.6 *Let E be a set of collapse-free equations, let $F_I \supseteq F_E$, let $\succeq_p$ be a precedence that is TPO-admissible for F_I, and let $\succeq^{st}$ be a TPO-status.*

1. *Then $\succeq_T^{st}(\succeq_p, \succeq^{st})$ is a prestatus.*
2. *If $\succeq_p$ is total and if $\succeq^{st}$ preserves totality on $\mathcal{I}^2$ then $\succeq_T^{st}(\succeq_p, \succeq^{st})$ preserves totality.*
3. *If $\succeq^{st}$ preserves E-antisymmetry on $\mathcal{I}^2$ then $\succeq_T^{st}(\succeq_p, \succeq^{st})$ preserves E-antisymmetry.*
4. *If $\succeq^{st}$ is prepared for E-compatibility then $\succeq_T^{st}(\succeq_p, \succeq^{st})$ is prepared for E-compatibility.*

Proof: (*Prestatus*) We have to show that $\succeq_T^{st}(\succeq_p, \succeq^{st})$ preserves quasi-orderings, is subterm founded and decreases infinite derivations.

(*Preserves quasi-orderings*) For reflexivity we observe that $\succeq_p$ is reflexive and that the lexicographic extension and $\succeq^{st}$ preserve reflexivity.

For transitivity suppose that $\succeq$ is a transitive relation and consider ground terms $s = g(s_1, \ldots, s_m)$, $t = h(t_1, \ldots, t_n)$ and $u = f(u_1, \ldots, u_k)$ such that $s \succeq_T^{st}(\succeq_p, \succeq^{st})(\succeq)\ t$ and $t \succeq_T^{st}(\succeq_p, \succeq^{st})(\succeq)\ u$.

(1) Suppose $g \sim_p h$ and $h \sim_p f$. Then by transitivity of $\succeq_p$ also $g \sim_p f$, and all three root symbols must be either equal or from F_I.

(1.1) Suppose $g = h = f \notin F_I$. Then by assumption $\langle s_1, \ldots, s_m \rangle \succeq_{lex}(\succeq) \langle t_1, \ldots, t_n \rangle$ and $\langle t_1, \ldots, t_n \rangle \succeq_{lex}(\succeq) \langle u_1, \ldots, u_k \rangle$. Since $\succeq_{lex}$ preserves transitivity, $\langle s_1, \ldots, s_m \rangle \succeq_{lex}(\succeq)$ $\langle u_1, \ldots, u_k \rangle$ holds as well. This implies $s \succeq_T^{st}(\succeq_p, \succeq^{st})(\succeq)\ u$.

(1.2) Otherwise f, g and h are in F_I. Then $s \succeq^{st}(\succeq)\ t$ and $t \succeq^{st}(\succeq)\ u$, and since $\succeq^{st}$ preserves transitivity $s \succeq^{st}(\succeq)\ u$. Hence $s \succeq_T^{st}(\succeq_p, \succeq^{st})(\succeq)\ u$.

(2) Otherwise either $g \succ_p h$ or $h \succ_p f$ and we get $g \succ_p f$ by transitivity of $\succeq_p$, which in turn implies $s \succeq_T^{st}(\succeq_p, \succeq^{st})(\succeq)\ u$.

(*Subterm founded*) Let $s = g(s_1, \ldots, s_m)$ and $t = h(t_1, \ldots, t_n)$ be ground terms. We have to show that

$$s \succeq_T^{st}(\succeq_p, \succeq^{st})(\succeq)\ t \quad \text{if and only if} \quad s \succeq_T^{st}(\succeq_p, \succeq^{st})(\succeq \cap \operatorname{fin}(s,t))\ t.$$

(1) Suppose $g \not\sim_p h$. Then

$$\begin{aligned} s \succeq_T^{st}(\succeq_p, \succeq^{st})(\succeq)\ t \quad & \text{if and only if} \quad g \succ_p h \\ & \text{if and only if} \quad s \succeq_T^{st}(\succeq_p, \succeq^{st})(\succeq \cap \operatorname{fin}(s,t))\ t. \end{aligned}$$

(2) Suppose $g \sim_p h \notin F_I$. Then $g = h$ and $m = n$, and

$$\begin{aligned} s \succeq_T^{st}(\succeq_p, \succeq^{st})(\succeq)\ t \quad & \text{if and only if} \quad \langle s_1, \ldots, s_m \rangle \succeq_{lex}(\succeq) \langle t_1, \ldots, t_n \rangle \\ & \text{if and only if} \quad \langle s_1, \ldots, s_m \rangle \succeq_{lex}(\succeq \cap \operatorname{fin}(s,t)) \langle t_1, \ldots, t_n \rangle \\ & \text{if and only if} \quad s \succeq_T^{st}(\succeq_p, \succeq^{st})(\succeq \cap \operatorname{fin}(s,t))\ t, \end{aligned}$$

since $s \rhd s_1, \ldots, s_m$ and $t \rhd t_1, \ldots, t_n$.

(3) Otherwise $g \sim_p h \in F_I$. Then also $g \in F_I$, and

$$\begin{array}{rll} s \succeq^{st}_T(\succeq_p, \succeq^{st})(\succeq)\; t & \text{if and only if} & s \succeq^{st}(\succeq)\; t \\ & \text{if and only if} & s \succeq^{st}(\succeq \cap \operatorname{fin}(s,t))\; t \\ & \text{if and only if} & s \succeq^{st}_T(\succeq_p, \succeq^{st})(\succeq \cap \operatorname{fin}(s,t))\; t, \end{array}$$

since $\succeq^{st}$ is subterm founded in case both terms have a symbol from F_I at their root.

(*Decreases infinite derivations*) Suppose there is some infinite descending chain

$$t_1 \succ^{st}_T(\succeq_p, \succeq^{st})(\succeq)\; t_2 \succ^{st}_T(\succeq_p, \succeq^{st})(\succeq)\; \dots\; .$$

Since $\succ_p$ is well-founded, there exists a minimal function symbol f in the set of root symbols of t_1, t_2, ..., and there exists some $i \geq 1$ such that $root(t_i) \sim_p f$ for all $j \geq i$.

(1) Suppose f is not in F_I. Then for $j \geq i$ each term t_j has the root symbol f and we have an infinite descending chain

$$\langle t_{i,1}, \dots, t_{i,n}\rangle \succ_{lex}(\succeq)\; \langle t_{i+1,1}, \dots, t_{i+1,n}\rangle \succ_{lex}(\succeq)\; \dots\; .$$

Then there exists an infinite descending subchain $t_{j_1 l} \succ_l t_{j_2 l} \succ_l \dots$ for some $l = 1, \dots, n$ and $i \leq j_1 < j_2 < \dots$, which satisfies $t_{j_1 l} \lhd t_{j_1}$.

(2) Otherwise $f \in F_I$. Then all terms are in $\mathcal{I}$, and we have an infinite descending chain $t_i \succ^{st}(\succeq)\; t_{i+1} \succ^{st}(\succeq)\; \dots$. Since $\succeq^{st}$ decreases infinite derivations on $\mathcal{I}$ we get the desired infinite descending chain in $\succ$.

(*Preserves totality*) Let $\succeq$ be a total quasi-ordering. We consider ground terms $s = g(s_1, \dots, s_m)$ and $t = h(t_1, \dots, t_n)$.

(1) Suppose $g \not\sim_p h$. Since $\succeq_p$ is total, either $g \succ_p h$ or $h \succ_p g$, and hence either $s \succ^{st}_T(\succeq_p, \succeq^{st})(\succeq)\; t$ or $t \succ^{st}_T(\succeq_p, \succeq^{st})(\succeq)\; s$, respectively.

(2) Suppose $g = h \notin F_I$. Since the lexicographic extension preserves totality, we have either $\langle s_1, \dots, s_m\rangle \succeq_{lex}(\succeq)\; \langle t_1, \dots, t_n\rangle$ or $\langle t_1, \dots, t_n\rangle \succeq_{lex}(\succeq)\; \langle s_1, \dots, s_m\rangle$. We conclude that $s \succeq^{st}_T(\succeq_p, \succeq^{st})(\succeq)\; t$ or $t \succeq^{st}_T(\succeq_p, \succeq^{st})(\succeq)\; s$, respectively.

(3) It remains to consider the case when g and h are in F_I. Then we use that $\succeq^{st}$ preserves totality.

(*Preserves E-antisymmetry*) Let $\succeq$ be E-antisymmetric, $s = g(s_1, \dots, s_m)$ and $t = h(t_1, \dots, t_n)$ ground terms, and suppose $s \sim^{st}_T(\succeq_p, \succeq^{st})(\succeq)\; t$. Then $g \sim_p h$.

(1) Suppose $g = h \notin F_I$. Then $m = n$, and $\langle s_1, \dots, s_m\rangle \sim_{lex}(\succeq)\; \langle t_1, \dots, t_n\rangle$, which implies $s_i \sim t_i$ for $i = 1, \dots, m$. Since $\succeq$ is E-antisymmetric, this implies $s_i =_E t_i$ for $i = 1, \dots, m$, and since $=_E$ is a congruence we conclude $s =_E t$.

(2) Otherwise both g and h are in F_I. Since $\succeq^{st}$ preserves E-antisymmetry, $\succeq^{st}(\succeq)$ is E-antisymmetric, and $s \sim^{st}(\succeq)\; t$ implies $s =_E t$.

(*Prepared for E-compatibility*) Let $s = \hat{s}\sigma \approx \hat{t}\sigma = t$ be a ground instance of an equation $\hat{s} \approx \hat{t}$ in E, where $s = g(s_1, \dots, s_m)$ and $t = h(t_1, \dots, t_n)$. Since E is collapse-free, g and h are both in F_I. Then $g \sim_p h$ and $s \succeq^{st}(\succeq)\; t$ for any quasi-ordering $\succeq$, since $\succeq^{st}$ is prepared for E-compatibility. This implies $s \succeq^{st}_T(\succeq_p, \succeq^{st})(\succeq)\; t$. □

Lemma 3.7 *Let $\succeq_p$ be a precedence that is TPO-admissible for F_I, and let $\succeq^{st}$ be a quasi-ordering functional that is subterm founded on $\mathcal{I}^2$ and has the multiset properties for F_I. Then $s \succeq_{tpo} t$ if and only if $s \succeq^{st}(\succeq_{tpo})\; t$.*

Proof: We use induction on the pairs $\langle s, t\rangle$ with respect to $\rhd_{mul}$. Let $s = g(s_1, \dots, s_m)$ and $t = h(t_1, \dots, t_n)$.

(1) Suppose g and h are not in F_I. Then $s \succeq_{tpo} t$ if and only if $s \succeq^{st}(\succeq_{tpo})\ t$ by (M0).

(2) Suppose g is in F_I and h is not, which implies $h \succ_p g$. Then $s \succeq_{tpo} t$ if and only if there exists some $i = 1, \dots, m$ such that $s_i \succeq_{tpo} t$. This is equivalent to $s_i \succeq^{st}(\succeq_{tpo})\ t$ by the induction hypothesis, and to $s \succeq^{st}(\succeq_{tpo})\ t$ by (M1) and (M2).

(3) Suppose g is not in F_I and h is, which implies $g \succ_p h$. Then $s \succeq_{tpo} t$ is equivalent to $s \succ_{tpo} t_j$ for all $j = 1, \dots, n$ by case (ii) of the definition of $\succeq_{tpo}$. Note that case (i) implies case (ii) in this context. By induction hypothesis this is equivalent to $s \succ^{st}(\succeq_{tpo})\ t_j$ for all $j = 1, \dots, n$, and to $s \succeq^{st}(\succeq_{tpo})\ t$ by (M3) and (M4).

(4) Otherwise g and h are in F_I.

For the only-if-direction suppose $s \succeq_{tpo} t$. If $s \succeq_{tpo} t$ by case (i) of the definition of $\succeq_{tpo}$ then there exists some $i = 1, \dots, m$ such that $s_i \succeq_{tpo} t$. This implies $s_i \succeq^{st}(\succeq_{tpo})\ t$ by induction hypothesis, and $s \succeq^{st}(\succeq_{tpo})\ t$ by (M2). Otherwise case (ii) of the definition of $\succeq_{tpo}$ holds, which explicitly includes $s \succeq^{st}(\succeq_{tpo})\ t$.

For the if-direction assume $s \succeq^{st}(\succeq_{tpo})\ t$. Then $s \succ^{st}(\succeq_{tpo})\ t_j$ for all $j = 1, \dots, n$ by (M4), and $s \succeq_{tpo} t_j$ for all $j = 1, \dots, n$ by induction hypothesis. Together with $s \succeq^{st}(\succeq_{tpo})\ t$ this satisfies case (ii) of the definition of $\succeq_{tpo}$, hence $s \succeq_{tpo} t$. $\square$

Theorem 3.8 *Let E be a set of collapse-free equations, let $F_I \supseteq F_E$, let $\succeq_p$ be a precedence that is TPO-admissible for F_I, and let $\succeq^{st}$ be a TPO-status for F_I.*

1. *Then $\succeq_{tpo}$ is a simplification quasi-ordering.*
2. *If $\succeq_p$ is total and $\succeq^{st}$ preserves totality on $\mathcal{I}^2$ then $\succeq_{tpo}$ is total.*
3. *If $\succeq^{st}$ preserves E-antisymmetry on $\mathcal{I}^2$ then $\succeq_{tpo}$ is E-antisymmetric.*
4. *If $\succeq^{st}$ is prepared for E-compatibility then $\succeq_{tpo}$ is E-compatible.*

Proof: (*Simplification quasi-ordering*) By Lemma 3.6 $\succeq^{st}_T(\succeq_p, \succeq^{st})$ is a prestatus. Thus $\succeq_{tpo}$ is a quasi-ordering on ground terms that has the subterm property and $\succ_{tpo}$ is well-founded. It remains to show compatibility with contexts and strict compatibility with contexts.

Let $s = g(s_1, \dots, s_m)$, $t = h(t_1, \dots, t_n)$, $s' = f(u_1, \dots, u_i, s, u_{i+1}, \dots, u_k)$, $t' = f(u_1, \dots, u_i, t, u_{i+1}, \dots, u_k)$, and suppose $s \succeq_{tpo} t$. By the subterm property $s' \succ_{gpo} u_j$ for $j = 1, \dots, k$, and $s' \succ_{tpo} t$ by Lemma 3.1.

(*Compatibility with contexts*) We have to show $s' \succeq_{tpo} t'$.

(1) Suppose $f \notin F_I$. Then $s' \succeq^{st}_T(\succeq_p, \succeq^{st})(\succeq_{tpo})\ t'$ because

$$\langle u_1, \dots, u_i, s, u_{i+1}, \dots, u_k\rangle \succeq_{lex}(\succeq_{tpo})\ \langle u_1, \dots, u_i, t, u_{i+1}, \dots, u_k\rangle$$

by definition of $\succeq_{lex}$.

(2) Otherwise $f \in F_I$. From $s \succeq_{tpo} t$ we get $s \succeq^{st}(\succeq_{tpo})\ t$ by Lemma 3.7, and by internal preparedness for contexts $s' \succeq^{st}(\succeq_{tpo})\ t'$, which is equivalent to $s' \succeq^{st}_T(\succeq_p, \succeq^{st})(\succeq_{tpo})\ t'$ for f in F_I.

(*Strict compatibility with contexts*) We have to show $s' \succ_{tpo} t'$ under the assumption $s \succ_{tpo} t$. Since there cannot exist a $j = 1, \dots, n$ such that $t_j \succ_{tpo} s$, case (i) of the definition of $\succeq_{tpo}$ cannot be used to obtain $t' \succeq_{tpo} s'$. Hence it suffices to show $s' \succ^{st}_T(\succeq_p, \succeq^{st})(\succeq_{tpo})\ t'$ in order to conclude $s' \succ_{tpo} t'$.

(1) Suppose $f \notin F_I$. Then $s' \succ_T^{st}(\succeq_p, \succeq^{st})(\succeq_{tpo})\ t'$ if and only if

$$\langle u_1, \ldots, u_i, s, u_{i+1}, \ldots, u_k\rangle \succ_{lex}(\succeq_{tpo})\ \langle u_1, \ldots, u_i, t, u_{i+1}, \ldots, u_k\rangle,$$

which follows from the definition of $\succ_{lex}$.

(2) Otherwise $f \in F_I$. From $s \succ_{tpo} t$ we get $s \succ^{st}(\succeq_{tpo})\ t$ by Lemma 3.7, and by strict internal preparedness for contexts $s' \succ^{st}(\succeq_{tpo})\ t'$, which implies $s' \succ_T^{st}(\succeq_p, \succeq^{st})(\succeq_{tpo})\ t'$ and in turn $s' \succ_{tpo} t'$.

(*Total*) Since $\succeq_p$ is total and $\succeq^{st}$ preserves totality on $\mathcal{I}^2$, we get that $\succeq_T^{st}(\succeq_p, \succeq^{st})$ preserves totality by Lemma 3.6(2). Hence $\succeq_{tpo}$ is total by Lemma 3.3.

(*E-antisymmetry*) $\succeq^{st}$ preserves E-antisymmetry on $\mathcal{I}^2$, hence $\succeq_T^{st}(\succeq_p, \succeq^{st})$ preserves E-antisymmetry by Lemma 3.6(3), and $\succeq_{tpo}$ is E-antisymmetric by Lemma 3.4.

(*E-compatibility*) Since $\succeq^{st}$ is prepared for E-compatibility we get that $\succeq_T^{st}(\succeq_p, \succeq^{st})$ is prepared for E-compatibility by Lemma 3.6(4). Since $\succeq_{tpo}$ is compatible with contexts it is E-compatible by Lemma 3.5. □

3.3 From extension function to TPO-status

When defining a TPO-status it is often useful to represent atomic subterms by constants. This ensures that the ordering obtained from the status for atomic subterms is determined only by its argument quasi-ordering. For instance, the definition of a TPO-status often involves normalizing with respect to some distributivity rules. By hiding atomic subterms in constants no rewriting can take place in atomic subterms. Also, extending an ordering on constants to terms is more natural and allows to reuse known simplification quasi-orderings in a status function.

We let F_C be the set of new constants $\{c_t \mid t \in \mathcal{A}\}$. That is, we assume that F_C and F are disjoint, where $\mathcal{A}$ contains only terms over F. Then for a given ordering $\succeq$ on terms over F we define the ordering $\succeq_c (\succeq)$ on constants in F_C by $c_s \succeq_c (\succeq) c_t$ if and only if $s \succeq t$. We will pack atomic subterms into constants from F_C and compare them according to $\succeq_c (\succeq)$. Technically, we let U be the convergent term rewriting system $\{c_t \Rightarrow t \mid t \in \mathcal{A}\}$, and write $U(t)$ for the normal form of t with respect to U. We use the term rewriting system

$$P_{F_I} = \{t \Rightarrow c_{t'} \mid t \text{ ground term over } F \cup F_C,\ \ t \notin F_C,\ \ t' = U(t) \text{ and } t' \in \mathcal{A}\}$$

for packing atomic subterms into constants. The unpacking of t in the definition of P_{F_I} is needed to remove nested constants in t. For termination of P_{F_I} observe that the number of symbols from F decreases in each step. For confluence observe that P_{F_I} contains a rule $u[c_{s'}] \Rightarrow c_{t'}$ for each critical pair

$$c_{t'} \Leftarrow_{P_{F_I}} u[s] \Rightarrow_{P_{F_I}} u[c_{s'}]$$

where $t' = U(u[s])$, $s' = U(s)$ and u is a nonempty F-context, since $U(u[c_{s'}]) = U(u[s]) = t' \in \mathcal{A}$.

It remains to obtain a quasi-ordering on the packed terms. Let $\succeq_t$ be a function that maps any quasi-ordering $\succeq_c$ on constants F_C to a quasi-ordering on terms over $F_I \cup F_C$, with the following properties:

1. $\succeq_t(\succeq_c)$ extends $\succeq_c$, is strictly compatible with contexts and has the subterm property.

2. Whenever there is an infinite descending chain $t_1 \succ_t(\succeq_c)\ t_2 \succ_t(\succeq_c) \ldots$ of terms over $F_I \cup F_C$ then there exists an infinite descending chain $c_1 \succ_c c_2 \succ_c \ldots$ of constants in F_C such that c_1 occurs in some t_j for $j \geq 1$.

3. Let c be a constant in F_C. If $c \succ_c c'$ for all constants c' occurring in a term t then $c \succ_t(\succeq_c)\ t$.

4. If $t \succeq_t(\succeq_c)\ c$ then $c' \succeq_c c$ for some constant c' in t.

Then we will call $\succeq_t$ an *extension function*. Property 3 is the constant dominance condition of Baader (1997). Note that it implies property 4 for total $\succeq_c$. We can now define a TPO-status $\succeq_t^{st}$ by $s \succeq_t^{st}(\succeq)\ t$ if and only if $P_{F_I}(s) \succeq_t(\succeq_c(\succeq))\ P_{F_I}(t)$.

Lemma 3.9 *Let $\succeq_t$ be an extension function. Then $\succeq_t^{st}$ is a TPO-status for F_I.*

Proof: (*Preserves quasi-orderings*) If $\succeq$ is a quasi-ordering on terms then $\succeq_c$ is a quasi-ordering on constants which is extended to a quasi-ordering on terms by $\succeq_t$.

(*Subterm founded*) $\succeq_t^{st}$ is subterm founded on $\mathcal{I}^2$, since for any nonatomic term only proper subterms are packed into constants, and $\succeq$ is only queried via $\succeq_c$.

(*Decreases infinite derivations*) Suppose there exists some infinite descending chain

$$s_1 \succ_t^{st}(\succeq)\ s_2 \succ_t^{st}(\succeq) \ldots$$

of nonatomic terms in $\mathcal{I}$. Then by definition of $\succ_t^{st}$

$$P_{F_I}(s_1) \succ_t(\succeq_c(\succeq))\ P_{F_I}(s_2) \succ_t(\succeq_c(\succeq)) \ldots,$$

and $c_{t_1} \succ_c (\succeq) c_{t_2} \succ_c (\succeq) \ldots$ by property 2 of $\succeq_t$, where c_{t_1} occurs in $P_{F_I}(s_j)$ for some $j \geq 1$. Finally $t_1 \succ t_2 \succ \ldots$ by the definitions of $\succeq_c (\succeq)$ and P, where t_1 is a proper subterm of s_j.

(*Strictly internally prepared for contexts*) This follows from $\succeq_t(\succeq_c)$ being strictly compatible with contexts.

(*Multiset properties*) (M0) follows directly from the definitions and from the fact that $\succeq_t$ extends $\succeq_c$. (M2) and (M4) follow by the subterm property of $\succeq_t$ and by transitivity. To show (M1) suppose $f(s_1, \ldots, s_m) \succeq_t^{st}(\succeq)\ t$, where t is atomic and thus $P_{F_I}(t) = c_t$. By property 4 a constant $c_{t'} \succeq_c c_t$ occurs in some $P_{F_I}(s_i)$, and possibly using the subterm property we get $s_i \succeq_t^{st}(\succeq)\ t$. For (M3) observe that c_s is greater than any constant occurring in some $P_{F_I}(t_i)$, and hence in $P_{F_I}(f(t_1, \ldots, t_n))$. □

Proposition 3.10 *Let $\succeq_t$ be an extension function such that $\succeq_t(\succeq_c)$ is total for any total quasi-ordering $\succeq_c$ on constants. Then $\succeq_t^{st}$ preserves totality.*

Lemma 3.11 *Let $\succeq_t$ be an extension function such that $\succeq_t(\succeq_c)$ is $(E \cup \sim_c)$-antisymmetric for any quasi-ordering $\succeq_c$ on constants. Then $\succeq_t^{st}$ preserves E-antisymmetry.*

Proof: Suppose $\succeq$ is E-antisymmetric and $s \sim_t^{st}(\succeq)\ t$. Then $P_{F_I}(s) \sim_t(\succeq_c)\ P_{F_I}(t)$ and hence $P_{F_I}(s) =_{E \cup \sim_c} P_{F_I}(t)$. Since $\succeq$ is assumed to be E-antisymmetric, $c_{s'} \sim_c c_{t'}$ implies $s' =_E t'$ for any atomic subterms s' and t' of s and t. Hence $s =_E t$. □

Lemma 3.12 *Suppose $F_I \supseteq F_E$, and let $\succeq_t$ be an extension function such that $\succeq_t(\succeq_c)$ is E-compatible for any quasi-ordering $\succeq_c$ on constants. Then $\succeq_t^{st}$ is prepared for E-compatibility.*

Proof: Observe that due to $F_I \supseteq F_E$ contexts at the root consisting of function symbols in E are left intact by packing. Hence for any instance of an equation in E packing both sides results again in an instance of the same equation. Thus for an equation $s \approx t$ in E we have $P_{F_I}(s\sigma) = s\sigma' \succeq_t(\succeq_c) t\sigma' = P_{F_I}(t\sigma)$ by E-compatibility of $\succeq_t$, where we define σ' by $x\sigma' = P_{F_I}(x\sigma)$ for all variables x in E. We conclude that $s\sigma \succeq_t^{st}(\succeq) t\sigma$ for any ground instance $s\sigma \approx t\sigma$ of an equation $s \approx t$ in E. □

The following theorem summarizes the construction of a TPO:

Theorem 3.13 *Let E be a set of collapse-free equations, let $\succeq_t$ be an extension function such that $\succeq_t(\succeq_c)$ is total, $(E \cup \sim_c)$-antisymmetric and E-compatible for any total quasi-ordering $\succeq_c$ on F_C, and let $\succeq_p$ be a precedence that is TPO-admissible for F_I. Then $\succeq_{tpo}(\succeq_p, \succeq_t^{st})$ is a total, E-antisymmetric and E-compatible simplification quasi-ordering.*

3.4 Examples of theory path orderings

In a trivial way any simplification quasi-ordering can be constructed as a TPO, by taking $F_I = F$ and letting $\succeq_t$ be the original ordering. Then $\mathcal{A}$ and F_C are empty, and properties (3) and (4) of an extension function become void. Being a simplification quasi-ordering, $\succeq_t$ satisfies properties (1) and (2).

On the other end of the spectrum is the lexicographic path ordering, which is obtained for $F_I = \emptyset$.

The simplest nontrivial example is an associative path ordering for a single associative and commutative symbol f. That is, we have $E = \mathrm{AC}(f)$ and $F_I = F_E = \{f\}$. Terms over $F_I \cup F_C$ are ordered according to the multiset of constants from F_C they contain. Formally, we associate a *complexity* $\kappa(t)$ to each term t over $F_I \cup F_C$, where

$$\kappa(t) = \begin{cases} \kappa(t_1) \cup \kappa(t_2) & \text{for } t = f(t_1, t_2), \\ \{c\} & \text{for } t = c \in F_C, \end{cases}$$

and define $\succeq_t(\succeq_c)$ by $s \succeq_t(\succeq_c) t$ if and only if $\kappa(s) \succeq_{mul}(\succeq_c) \kappa(t)$. We now show that this extension function satisfies the requirements of Theorem 3.13. Associativity and commutativity are collapse-free. Clearly $\succeq_t(\succeq_c)$ extends $\succeq_c$ and satisfies properties 3 and 4. The multiset extension of a quasi-ordering has the following properties:

$$M_1 \succeq_{mul} M_2 \text{ implies } N \cup M_1 \succeq_{mul} N \cup M_2 \tag{3.1}$$

$$M_1 \succ_{mul} M_2 \text{ implies } N \cup M_1 \succ_{mul} N \cup M_2 \tag{3.2}$$

$$M_1 \subsetneq M_2 \text{ implies } M_1 \succ_{mul} M_2. \tag{3.3}$$

Strict compatibility with F_I-contexts is a consequence of (3.1) and (3.2). The subterm property follows by (3.3). Since the multiset extension preserves well-foundedness, an infinite descending chain in $\succeq_t(\succeq_c)$ can only arise from an infinite descending chain in $\succeq_c$. Since we can restrict the ordering to the constants occurring in the infinite descending chain of multisets, there is an infinite descending chain of constants occurring in these multisets. Thus $\succeq_t$ is an extension function. If $\succeq_c$ is total then $\succeq_t(\succeq_c)$ is total, since the multiset extension preserves totality. AC-compatibility is obvious from the construction. Finally we show that the quasi-ordering $\succeq_t(\succeq_c)$ is $(\mathrm{AC} \cup \sim_c)$-antisymmetric. Suppose $s \sim_t(\succeq_c) t$. To take care of $\sim_c$ we select a representative $\mathrm{rep}(c)$ for each $\sim_c$-equivalence class in F_C. That is, $c \sim_c d$ if and only if $\mathrm{rep}(c) = \mathrm{rep}(d)$ for any two constants c and d

in F_C. We replace each constant c in s and t by its representative and obtain terms s' and t', respectively. Then $s' \sim_t(\succeq_c) t'$,

$$\kappa(s') = M_1 = \{c_1, \ldots, c_k\} \text{ and}$$
$$\kappa(t') = M_2 = \{d_1, \ldots, d_l\},$$

and $M_1 \sim_{mul}(\succeq_c) M_2$. Since $\sim_c$-equivalent constants are equal in s' and t', we even have $M_1 = M_2$ and hence $s' =_{\mathrm{AC}} t'$. Combining this with $s =_{\sim_c} s'$ and $t =_{\sim_c} t'$ we get $s =_{\mathrm{AC} \cup \sim_c} t$. We conclude that $\succeq_t(\succeq_c)$ is $\mathrm{AC} \cup \sim_c$-antisymmetric. Thus $\succeq_{tpo}(\succeq_p, \succeq_t^{st})$ is a total, E-antisymmetric and E-compatible simplification quasi-ordering.

4

Superposition for Convergent Theories

In this chapter we show how to systematically develop a refutationally complete inference system for a theory given by a convergent term rewriting system.

4.1 The term rewriting system

We require that a theory is represented by a ground term rewriting system T that is convergent modulo an equational theory E. That is, T is terminating and Church-Rosser modulo E. Then for any equational proof $s \overset{*}{\Leftrightarrow}_{T \cup E} t$ there exists a valley proof

$$s \overset{*}{\Rightarrow}_T s' \overset{*}{\Leftrightarrow}_E t' \overset{*}{\Leftarrow}_T t.$$

To avoid explicitly mentioning E-matching everywhere we assume that it is included in T. That is, $T = E \backslash T'$ for some term rewriting system T'. We assume a fixed set of function symbols F. A function symbol f is *free* in T if there exists a possibly nonground term rewriting system $\widehat{T}$ such that $T = gnd(\widehat{T})$ and f does not occur in $\widehat{T}$. Function symbols which are not free are called *interpreted*. The set of interpreted function symbols is denoted by F_T. A term with a free function symbol at the root position is called T-*atomic* Atomic terms will be denoted by α. We let $T_1 = T \cup E \cup \mathrm{Eq}$ denote the logical contents of T, where the rules in T are understood as equations and Eq is the first-order axiomatization of equality for F.

For example, consider the theory $\mathrm{ACU}(+, 0)$ of commutative monoids. For this case we let $E = \mathrm{AC}(+)$ and $T = E \backslash gnd(\widehat{T})$ where $\widehat{T} = \{x + 0 \Rightarrow x\}$. Let $F = \{f, a, +, 0\}$. Then f and a are free and $+$ and 0 are interpreted function symbols, and $f(a)$ as well as $f(a + 0)$ are atomic terms.

Furthermore, we require an E-antisymmetric and E-compatible simplification quasi-ordering $\succeq_T$ that is total on ground terms such that $\succ_T$ contains T. Then $s \sim_T t$ if and only if $s =_E t$ for ground terms s and t. We will usually omit the subscript T as the ordering used will be clear from the context. An atomic term α is called a *maximal atomic term* in s if $s = u[\alpha, \alpha_1, \ldots, \alpha_n]$ where $n \geq 0$, $\alpha_1, \ldots, \alpha_n$ are atomic, u is an F_T-context, and $\alpha \succeq \alpha_i$ for $i = 1 \ldots n$.

For the example we let $F_I = F_{\mathrm{AC}} = \{+\}$ and use the precedence

$$f \succ_p a \succ_p 0 \succ_p (+),$$

which is TPO-admissible for F_I, together with the AC-status $\succeq^{st}_{\mathrm{AC}(+)}$. Then the TPO $\succeq_{tpo}(\succeq_p, \succeq^{st}_{\mathrm{AC}(+)})$ is a suitable term ordering. This is an APO that uses multiset status for flattened $+$-contexts and lexicographic status for the other function symbols. Then for instance $f(a) \succ f(0+0)+a$, and $f(a+0)$ is a maximal atomic subterm in $f(a+0)+(f(0+a)+a)$.

4.2 The symmetrization function

The symmetrization function is at the heart of our approach. It maps some given rewrite rule into a set of rewrite rules that encodes a special rewrite relation appropriate for the theory. That is, we will construct terminating term rewriting systems of the form

$$T \cup \bigcup_{s \approx t} \mathcal{S}_T(s \approx t),$$

where the set of rules $\mathcal{S}_T(s \approx t)$ is designed such that $s \approx t$ becomes true and as much as possible of T and the equality axioms are preserved. It turns out that this works well for all axioms except transitivity, which causes problems for certain theories.

We start by the notion of a set of rules being (strongly) symmetrized. Being symmetrized is a rather technical notion that is required by our general superposition calculus. It amounts to the convergence of critical pairs that involve a rule from T or equation from E, and hence validity of the corresponding instances of transitivity. The notion of a strongly symmetrized set of rewrite rules becomes important when we later instantiate the general framework by specific theories. It will allow to manipulate equational proofs by normalizing the terms in the proof (see Section 4.9).

A set of rewrite rules S is *symmetrized* with respect to T modulo E if for all peaks $t_1 \Leftarrow_T t \Rightarrow_S t_2$ and for all cliffs $t_1 \Leftrightarrow_E t \Rightarrow_S t_2$ we have $t_1 \Downarrow_{T \cup S} t_2$.

The set S is called *strongly symmetrized* with respect to T modulo E if S can be partitioned into sets S_i, $i \in I$, such that $T \cup S_i$ is convergent modulo E for all $i \in I$.

Proposition 4.1 *If a set of rewrite rules S is strongly symmetrized with respect to T modulo E then S is symmetrized with respect to T modulo E.*

Proof: Consider some peak $t_1 \Leftarrow_T t \Rightarrow_S t_2$ or cliff $t_1 \Leftrightarrow_E t \Rightarrow_S t_2$. The rule from S is in some S_i, and by convergence of $T \cup S_i$ we get the desired valley proof. □

Note that S being strongly symmetrized implies that peaks of the form $t_1 \Leftarrow_{S_i} t \Rightarrow_{S_i} t_2$ converge, which is not guaranteed if S is symmetrized but not strongly symmetrized. However, this is still much weaker than convergence, as peaks of the form $t_1 \Leftarrow_{S_i} t \Rightarrow_{S_j} t_2$ need not converge for $i \neq j$.

Our goal is to derive (strongly) symmetrized sets of rules directly for some given equation, so that the equation becomes true in the rewrite system. We break this into two steps. First the equation is brought into a certain T-normal form by simplification, and then for any such equation a symmetrized set is obtained by applying a symmetrization function. For now we only assume that a set Norm_T of equations in *T-normal form* is given, and postpone the discussion of simplification. We continue by discussing symmetrization functions.

A *(strong) symmetrization function $\mathcal{S}_T$* (for T) maps any equation $l \approx r$ in T-normal form to a (strongly) symmetrized set of rewrite rules $\mathcal{S}_T(l \approx r)$ such that

$$T_1 \cup \{l \approx r\} \models \mathcal{S}_T(l \approx r) \tag{4.1}$$

$$l \Downarrow_{T \cup \mathcal{S}_T(l \approx r)} r \tag{4.2}$$

$$\mathcal{S}_T(l \approx r) \subseteq (\succ) \tag{4.3}$$

$$l' \succeq l \text{ for any } l' \Rightarrow r' \text{ in } \mathcal{S}_T(l \approx r) \tag{4.4}$$

(4.1) ensures soundness, (4.2) ensures that $l \approx r$ becomes true, (4.3) ensures termination, and (4.4) ensures that terms smaller than l cannot be rewritten by $\mathcal{S}_T(l \approx r)$. We call a rule $l' \Rightarrow r'$ in $\mathcal{S}_T(l \Rightarrow r) \setminus \{l \Rightarrow r\}$ an *extension* (of $l \Rightarrow r$). The symmetrization function is extended to sets of equations in T-normal form by

$$\mathcal{S}_T(R) = \bigcup_{l \approx r \in R} \mathcal{S}_T(l \approx r).$$

Assumption 4.2 *We assume from now on that $\mathcal{S}_T$ is a symmetrization function for T modulo E.*

To obtain a symmetrization function one considers critical peaks of the form $t_1 \Leftarrow_T s \Rightarrow_{S_i} t_2$, in a way very similar to Knuth-Bendix completion. To obtain a strong symmetrization function one also has to consider critical peaks of the form $t_1 \Leftarrow_{S_i} s \Rightarrow_{S_i} t_2$. For the commutative theories that we consider here it turns out that the symmetrization function obtained by considering the first kind of peaks also makes the second kind converge. Thus the strong symmetrization property requires no extra effort in these cases. Without commutativity, however, an equation may have nontrivial overlaps with variants of itself. It is infeasible to derive a strong symmetrization function in that case, hence for instance Le Chenadec (1986) uses ordinary symmetrization for nonabelian groups.

For our example theory of commutative monoids we let an equation be in T-normal form if both sides are irreducible with respect to T. Then $f(a) \approx a+a$ is in T-normal form, and $f(a+0) \approx a$ is not. The strong symmetrization function for this theory maps any nontrivial equation $l \approx r$ with $l \succ r$ to the set consisting of $l \Rightarrow r$ and the AC-extensions of $l \Rightarrow r$. Trivial equations $s \approx t$ where $s =_{\mathrm{AC}} t$ are mapped to the empty set. That is,

$$\begin{aligned}
&\mathcal{S}_T(s \approx t) = \emptyset && \text{if } s =_{\mathrm{AC}} t,\\
&\mathcal{S}_T(l \approx r) = \{l \Rightarrow r\} && \text{if } l \succ r \text{ and } l \text{ is not a proper sum},\\
&\mathcal{S}_T(l \approx r) = \{l \Rightarrow r\} \cup gnd(\{x + l \Rightarrow x + r\}) && \text{if } l \succ r \text{ and } l \text{ is a proper sum}.
\end{aligned}$$

The sets of rules in the range of this function are convergent modulo AC, and the function satisfies the other properties of a symmetrization function. Hence it is a strong symmetrization function.

4.3 Ordering literals and clauses

To extend the term ordering $\succeq_T$ to equations, literals and clauses we assign to each of these a *complexity* c. For an equation $s \approx t$, the complexity is the multiset $\{s, t\}$ and equations are compared by the multiset extension of $\succ$. For literals we let

$$c(s \approx t) = \{\{s\}, \{t\}\} \tag{4.5}$$

$$c(s \not\approx t) = \{\{s, t\}\} \tag{4.6}$$

and $\succ$ on literals is the two-fold multiset extension of $\succ$ on terms applied to these complexities. This has the effect that the ordering on literals is the lexicographic combination of $\succ$ on the maximal term, the ordering $- \succ +$ on the polarity of the literal and $\succ$ on the minimal term. For a clause $C = L_1 \vee L_n$ that is not an instance of transitivity we let

$$c(C) = \langle \{c(L_1), \ldots, c(L_n)\}, \emptyset \rangle \tag{4.7}$$

That is, the complexity of a nontransitivity clause is the multiset of the complexities of its literals. The *middle term* of a ground instance

$$D = t_1 \not\approx s \lor s \not\approx t_2 \lor t_1 \approx t_2$$

of transitivity is s. In this case we let

$$c(D) = \langle \{\{\{s\}\}\}, \{t_1, t_2\} \rangle \tag{4.8}$$

Then the quasi-ordering on clauses is

$$(\succeq) = \succeq_{lex}(\succeq_{mul}(\succeq_{mul}(\succeq_{mul}(\succeq))), \succeq_{mul}(\succeq)), \tag{4.9}$$

where the inner $\succeq$ is the quasi-ordering on terms. That is, the ordering on clauses is the lexicographic combination of the three-fold multiset extension of the term ordering and the multiset extension of the term ordering, applied to the complexities. By this definition transitivity instances with a middle term s are immediately below nontransitivity clauses with maximal term s in the term ordering. We call the middle term of transitivity instances and the maximal term of other clauses the *dominating term* of the clause, since it dominates the term ordering. Note that the ordering is not only E-compatible. The extensions to equations and literals are compatible with the symmetry of equality, and the extension to clauses is compatible with associativity and commutativity of $\lor$.

4.4 Candidate models

In this section we define a model functor I that maps any set N of ground clauses to an interpretation I_N. We show that I_N satisfies the theory and the equality axioms except for transitivity.

The construction of the interpretation extends the standard one by Bachmair and Ganzinger (1998a) in several respects.

Firstly, rewriting is modulo E. Secondly, the built-in term rewriting system T is always included when constructing the interpretation. This ensures that these interpretations satisfy T. Thirdly, we have the additional restriction that a clause can be productive only if the rule it produces is in T-normal form. Finally, the term rewriting systems are built from symmetrizations of rules, which ensures that they are is always symmetrized.

A ground clause $C \lor s \approx t$ is called *reductive* for $s \Rightarrow t$ if $s \approx t$ is strictly maximal in C and $s \succ t$. Only reductive clauses can contribute to an interpretation. Given a set N of ground clauses, we let N_C be the set of ground clauses in N which are smaller than C. For any set N of ground clauses we inductively define a set R_N of ground rules, a symmetrized set $S_N = \mathcal{S}_T(R_N)$ of ground rules, and the corresponding interpretation $I_N = (T \cup S_N)^{\Downarrow}$. We may regard R, S and I as functions which map sets of clauses to sets of rewrite rules or equations. A rule $\{l \Rightarrow r\}$ is in R_N if there exists a clause $C = C' \lor l \approx r$ in N such that (i) C is false in I_{N_C}, (ii) C is reductive for $l \Rightarrow r$, (iii) $l \Rightarrow r$ is in T-normal form, (iv) l is irreducible by S_{N_C}, and (v) C' is false in $(T \cup S_{N_C} \cup \mathcal{S}_T(l \Rightarrow r))^{\Downarrow}$. In this case we say that C *produces* $l \Rightarrow r$ in R_N, or that C is *productive*. The set R_N is well-defined, since for any ground clause C only the interpretation for smaller clauses in N_C determines whether C produces a rule. Where N is clear from the context we write R_C for R_{N_C}, S_C for S_{N_C} and I_C for I_{N_C}.

Lemma 4.3 *Let $C = C' \lor l \approx r$ be a clause that produces $l \Rightarrow r$ in R_N. Then C' is false in I_N.*

Proof: Since C is false in I_{N_C}, the subclause C' is also false. By condition (v) C' is also false in $(T \cup S_{N_C} \cup \mathcal{S}_T(l \Rightarrow r))^{\Downarrow}$. Since the maximal term of C' cannot be greater than l, and since (iv) in combination with left-minimality prevents that clauses greater than C produce a rule with left-hand side smaller than or equal to l, equations in C' have the same truth-value in I_C and in I_N. Hence C' is false in I_N. □

We let $T_0 = \mathrm{Refl} \cup \mathrm{Symm} \cup \mathrm{Mon} \cup E \cup T$.

Lemma 4.4 *Let N be a set of ground clauses. Then $I_N \models T_0$.*

Proof: The convergent term rewriting system T modulo E is included in the rules used to define I_N, hence $s \Downarrow t$ for all axioms in $T \cup E$. Reflexivity, symmetry and monotonicity follow from the definition of $\Downarrow$. □

It remains to consider instances of transitivity and clauses in N. These are in general not true in I_N. For instance, for commutative rings there are cases where two extended rules overlap in such a way that the resulting critical pair does not converge. Then $S_C \cup T$ is not confluent and transitivity does not hold. We say that a clause C in $\mathrm{Trans} \cup N$ is a *counterexample* for I_N if C is false in I_N.

To illustrate these concepts we continue the commutative monoid example. Consider the set N of the ground clauses

$$\begin{aligned}
C_1 &= f(0) \approx 0 \\
C_2 &= f(0) + a \approx 0 \\
C_3 &= f(0) + f(0) \not\approx 0 \ \vee\ a \approx 0 \\
C_4 &= f(0+0) \approx f(0) \\
C_5 &= f(a) \approx f(0) \ \vee\ f(a) \approx 0 \\
C_6 &= f(a) + a \approx f(a) \\
C_7 &= f(a) + f(a) \approx f(0).
\end{aligned}$$

The clauses are listed in ascending order with respect to $\succ$. The interpretation I_N is constructed as follows. Clause C_1 produces the rule $f(0) \Rightarrow 0$. Clause C_2 is not productive, since $f(0)+a$ is reducible by $f(0) \Rightarrow 0$. Clause C_3 is not productive, since its positive literal is not maximal; hence C_3 is not reductive. Clause C_4 is not productive, since it is true in I_{C_4} and not in T-normal form. Clause C_5 is not productive, since adding the rule $f(a) \Rightarrow f(0)$ would also make $f(a) \approx 0$ true, which violates condition (v). Clauses C_6 and C_7 produce the rules $f(a) + a \Rightarrow f(a)$ and $f(a) + f(a) \Rightarrow f(0)$, respectively. Their symmetrizations contain the AC-extensions $f(a) + f(a) + a \Rightarrow f(a) + f(a)$ and $f(a) + f(a) + a \Rightarrow f(0) + a$, which thus are true in I_N when seen as equations. However, $f(a) + f(a) \approx f(0) + a$ is false in I_N, since $f(a) + f(a)$ reduces to 0 and $f(0) + a$ reduces to a. Thus we have the transitivity counterexample

$$f(a) + f(a) \not\approx f(a) + f(a) + a \ \vee\ f(a) + f(a) + a \not\approx f(0) + a \ \vee\ f(a) + f(a) \approx f(0) + a.$$

The clauses C_2, C_3 and C_5 are the counterexamples in N.

4.5 Redundancy of clauses and simplification

We will later need to refer to the specific construction of candidate models when we prove that certain clauses or inferences are redundant. In particular, we need that candidate models are built from (strongly) symmetrized sets of rewrite rules S_N, and we need to refer to the presence of certain rewrite rules in R_N. We achieve this in a nice and coherent way by defining a special notion of consequence that takes into account only interpretations constructed by the model functor I, and by introducing a new atomic formula $s \Rightarrow t$ that is true in such an interpretation I_N whenever the rule $s \Rightarrow t$ is in R_N. Note that the R_N corresponding to I_N will always be known from the context via the set N of clauses. These atoms will be used only as unit clauses, and we will refer to them as rewrite rules. For sets of clauses or rewrite rules N_1 and N_2 we say that N_2 is an *I-consequence* of N_1, in symbols $N_1 \models_I N_2$, if $I_M \models N_1$ implies $I_M \models N_2$ for all sets of ground clauses M. Lemma 4.4 can then be rephrased as $\models_I T_0$.

The distinction between $\models$ and $\models_I$ is that for the latter only validity in interpretations from the set $\{I_N \mid N \text{ set of ground clauses}\}$ is considered. Basic properties of classical logic like reflexivity, transitivity, monotonicity and the deduction theorem also hold for $\models_I$. The following observation allows to approximate $\models_I$:

Proposition 4.5 *Let M be a set of ground clauses or rewrite rules, let M' be the sets resulting from M by replacing any rewrite rule $l \Rightarrow r$ by a unit clause $l \approx r$, and let N be a set of ground clauses such that $M' \cup T_0 \models N$. Then $M \models_I N$.*

Proof: Observe that $M \models_I M'$ and $\models_I T_0$ hold and that $\models_I$ is transitive. □

After having proved refutational completeness of our calculus we will be able to show that $M \models_I N$ implies $M' \cup T_1 \models N$ (Proposition 4.21 on page 49), which approximates $\models_I$ from above.

Let C be some ground clause. We write Trans_C for the set of ground instances of transitivity in Trans which are smaller than C.The middle term of such an instance of transitivity is smaller than or equal to the dominating term of C. Then C is *redundant* (with respect to T) in a set of ground clauses N if

$$N_C \cup \text{Trans}_C \models_I C.$$

A (possibly nonground) clause is called *redundant* in a set of clauses N if all its ground instances are redundant in the set of ground instances of N. A clause is called *redundant* if it is redundant in $\emptyset$. A clause that is redundant in N cannot be the minimal counterexample for I_N, because some smaller clause in $N_C \cup \text{Trans}_C$ would have to be a counterexample for I_N as well. Note that by Proposition 4.5 we can use

$$N_C \cup \text{Trans}_C \cup T_0 \models C$$

as a sufficient criterion for redundancy. This criterion corresponds to the notion of redundancy used by Bachmair and Ganzinger (1998a).

As an example of a redundancy criterion consider the following lemma. It corresponds to the well-known fact that peaks with a reducible middle term are redundant (Buchberger 1979, Winkler and Buchberger 1983, Kapur, Musser and Narendran 1988).

Lemma 4.6 *Let $D = t_1 \not\approx s \vee s \not\approx t_2 \vee t_1 \approx t_2$ be a ground instance of transitivity such that the middle term s is reducible by T. Then D is redundant.*

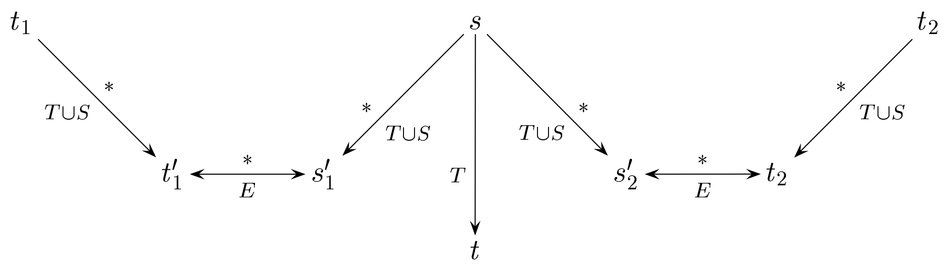


Figure 4.1: Transitivity with a reducible middle term

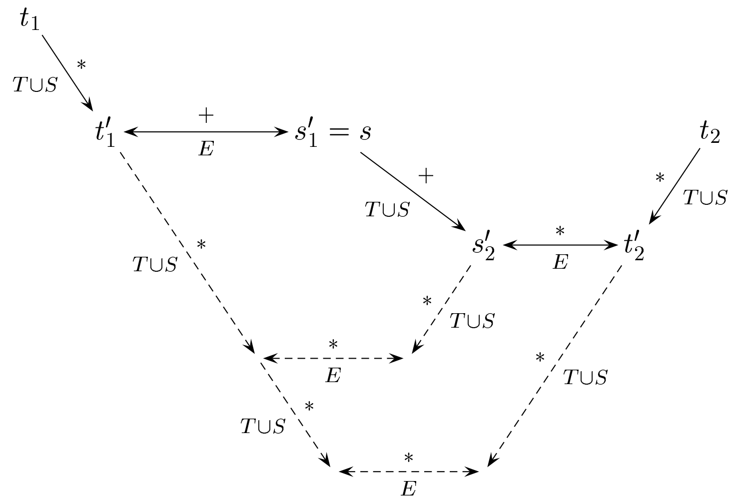


Figure 4.2: Transitivity with a nonpeak middle term

Proof: Let I_N be some interpretation that satisfies Trans_D and suppose that $t_1 \approx s$ and $s \approx t_2$ are true in I_N. That is, there exists valley proofs $t_1 \Downarrow_{T \cup S_N} s$ and $s \Downarrow_{T \cup S_N} t_2$ and we have the situation of Figure 4.1. We have to find a valley proof $t_1 \Downarrow t_2$ to show $I_N \models t_1 \approx t_2$.

(1) If s is not at a peak then either the proof is already a valley proof or coherence can be used at s, as sketched in Figure 4.2.

(2) Otherwise it suffices to cut off the peak as indicated in Figure 4.3, in order to obtain a proof $t_1' \overset{*}{\Leftrightarrow} t_2'$ that stays below s and hence gives rise to a valley proof $t_1 \Downarrow t_2$. $\square$

Based on our notion of redundancy, we say that a ground clause D is a *simplification* (with respect to T) of a ground clause C if $\{C\} \cup T_1 \models D$ and C is redundant in $\{D\}$. That is, $C \succ D$, $\{C\} \cup T_1 \models D$, and $\{D\} \cup \mathrm{Trans}_C \models_I C$. We write

$$\frac{C}{D}\!\!\rightarrow$$

or $C \Longrightarrow D$ to indicate that C can be simplified to D.

We will now define a slightly stronger notion of simplification that most of our simplification rules satisfy. This allows to combine simplifications more freely, disregarding

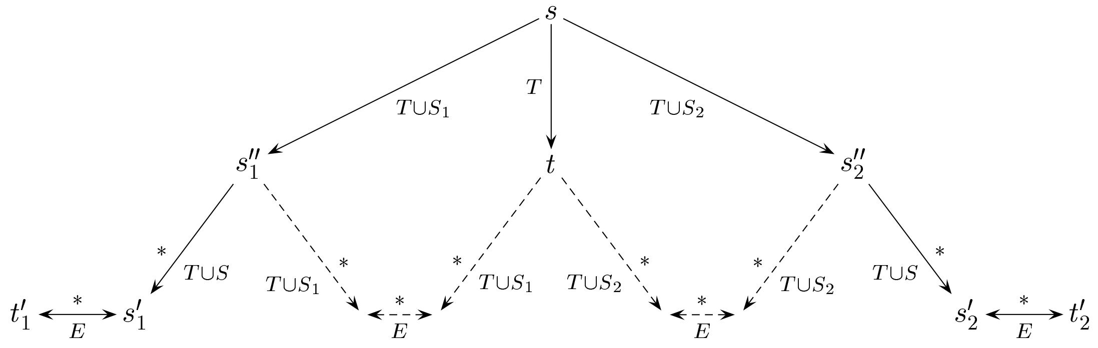


Figure 4.3: Transitivity with a peak middle term that is reducible

their direction as long as the bounds on transitivity are obeyed and the final clause is smaller than the original. Given ground clauses C_1, C_2 and D we write $C_1 \Longleftrightarrow_D C_2$ if $\{C_1\} \cup \mathrm{Trans}_D \models_I C_2$ and $\{C_2\} \cup \mathrm{Trans}_D \models_I C_1$. If additionally $C_1 \succ C_2$ we write $C_1 \Longleftrightarrow\!\!\!\!\!\!\!\!\!\!\rightarrow_D C_2$. From the definition it is clear that $\Longleftrightarrow_D$ is an equivalence relation and $\Longleftrightarrow\!\!\!\!\!\!\!\!\!\!\rightarrow_D$ is a strict partial ordering on ground clauses. Furthermore, $D_1 \succeq D_2$ implies $\Longleftrightarrow_{D_2} \subseteq \Longleftrightarrow_{D_1}$. Finally, $C \Longleftrightarrow\!\!\!\!\!\!\!\!\!\!\rightarrow_C D$ implies that C can be simplified to D.

We say that a set of simplifications on literals is *uniform with respect to polarity* if $L_1 \Longrightarrow L_2$ is in the set if and only if $L'_1 \Longrightarrow L'_2$ is in the set, where L'_i is the negation of L_i for $i = 1, 2$.

Lemma 4.7 *Let S be a set of simplification rules on literals that is uniform with respect to polarity. Then for any rule $L_1 \Longrightarrow L_2$ in S we have $L_1 \Longleftrightarrow\!\!\!\!\!\!\!\!\!\!\rightarrow_{L_1} L_2$.*

Proof: Suppose $L_1 \Longrightarrow L_2$ is in S. Then also $L'_1 \Longrightarrow L'_2$ is in S and both are simplifications. Hence $\{L_2\} \cup \mathrm{Trans}_{L_1} \models_I L_1$ and $\{L'_2\} \cup \mathrm{Trans}_{L'_1} \models_I L'_1$, which is equivalent to $\{L_1\} \cup \mathrm{Trans}_{L'_1} \models_I L_2$. Moreover, since L_1 and L'_1 are equal except for polarity and Trans_C depends only on the maximal term of C for any clause C, $\mathrm{Trans}_{L_1} = \mathrm{Trans}_{L'_1}$. We conclude $L_1 \Longleftrightarrow\!\!\!\!\!\!\!\!\!\!\rightarrow_{L_1} L_2$. □

To extend simplification rules from literals to clauses we have the following lemma:

Lemma 4.8 *Let C, D and C' be ground clauses.*

1. *If C simplifies to D then $C \vee C'$ simplifies to $D \vee C'$.*
2. *If $C \Longleftrightarrow\!\!\!\!\!\!\!\!\!\!\rightarrow_C D$ then $C \vee C' \Longleftrightarrow\!\!\!\!\!\!\!\!\!\!\rightarrow_C D \vee C'$.*

Proof: This follows immediately from $\{C\} \models_{\mathcal{M}} D$ implying $\{C \vee C'\} \models_{\mathcal{M}} D \vee C'$, where $\mathcal{M}$ is an arbitrary class of models. □

We now show that rewriting with T is a simplification.

Lemma 4.9 *If $L_1 \Rightarrow_T L_2$ then L_1 simplifies to L_2.*

Proof: $L_1 \overset{+}{\Rightarrow}_T L_2$ implies $L_1 \succ L_2$ and $\{L_1\} \cup T_1 \models L_2$. To show $\{L_2\} \cup \mathrm{Trans}_{L_1} \models_I L_1$ consider some interpretation I_N that satisfies L_2 and Trans_{L_1} and let $L_i = [\neg](s_i \approx t_i)$ for $i = 1, 2$.

(1) Suppose L_1 and L_2 are positive. Then $s_2 \Downarrow_{T \cup S_N} t_2$ and hence $s_1 \overset{*}{\Rightarrow}_T s_2 \Downarrow t_2 \overset{*}{\Leftarrow} t_1$ and L_1 is also true in I_N.

(2) Otherwise L_1 and L_2 are negative. Suppose L_1 is false in I_N, that is, $s_1 \Downarrow_{T \cup S_N} t_1$. Then $s_2 \overset{*}{\Leftarrow}_T s_1 \Downarrow t_1 \overset{*}{\Rightarrow}_T t_2$. Since all terms of this proof are bounded by s_1 or t_1, we get $s_2 \Downarrow_{T \cup S_N} t_2$ and L_2 would be false in I_N as well, a contradiction. So L_1 is true in I_N. □

Lemma 4.10 *If $C \Rightarrow_T D$ then $C \Longleftrightarrow\!\!\!\!\!\Rightarrow_C D$.*

Proof: By combining Lemmas 4.9, 4.7 and 4.8. □

By this lemma the following is a simplification rule:

T-Rewriting $\qquad \dfrac{C}{D}\!\!\rightarrow$

if $C \Rightarrow_T D$.

Apart from using rewriting directly as a simplification we can also combine rewrite sequences and simplifications to obtain new or more general simplifications. This is useful when showing that certain transformations are simplifications. It allows to extend simplifications between clauses where the terms in question are irreducible with respect to T to clauses where this is not the case. This uniformity is useful in particular for lifting, where instances may be T-reducible.

Lemma 4.11 *Let C_1, C_2, D_1 and D_2 be ground clauses such that $C_1 \overset{*}{\Rightarrow}_T C_2$, $D_1 \overset{*}{\Rightarrow}_T D_2$ and $C_1 \succ D_1$ and suppose that C_2 simplifies to D_2 with respect to T. Then C_1 simplifies to D_1 with respect to T.*

Proof: Clearly $C_1 \Rightarrow_{C_1} C_2$ and $D_1 \Rightarrow_{C_1} D_2$. Chaining these together with $\{D_2\} \cup \mathrm{Trans}_{C_1} \models_I C_2$ we conclude $\{D_1\} \cup \mathrm{Trans}_{C_1} \models_I C_1$. □

Remember that the symmetrization function is only defined on equations in T-normal form. Equations not in this form need to be simplified before symmetrization can be applied. However, we want to restrict simplifications as much as possible, since ground simplifications become inferences when they are lifted. We formalize this by assuming that there exists a function Simp_T which maps ground literals to sets of ground literals, such that L' is a simplification of L for all $L' \in \mathrm{Simp}_T(L)$. We say that Simp_T is *admissible* with respect to $\mathcal{S}_T$ if

$$\{L \mid \mathrm{Simp}_T(L) = \emptyset\} \subseteq \mathrm{Norm}_T,$$

where Norm_T is the set of T-normal forms on which the symmetrization function $\mathcal{S}_T$ is defined. That is, any equation for which no symmetrization is given by $\mathcal{S}_T$ must be simplifiable by Simp_T. Since $\succ$ is well-founded, it suffices to nondeterministically apply simplifications in Simp_T to eventually reach a literal in T-normal form. Note that T-normal forms of literals need not be unique. Such a requirement would lead to unnecessary simplifications of the smaller side of equations.

Assumption 4.12 *We assume from now on that* Simp_T *is admissible with respect to* $\mathcal{S}_T$.

The definitions of $\mathcal{S}_T$ and Simp_T impose certain properties on Norm_T. Since any literal can be simplified to some literal in T-normal form, and since simplification preserves T_1-equivalence, any literal has a T_1-equivalent T-normal form. Moreover, for strong symmetrization functions the requirement that l is minimal among the left-hand sides of rules in $\mathcal{S}_T(l \Rightarrow r)$ translates into the requirement that T-normal forms of equations are *left-minimal*. That is, l is minimal among the greater sides of T_1-equivalent equations. For if this were not the case, say there exists $l' \approx r'$ with $l' \succ r'$ and $l \succ l'$ then l' must be reducible by $\mathcal{S}_T(l \approx r)$, by some rule with left-hand side smaller than l.

For the example of commutative monoids Simp_T consists of rewriting with T. Additionally, we can impose some strategy such as leftmost-innermost rewriting.

4.6 The inference system

We present a ground inference system that is based on the parameters introduced in the previous sections, namely the term rewriting system T, the ordering $\succ$, the set of T-normal forms Norm_T, the symmetrization function $\mathcal{S}_T$ and the simplification function Simp_T.

We assume that in each ground clause a literal is selected; either some arbitrary negative literal, or a positive literal that is maximal in the entire clause. An *inference system* is a set of inferences. Each *inference* has a *main premise* C, *side premises* $C_1, \ldots, C_n$, and a conclusion D. The main premise may either be a clause supposed to be from N, then we write

$$\frac{C_1 \quad \ldots \quad C_n \qquad C}{D}$$

for the inference, with the main premise at the right. Or the main premise may be an instance of transitivity, then we omit it and write

$$\frac{C_1 \quad \ldots \quad C_n}{D},$$

for the inference. In this case we state the main premise in the text. This allows uniform definitions of reduction property and redundancy of inferences. An inference is *strictly decreasing* if the conclusion is smaller than the main premise in the clause ordering. All inferences that we present are sound and strictly decreasing.

We let Sup_T be the set of the following inferences:

Let $l_1 \Rightarrow r_1$ and $l_2 \Rightarrow r_2$ be rules in Norm_T and $S_i = \mathcal{S}_T(l_i \Rightarrow r_i)$ for $i = 1, 2$. An *extension peak* between $l_1 \Rightarrow r_1$ and $l_2 \Rightarrow r_2$ with respect to T is a rewrite sequence

$$r_1' \Leftarrow_{S_1} l_1[l_2'] \Rightarrow_{S_2} l_1'[r_2']$$

such that $l_i' \Rightarrow r_i'$ is a rule in S_i for $i = 1, 2$, l_1 is irreducible by $T \cup S_2$, and l_2 is irreducible by $T \cup S_1$.

T-Extension Superposition $$\frac{l_1 \approx r_1 \lor C_1 \qquad l_2 \approx r_2 \lor C_2}{r_1' \approx l_1'[r_2'] \lor C_1 \lor C_2}$$

if (i) $l_i \approx r_i$ is selected in $l_i \approx r_i \lor C_i$ and in T-normal form for $i = 1, 2$, and (ii) there exists an extension peak $r_1' \Leftarrow_{\mathcal{S}_T(l_1 \approx r_1)} l_1'[l_2'] \Rightarrow_{\mathcal{S}_T(l_2 \approx r_2)} l_1'[r_2']$ between $l_1 \Rightarrow r_1$ and $l_2 \Rightarrow r_2$.

The transitivity instance corresponding to the peak,

$$r'_1 \not\approx l'_1[l'_2] \vee l'_1[l'_2] \not\approx l'_1[r'_2] \vee r'_1 \approx l'_1[r'_2],$$

is the main premise of this inference. The explicit premises are side premises.

T-Theory Simplification $\dfrac{L \vee C}{L' \vee C}$

if (i) L is selected in $L \vee C$, and (ii) $L' \in \mathrm{Simp}_T(L)$.

T-Reflexivity Resolution $\dfrac{p \not\approx q \vee C}{C}$

if (i) $p \not\approx q$ is in T-normal form and selected in $p \not\approx q \vee C$, and (ii) $p =_E q$.

T-Equality Factoring $\dfrac{s \approx t \vee s' \approx t' \vee C}{t \not\approx t' \vee s' \approx t' \vee C}$

if (i) $s \approx t$ is in T-normal form and selected in $s \approx t \vee s' \approx t' \vee C$, (ii) $s =_E s'$.

The single premise of Theory Simplification, Reflexivity Resolution and Equality Factoring is their main premise, they have no side premises.

T-Superposition $\dfrac{l \approx r \vee D \qquad [\neg](s[l''] \approx t) \vee C}{[\neg](s[r'] \approx t) \vee C \vee D}$

if (i) $l' \Rightarrow r'$ is the rule with minimal right-hand side among the rules with left-hand side l' in $\mathcal{S}_T(l \approx r)$, (ii) $l' =_E l''$, (iii) $[\neg](s[l''] \approx t)$ is selected in $[\neg](s[l''] \approx t) \vee C$ and in T-normal form, and (iv) $l \approx r$ is selected in $l \approx r \vee D'$ and in T-normal form.

Superposition has the main premise $[\neg](s[l'] \approx t) \vee C$ and the side premise $l \approx r \vee D'$. The restriction to rules with minimal right-hand side in (i) allows to use highly nondeterministic symmetrization functions that contain many rules with the same left-hand side. These are convenient in certain confluence proofs (see Sections 7.2 and 8.2), but would lead to many unnecessary Superposition inferences without this restriction.

An inference with main premise C, conclusion D and side premises $C_1, \ldots, C_n$, where each side premise $C_i = C'_i \vee l_i \approx r_i$ is reductive for $l_i \Rightarrow r_i$, is *redundant* in N if

$$N_C \cup \mathrm{Trans}_C \cup \{l_i \Rightarrow r_i \mid i = 1, \ldots, n\} \cup \{\neg C'_i \mid i = 1, \ldots, n\} \models_I D.$$

Here we exploit that side premises arise from productive clauses. Hence each side premise C_i has the form $C'_i \vee l_i \approx r_i$, such that it is reductive for $l_i \Rightarrow r_i$ and $l_i \Rightarrow r_i$ is in T-normal form. We may assume that $l_i \Rightarrow r_i$ is in R_N and that C'_i is false in I_N. An inference is called *redundant* if it is redundant in $\emptyset$.

Let C be the minimal counterexample for I_N and let π be an inference with main premise C, conclusion D and side premises $C_1, \ldots, C_k$ such that $C \succ D$, and each side premise C_i is smaller than C, has the form $C_i = l_i \approx r_i \vee C'_i$ and is reductive for $l_i \Rightarrow r_i$. We say that π *reduces* C (with respect to I_N) if

$$I_N \models \neg D \wedge l_1 \Rightarrow r_1 \wedge \ldots \wedge l_k \Rightarrow r_k \wedge \neg C'_1 \wedge \ldots \wedge \neg C'_k.$$

An inference system Sup has the *reduction property for counterexamples* (with respect to I) if Sup contains an inference that reduces C with respect to I_N for any set N of ground clauses such that I_N has a minimal counterexample $C \neq \bot$.

Lemma 4.13 (Extension Superposition) *Let N be a set of ground clauses such that N does not contain the empty clause. Suppose that the minimal counterexample C for I_N is an instance of transitivity. Then Sup_T contains an Extension Superposition inference that reduces C.*

Proof: Let C be the minimal counterexample and let s be its middle term. Since C is minimal, instances of transitivity with smaller middle terms are true in I_N. By Lemma 2.13 this implies that $T \cup S_C$ is Church-Rosser modulo E below s, but that there exists some peak $t_1 \Leftarrow s \Rightarrow t_2$ such that t_1 and t_2 do not converge and $t_1 \approx t_2$ is false in I_N. As T is convergent and S is symmetrized modulo E with respect to T, all peaks involving T and all cliffs with E converge, so both rules used in the peak are from S. If the rewrite steps in the peak were in parallel positions of s then t_1 and t_2 would converge, which is not the case. Let $l'_1 \Rightarrow r'_1$ and $l'_2 \Rightarrow r'_2$ be the rules from S_C used in the peak. For $i = 1, 2$ the rule $l'_i \Rightarrow r'_i$ is from some symmetrization $\mathcal{S}_T(l_i \Rightarrow r_i)$ where $l_i \Rightarrow r_i$ is a rule in R_C that has been produced by some clause $C_i = l_i \approx r_i \vee C'_i$. If we suppose without loss of generality that $l_1 \succ l_2$ then l_1 is irreducible by $\mathcal{S}_T(l_2 \Rightarrow r_2)$ because this is a condition for D_1 being productive, and l_2 is irreducible by $\mathcal{S}_T(l_1 \Rightarrow r_1)$ because $l_1 \succ l_2$ and l_2 is minimal among the left-hand sides in $\mathcal{S}_T(l_2 \Rightarrow r_2)$. Hence this is an extension peak of the form

$$t_1 = s[r'_1] \Leftarrow s[l'_1[l'_2]] \Rightarrow s'[r'_2] = t_2.$$

Since C is the minimal counterexample, the context must be empty, and the peak has the form

$$r'_1 \Leftarrow l'_1[l'_2] \Rightarrow l_1[r'_2].$$

For such a peak Sup_T contains the Extension Superposition inference

$$\frac{l_1 \approx r_1 \vee C'_1 \qquad l_2 \approx r_2 \vee C'_2}{r'_1 \approx l'_1[r'_2] \vee C'_1 \vee C'_2}$$

where C'_i is false in I_C and $l_i \Rightarrow r_i$ is a rule in R_N for $i = 1, 2$. Since C'_1, C'_2 and $r'_1 \approx l'_1[r'_2]$ are false in I_N, the conclusion is false in I_N. Hence the inference reduces C. $\square$

Lemma 4.14 (Theory Simplification) *Let N be a set of ground clauses such that N does not contain the empty clause. Suppose that the minimal counterexample C for I_N is a clause in N and that the selected literal $[\neg](p \approx q)$ of C is not in T-normal form. Then Sup_T contains a Theory Simplification inference that reduces C.*

Proof: Let $C = L \vee C'$ where L is selected in C and L is not in T-normal form. Without loss of generality we may assume $p \succeq q$. We may further assume that C' is false in I_C, since otherwise C would already be true in I_C. Since L is not in T-normal form and Simp_T is admissible, there exists some simplified literal L' in $\mathrm{Simp}_T(L)$. Then the following Theory Simplification inference is in Sup_T:

$$\frac{L \vee C'}{L' \vee C'}$$

Since L' is a simplification of L we know that $L' \vee C'$ is a simplification of $L \vee C'$ by Lemma 4.8. If $L' \vee C'$ were true in I_N this would imply that $L \vee C'$ is true in I_N, a contradiction. So $L' \vee C'$ is false in I_N. □

Lemma 4.15 (Superposition) *Let N be a set of ground clauses such that N does not contain the empty clause. Suppose that the minimal counterexample C for I_N is a clause in N and that the selected literal $[\neg](p \approx q)$ of C is in T-normal form, that $p \succ q$ and that p is reducible by S_N. Then Sup_T contains a Superposition inference that reduces C.*

Proof: Let $C = L \vee C'$ where $L = [\neg](p \approx q)$ is selected in C, L is in T-normal form, $p \succ q$ and p is reducible by S_C. Also, we may assume that C' is false in I_N, since otherwise C would already be true in I_N. Since p is reducible there exists rules $l \Rightarrow r \in R_N$ and $l' \Rightarrow r'$ in $\mathcal{S}_T(l \Rightarrow r)$ such that $p = u[l'']$ and $l' =_E l''$. We may assume that $l' \Rightarrow r'$ is chosen such that r' is minimal among the rules in $\mathcal{S}_T(l \Rightarrow r)$ with left-hand side l'. This rule has been produced by some ground clause D in N_C. Consider the following Superposition inference:

$$\frac{l \approx r \vee D' \qquad [\neg](u[l'] \approx q) \vee C'}{[\neg](u[r'] \approx q) \vee C' \vee D'}$$

By assumption $L = [\neg](u[l'] \approx q)$ is false in I_N, and by using $l' \approx r'$, the congruence laws and transitivity we obtain that L' is false in I_N. The instances of transitivity used have as their middle term at most the maximal term of C and are thus smaller than C. Since C is the minimal counterexample they are true in I_N. As C' and D' are also false in I_N the conclusion is false in I_N. We conclude that the Superposition inference reduces C. □

Lemma 4.16 (Selected literals) *Let N be a set of ground clauses such that N does not contain the empty clause. Suppose that the minimal counterexample C for I_N is a clause in N and that a negative literal is selected in C. Then Sup_T contains an inference that reduces C.*

Proof: Let $C = L \vee C'$ where $L = \neg A$ is selected.

(1) Suppose L is not in T-normal form. Then by Lemma 4.14 there exists a Theory Simplification inference that reduces C.

(2) Otherwise L is in T-normal form.

(2.1) If $p =_E q$ then consider the following Reflexivity Resolution inference in Sup_T:

$$\frac{p \not\approx q \vee C'}{C'}$$

The conclusion C' is false in I_N.

(2.2) Otherwise assume without loss of generality $p \succ q$. Since $p \approx q$ is true in I_N, the left-hand side p is reducible by S_N, and by Lemma 4.15 Sup_T contains a Superposition inference that reduces C. □

Theorem 4.17 Sup_T *has the reduction property for counterexamples.*

Proof: Let C be the minimal counterexample for I_N. If C is an instance of transitivity then by Lemma 4.13 Sup_T contains an Extension Superposition inference that reduces C.

Otherwise C is a ground clause in N which is not productive. We do a case analysis on the condition for productivity that is violated.

(ii) Suppose that C is not reductive for any positive literal $s \approx t$ in C. Then either (1) the strictly maximal positive literal is of the form $s \approx t$ and $s =_E t$, (2) there exist more than one maximal positive literal, or (3) the maximal literals are negative.

Case (1) cannot occur, since C is false in I_N.

In case (2) C has the form $s \approx t \vee s' \approx t' \vee C''$ where $s =_{\mathrm{AC}} s'$ and $t =_{\mathrm{AC}} t'$. The Equality Factoring inference

$$\frac{s \approx t \vee s' \approx t' \vee C''}{t \not\approx t' \vee s' \approx t' \vee C''}$$

reduces C.

In case (3) some negative literal in C must be selected, and we may apply Lemma 4.16 to infer that some inference in Sup_T reduces C.

(iii) Suppose $C = s \approx t \vee C'$ and C is reductive for $s \Rightarrow t$, but $s \Rightarrow t$ is not in T-normal form. Then by Lemma 4.14 some inference in Sup_T reduces C.

(iv) Suppose $C = s \approx t \vee C'$ and C is reductive for $s \Rightarrow t$, $s \approx t$ is in T-normal form, but s is reducible by S_C. Then by Lemma 4.15 some Superposition inference in Sup_T reduces C.

(v) Suppose $C = s \approx t \vee C'$ and C is reductive for $s \Rightarrow t$, $s \approx t$ is in T-normal form, s is irreducible by S_C, but C' is true in $(T \cup S_C \cup \mathcal{S}_T(s \approx t))^{\Downarrow}$. The only way that this can happen is that there is another positive equation with maximal term $s' =_E s$ in C', that is, $C' = s' \approx t' \vee C''$, such that $t \Downarrow_{T \cup S_N} t'$. Then the ground instance

$$\frac{s \approx t \vee s' \approx t' \vee C''}{t \not\approx t' \vee s' \approx t' \vee C''}$$

of Equality Factoring reduces C. □

We denote by $\mathcal{RC}(N)$ the set of ground clauses that are redundant in N, and by $\mathcal{RI}(N)$ the set of ground inferences that are redundant in N. By definition $\mathcal{RC}$ and $\mathcal{RI}$ are monotonic, i.e., $M \subseteq N$ implies $\mathcal{RC}(M) \subseteq \mathcal{RC}(N)$ and $\mathcal{RI}(M) \subseteq \mathcal{RI}(N)$. The following lemma allows to delete redundant clauses without loosing redundancy of clauses or inferences.

Lemma 4.18 *Let N be a set of ground clauses and $M = N \setminus \mathcal{RC}(N)$, and let C and D be ground clauses. Then $N_C \cup \mathrm{Trans}_C \models_I D$ implies $M_C \cup \mathrm{Trans}_C \models_I D$.*

Proof: Each clause C' in $N_C \setminus M_C$ is redundant in M, so $M_{C'} \cup \mathrm{Trans}_{C'} \models_I C'$. Since $\bigcup_{C' \prec C} M_{C'} \subseteq M_C$ and $\bigcup_{C' \prec C} \mathrm{Trans}_{C'} \subseteq \mathrm{Trans}_C$ we get that $M_C \cup \mathrm{Trans}_C \models_I N_C$. We conclude that $M_C \cup \mathrm{Trans}_C \models_I D$. □

Lemma 4.19

$$\mathcal{RC}(N) = \mathcal{RC}(N \setminus \mathcal{RC}(N)) \qquad \textit{and} \qquad \mathcal{RI}(N) = \mathcal{RI}(N \setminus \mathcal{RC}(N)).$$

A set N of ground clauses is *saturated up to redundancy* with respect to an inference system Sup if each inference in Sup with premises from $N \setminus \mathcal{RC}(N)$ is redundant in N. We say that an inference system Sup is *refutationally complete* for some theory T if whenever $N \cup T$ is inconsistent for some set N that is saturated up to redundancy with respect to Sup, then N contains the empty clause.

Theorem 4.20 *Suppose that T is a ground term rewriting system that is confluent modulo E, $\succ$ is an E-compatible and E-antisymmetric simplification quasi-ordering that is total on ground terms, $T \subseteq (\succ)$, Simp_T is a simplification function, and $\mathcal{S}_T$ is a symmetrization function for T modulo E with respect to $\succeq$ such that Simp_T is admissible with respect to $\mathcal{S}_T$. Then Sup_T is refutationally complete for T_1.*

Proof: Let N be some set of ground clauses that is saturated up to redundancy with respect to Sup_T, and assume that $N \cup T_1$ is inconsistent but does not contain the empty clause. Let M be the subset of clauses in N which are not redundant in N. We first show that I_M is a model of $M \cup T_1$. Suppose this is not the case and let C be the minimal counterexample for I_M, which cannot be the empty clause. Since Sup_T has the reduction property for counterexamples, Sup_T contains an inference with main premise C, conclusion D and side premises $C_1, \ldots, C_k$ such that each side premise has the form $C_i = l_i \approx r_i \vee C'_i$ and is reductive for $l_i \Rightarrow r_i$, and

$$I_M \models_I \neg D \wedge l_1 \Rightarrow r_1 \wedge \ldots \wedge l_k \Rightarrow r_k \wedge \neg C'_1 \wedge \ldots \wedge \neg C'_k.$$

Since N is saturated this inference is redundant in N. That is,

$$N_C \cup \mathrm{Trans}_C \cup \{l_i \Rightarrow r_i \mid i = 1, \ldots, n\} \cup \{\neg C'_i \mid i = 1, \ldots, n\} \models_I D.$$

Since C is the minimal counterexample M_C and Trans_C hold in I_M. All clauses in N_C are smaller than C and either in M_C or redundant in M_C by Lemma 4.19, hence $M_C \cup \mathrm{Trans}_C \models_I N_C$. Thus N_C holds in I_M, which in turn implies that D holds in I_M, a contradiction. So there exists no minimal counterexample for I_M and I_M is a model of $M \cup T_1$, in contradiction to its inconsistency. $\square$

We now prove the upper bound on the relation $\models_I$ promised on page 40:

Proposition 4.21 *Let M be a set of ground clauses, let N a set of ground clauses or rewrite rules, and let N' be the sets resulting from N by replacing any rewrite rule $l \Rightarrow r$ by a unit clause $l \approx r$. If $M \models_I N$ then $M \cup T_1 \models N'$.*

Proof: Suppose that $M \cup T_1 \not\models N'$. Then there exists some Herbrand interpretation I such that $M \cup T_1$ is true in I and N' is false in I. Let N_I be the set of ground clauses

$$\{C \mid C \text{ ground clause and } I \models C\}.$$

Since all inferences are sound with respect to T_1, the set N_I contains the conclusion of an inference whenever it contains all its premises. But this implies that all inferences are redundant in N_I, hence N_I is saturated and I_{N_I} is a model of N_I. Since N_I contains for each ground atom A exactly one of the single-literal clauses A or $\neg A$, a ground atom A is true in I if and only if it is true in I_{N_I}. Hence M is true and N' is false in I_{N_I}. This implies that N is false in I_{N_I} as well, and thus $M \not\models_I N$. $\square$

It remains to consider how to obtain a saturated set of clauses from an initial set that is not saturated. A *theorem proving derivation* (with respect to T) is a sequence of sets of clauses $N_0 \vdash N_1 \vdash \ldots$ such that for all $i \geq 0$ either $N_{i+1} = N_i \cup \{C\}$ for some clause C such that $N \cup T_1 \models C$, or $N_{i+1} = N_i \setminus \{C\}$ for some clause C which is redundant in N_i. By this definition whenever I is a model of T_1 and $N_i \vdash N_{i+1}$ then I is a model of N_i if and only if I is a model of N_{i+1}. For such a derivation the set of *persistent clauses* N_∞ is defined as $N_\infty = \bigcup_{i\geq 0}\bigcap_{j\geq i} N_j$. A theorem proving derivation is called *fair* with respect to a set of inferences Sup if all inferences in Sup from clauses in N_∞ are redundant in N_i for some $i \geq 0$.

Lemma 4.22 *Let $N_0 \vdash N_1 \vdash \ldots$ be a theorem proving derivation. Then*

$$\mathcal{RC}(\bigcup_i N_i) \subseteq \mathcal{RC}(N_\infty) \quad \textit{and} \tag{4.10}$$

$$\mathcal{RI}(\bigcup_i N_i) \subseteq \mathcal{RI}(N_\infty). \tag{4.11}$$

Proof: Let $N = \bigcup_i N_i$. Then any clause in $N \setminus N_\infty$ is redundant in some N_i and hence in N, which implies $N \setminus N_\infty \subseteq \mathcal{RC}(N)$. From $N \setminus (N \setminus N_\infty) = N_\infty$ we get

$$N \setminus \mathcal{RC}(N) \subseteq N_\infty \subseteq N.$$

By monotonicity of $\mathcal{RC}$ this implies

$$\mathcal{RC}(N \setminus \mathcal{RC}(N)) \subseteq \mathcal{RC}(N_\infty) \subseteq \mathcal{RC}(N).$$

By Lemma 4.19 $\mathcal{RC}(N) = \mathcal{RC}(N \setminus \mathcal{RC}(N))$, hence $\mathcal{RC}(N_\infty) = \mathcal{RC}(N)$.

The same applies to $\mathcal{RI}$. □

Theorem 4.23 *Let $N_0 \vdash N_1 \vdash \ldots$ be a theorem proving derivation that is fair with respect to* Sup. *Then N_∞ is saturated up to redundancy with respect to* Sup.

Proof: Redundancy of an inference in N_i implies redundancy in $\bigcup_i N_i$, and by the previous lemma it is redundant in N_∞. □

Since N_∞ is saturated, it has a model I_{N_∞} if it does not contain the empty clause. Since all clauses in $N_0 \setminus N_\infty$ are redundant in some N_i, they are also redundant in N_∞. Hence they are T_1-consequences of N_∞, and I_{N_∞} is a model of N_0. On the other hand, if N_0 is T_1-consistent then it has a model I that also satisfies T_1. This model I satisfies each N_i, hence also N_∞, and N_∞ is T_1-consistent and does not contain the empty clause. We conclude that $\bot \in N_\infty$ if and only if N_0 is T_1-inconsistent.

It remains to consider how to obtain a fair derivation. Since all inferences are strictly decreasing, they become redundant once their conclusion is added to N. Thus a fair derivation can be obtained by adding conclusions of inferences in a fair way. This can be improved by not adding conclusions of inferences which satisfy some sufficient criterion for redundancy.

Let us now look again at the commutative monoid example, and let us denote the set of clauses on page 39 by N_0. The minimal counterexample for I_{N_0} is C_2. Since its left-hand side is reducible by C_1 it can be reduced by the Superposition inference

$$\frac{f(0) \approx 0 \qquad f(0) + a \approx 0}{0 + a \approx 0}.$$

Adding the conclusion $C_8 = 0 + a \approx 0$ to N_0 we obtain $N_1 = N_0 \cup \{C_8\}$. Then $I_{N_1} = I_{N_0}$ as C_8 is not productive, because it is not in T-normal form. Using the Theory Simplification inference

$$\frac{0 + a \approx 0}{a \approx 0}$$

we obtain $C_9 = a \approx 0$, which becomes productive in I_{N_2} where $N_2 = N_1 \cup \{C_9\}$. With the two rules $f(0) \Rightarrow 0$ and $a \Rightarrow 0$ in R_{N_2} all ground term over F can be reduced to 0. Since all clauses in N_2 contain a positive literal, I_{N_2} is a model of N_2. Finally, we point out that the Extension Superposition inference

$$\frac{f(a) + a \approx f(a) \qquad f(a) + f(a) \approx f(0)}{f(a) + f(a) \approx f(0) + a}$$

takes care of the transitivity counterexample induced by C_6 and C_7. In N_2 this inference is redundant, since the conclusion follows from the smaller clauses $f(0) \approx 0$ and $a \approx 0$.

4.7 Extension peaks revisited

In the Extension Superposition rules stated above any extension peak between two rules leads to an inference, leading to a large or infinite number of inferences for any pair of clauses whose symmetrizations overlap. For specific theories we can do much better by exploiting the known structure of the symmetrizations. For the theories considered in this work either none or a single Extension Superposition inference suffices; we show that all other such inferences are redundant. We call the extension peaks that give rise to these Extension Superposition inferences *critical extension peaks.* Furthermore, critical extension peaks are the only cause of transitivity counterexamples. We exploit this to relax the bound on transitivity somewhat, by bounding only subterms that occur as the middle terms in critical extension peaks. We will need this extension in the cases of commutative rings and algebras, where the bound will be on single summands instead of on the whole sum. Finally, we will generalize a global redundancy criterion that is known for Gröbner basis computation and Knuth-Bendix completion (Buchberger 1979, Kapur, Musser and Narendran 1988).

Let $S_i = \mathcal{S}_T(l_i \Rightarrow r_i)$ for $i = 1, 2$. Consider some extension peak $t_1 \Leftarrow_{S_1} s \Rightarrow_{S_2} t_2$ between rules $l_1 \Rightarrow r_1$ and $l_2 \Rightarrow r_2$. We call such an extension peak *redundant* in N if all Extension Superposition inferences that have the peak as their main premise are redundant in N. We call the extension peak *redundant* if it is redundant in $\emptyset$.

Lemma 4.24 *Let $t_1 \Leftarrow_{S_1} s \Rightarrow_{S_2} t_2$ be an extension peak between rules $l_1 \Rightarrow r_1$ and $l_2 \Rightarrow r_2$. The peak is redundant in N if and only if the extension superposition inference*

$$\frac{l_1 \approx r_1 \qquad l_2 \approx r_2}{t_1 \approx t_2}$$

with main premise $t_1 \not\approx s \vee s \not\approx t_2 \vee t_1 \approx t_2$ is redundant in N.

Proof: If the peak is redundant in N then by definition the inference is redundant in N.

For the if-direction consider some extension superposition inference

$$\frac{C_1 \vee l_1 \approx r_1 \qquad C_2 \vee l_2 \approx r_2}{C_1 \vee C_2 \vee t_1 \approx t_2}$$

with main premise $C = t_1 \not\approx t \vee t \not\approx t_2 \vee t_1 \approx t_2$. The inference is redundant in N if

$$N_C \cup \mathrm{Trans}_C \cup \{l_1 \Rightarrow r_1, l_2 \Rightarrow r_2\} \cup \{\neg C_1, \neg C_2\} \models_I C_1 \vee C_2 \vee t_1 \approx t_2.$$

The inference

$$\frac{l_1 \approx r_1 \qquad l_2 \approx r_2}{t_1 \approx t_2}$$

with the same main premise C is redundant in N if and only if

$$N_C \cup \mathrm{Trans}_C \cup \{l_1 \Rightarrow r_1, l_2 \Rightarrow r_2\} \models_I t_1 \approx t_2,$$

which clearly implies the former. □

Lemma 4.25 *Let $t_1 \Leftarrow s \Rightarrow t_2$ be an extension peak between rules $l_1 \Rightarrow r_1$ and $l_2 \Rightarrow r_2$. If*

$$C = t_1 \not\approx t \vee t \not\approx t_2 \vee t_1 \approx t_2$$

is redundant in N then $t_1 \Leftarrow s \Rightarrow t_2$ is redundant in N.

Proof: By the definition of redundancy for clauses $N_C \cup \mathrm{Trans}_C \models_I C$, which by the deduction theorem is equivalent to $N_C \cup \mathrm{Trans}_C \cup \{t_1 \approx t, t \approx t_2\} \models t_1 \approx t_2$. From $l_i \Rightarrow r_i \models_I t \approx t_i$ for $i = 1, 2$, we conclude $N_C \cup \mathrm{Trans}_C \cup \{l_1 \Rightarrow r_1, l_2 \Rightarrow r_2\} \models t_1 \approx t_2$. □

An extension peak is called *critical* if it is not redundant. A ground instance of transitivity is called *critical* if it corresponds to a critical extension peak. A ground term is *critical* if it occurs as the middle term of a critical instance of transitivity.

The *critical closure* $\mathrm{cc}_T(t)$ of a ground term t is the greatest downward-closed set of ground terms that contains no critical term greater than t. For a ground clause C we let the *critical closure* $\mathrm{Trans}_{\mathrm{cc}_T(C)}$ be the greatest downward-closed set of ground instances of transitivity that contains no critical ground instance of transitivity greater than C. So $\mathrm{cc}_T(t)$ contains all terms below the smallest critical term greater than or equal to t with respect to the term, respectively. Analogously, $\mathrm{Trans}_{\mathrm{cc}_T(C)}$ contains all ground instances of transitivity below the smallest critical instance greater than or equal to C with respect to the clause ordering.

The notion of critical closure allows to increase the bound below which transitivity is known to hold. If we have an interpretation that satisfies Trans_C for ground clause C, then there are no critical extension peaks below $\mathrm{Trans}_{\mathrm{cc}_T(C)}$. The notion of the critical closure of a term helps to check this new bound:

Lemma 4.26 *Let C be a ground clause with maximal term s, and let D be an instance of transitivity that has a middle term in $\mathrm{cc}_T(s)$. Then D is in $\mathrm{Trans}_{\mathrm{cc}_T(C)}$.*

Proof: Suppose this is false and C is a ground clause with maximal term s, D is an instance of transitivity that has a middle term t in $\mathrm{cc}_T(t)$, but D is not in $\mathrm{Trans}_{\mathrm{cc}_T(C)}$. Then there is a critical instance D' of transitivity with $D \succeq D' \succ C$, and D' has a critical middle term t'. By the definition of the clause ordering and because the critical instance of transitivity D is the minimal clause containing t' we have $t \succeq t' \succ s$. But this contradicts t in $\mathrm{cc}_T(s)$. □

It remains to find sufficient criteria for a term t being in $\mathrm{cc}_T(s)$. Below we will give suitable ones for the particular theories.

Lemma 4.27 $\mathrm{Trans}_C \models_I \mathrm{Trans}_{\mathrm{cc}_T(C)}$.

Proof: Let $D = t_1 \not\approx t \vee t \not\approx t_2 \vee t_1 \approx t_2$ be the smallest instance of transitivity such that $\mathrm{Trans}_C \not\models_I D$. Then there exists a set of ground clauses N such that $I_N \models \mathrm{Trans}_C$ but $I_N \not\models D$, that is $t_1 \approx t$ and $t \approx t_2$ are true in I_N and $t_1 \approx t_2$ is false in I_N. There exist valley proofs $t_1 \Downarrow_{T \cup S_N} t$ and $t \Downarrow_{T \cup S_N} t_2$ but no valley proof $t_1 \Downarrow_{T \cup S_N} t_2$. Hence there exist rewrite steps in the two valley proofs that form a peak $t'_1 \Leftarrow t \Rightarrow t'_2$.

Suppose this peak converges. Then transitivity instances with a middle term strictly smaller than t imply $t_1 \approx t_2$, in contradiction to our assumption.

Suppose one of the rules used in the peak is from T. Since T is convergent and S_N symmetrized modulo E, this implies that the peak converges, which again contradicts the assumption.

Suppose one of the rules used is not extended. Since rules are produced only when their left-hand sides are irreducible, this cannot be the case. Thus the peak is an extension peak.

Suppose this extension peak were redundant. Since D is a minimal false instance of transitivity in I_N, the smaller instances in Trans_C would be true in I_N, and together with $t_1 \approx t$ and $t \approx t_2$ this would imply that $t_1 \approx t_2$ is true in I_N, a contradiction.

Suppose that both rules are applied strictly below the root, i.e.,

$$D = u[r'_1] \not\approx u[l'_1[l'_2]] \vee u[l'_1[l'_2]] \not\approx u[l'_1[r'_2]] \vee u[r'_1] \approx u[l'_1[r'_2]].$$

But then the smaller instance

$$D' = r'_1 \not\approx l'_1[l'_2] \vee l'_1[l'_2] \not\approx l'_1[r'_2] \vee r'_1 \approx l'_1[r'_2]$$

would imply D. Hence

$$D = r'_1 \not\approx l'_1[l'_2] \vee l'_1[l'_2] \not\approx l'_1[r'_2] \vee r'_1 \approx l'_1[r'_2]$$

where $t = l'_1[l'_2]$. This is a critical ground instance of transitivity. Since it is false in I_N, it cannot be in Trans_C, which holds in I_N. We conclude that D is not even in $\mathrm{Trans}_{\mathrm{cc}_T(C)}$. □

In our example theory of commutative monoids there is one critical extension peak for any two rules if the extensions overlap but the original rules don't. In such a minimal peak the flattened left-hand sides of the original unextended rules overlap maximally. E.g., consider the rules $f(a) + f(a) + a \Rightarrow f(a)$ and $f(a) + f(a) + f(a) \Rightarrow f(0)$. Their symmetrizations contain AC-extensions which overlap to form the peak

$$f(a) + f(a) + f(a) \Leftarrow f(a) + f(a) + f(a) + a \Rightarrow f(a) + f(0) + a.$$

This peak is indeed critical. However, other extension peaks of these two rules are redundant. For instance, the peak

$$f(a) + f(a) + f(a) + f(a) \Leftarrow f(a) + f(a) + f(a) + f(a) + a \Rightarrow f(a) + f(a) + f(0) + a$$

is redundant. Whenever the minimal peak converges, we can put a context around the valley proof to show that this peak converges, too. Note that here only one occurrence of $f(a)$ from the original rules overlaps. For the example theory a term can occur as the middle term of a critical extension peak if it is a sum of at least three summands and is irreducible with respect to T. The critical closure of such a term t contains all terms up to t, since any term $t + p$ is again critical. Consequently the move from Trans_C to $\mathrm{Trans}_{\mathrm{cc}_T(C)}$ is not useful for this theory. However, in later chapters we will handle theories where this extension is needed.

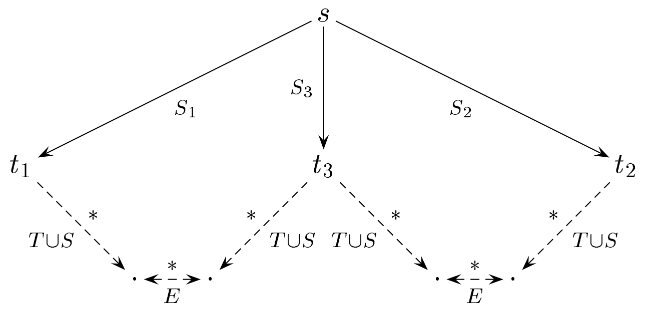


Figure 4.4: Reducible peaks are redundant

4.8 A redundancy criterion for extension peaks

We now show a redundancy criterion for extension peaks. If the middle term of the peak is reducible then the peak is redundant, provided that the two smaller peaks resulting from the overlaps with the reducing rule are redundant. This criterion was invented for Gröbner bases by Buchberger (1979) and adapted to term rewriting systems by Winkler and Buchberger (1983) and Kapur, Musser and Narendran (1988). The basic idea is illustrated by Figure 4.4. In this diagram the dashed proof shows that the peak $t_1 \Leftarrow s \Rightarrow t_2$ is redundant, provided that the peaks $t_1 \Leftarrow s \Rightarrow t_3$ and $t_3 \Leftarrow s \Rightarrow t_2$ converge.

Alternatively, we may describe this criterion in terms of transitivity. We are interested in showing that the instance

$$C = t_1 \not\approx s \ \vee\ s \not\approx t_2 \ \vee\ t_1 \approx t_2$$

holds in some interpretation I_N, under the assumption that all smaller instances of transitivity (and other clauses) hold. Then if t_3 is smaller than t_1 and t_2, in particular the transitivity instances

$$\begin{array}{ll} t_1 \not\approx s \ \vee\ s \not\approx t_3 \ \vee\ t_1 \approx t_3, & \\ t_3 \not\approx s \ \vee\ s \not\approx t_2 \ \vee\ t_3 \approx t_2, & \text{and} \\ t_1 \not\approx t_3 \ \vee\ t_3 \not\approx t_2 \ \vee\ t_1 \approx t_2 & \end{array}$$

hold in I_N, and hence also

$$s \not\approx t_3 \ \vee\ t_1 \not\approx s \ \vee\ s \not\approx t_2 \ \vee\ t_1 \approx t_2.$$

Now if $s \approx t_3$ holds in I_N, we can conclude that $I_N \models C$.

Theorem 4.28 *Let N be a set of ground clauses and let $C_i = l_i \approx r_i \vee C_i'$ be ground clauses that are reductive for $l_i \Rightarrow r_i$ for $i = 1, 2$. Furthermore, suppose there is an Extension Superposition inference between C_1 and C_2 with main premise*

$$C = t_1 \not\approx s \ \vee\ s \not\approx t_2 \ \vee\ t_1 \approx t_2$$

and $\{\neg C_1', \neg C_2'\} \cup N_C \cup \mathrm{Trans}_C \models_I s \approx t_3$ where $t_1 \succ t_3$ and $t_2 \succ t_3$.

Then the Extension Superposition inference is redundant in N.

Proof: We have to show

$$N_C \cup \mathrm{Trans}_C \cup \{l_1 \Rightarrow r_1,\ l_2 \Rightarrow r_2\} \cup \{\neg C_1',\ \neg C_2'\} \models_I C_1' \vee C_2' \vee t_1 \approx t_2$$

under the assumptions above. So let M be a set of ground clauses such that

$$I_M \models N_C \cup \mathrm{Trans}_C \cup \{l_1 \Rightarrow r_1,\ l_2 \Rightarrow r_2\} \cup \{\neg C_1',\ \neg C_2'\},$$

which implies that $s \approx t_1$ and $s \approx t_2$ are true in I_M. Then by the assumptions there exists a ground term t_3 such that $t_1 \succ t_3$, $t_2 \succ t_3$ and $s \approx t_3$ is true in I_M. By the argument above we conclude that $t_1 \approx t_2$ holds in I_M. □

To apply this in practice it remains to find criteria for the condition

$$\{\neg C_1', \neg C_2'\} \cup N_C \cup \mathrm{Trans}_C \models_I s \approx t_3$$

to hold. For instance, suppose that N_C contains a clause $C_3 = C_3' \vee l_3 \approx r_3$ that is reductive for $l_3 \Rightarrow r_3$ and $\mathcal{S}_T(l_3 \approx r_3)$ can be used to reduce s to a suitable t_3. Then

$$\{\neg C_1', \neg C_2'\} \cup N_C \cup \mathrm{Trans}_C \models_I \neg C_3'$$

implies that the condition is satisfied. Note that for empty C_3' this becomes trivial.

This criterion is useful for sets of clauses where each clause overlaps with many other clauses. As an extreme example, consider a set of n clauses such that there is a critical pair between any two clauses, resulting in $O(n^2)$ overlaps as a whole. Then there is some minimal rule that can reduce all peaks, and it suffices by this criterion to consider critical pairs with this rule, resulting in only $O(n)$ overlaps overall.

Buchberger (1979) reports a speedup of this magnitude for Gröbner bases, and Kapur, Musser and Narendran (1988) also report a substantial speedup for the case of Knuth-Bendix completion modulo AC. They report only a minor reduction in the number of critical pairs for the non-AC case, which does not lead to an overall speedup. It is not clear how useful our generalized criterion is for the case of nonunit clauses modulo AC.

4.9 Normalizing equational proofs

In this section we show that convergence with respect to normalized rewriting is equivalent to convergence with respect to unrestricted rewriting. We then use this result to show that strong symmetrization implies semicompatibility. That is, for every rewrite step between two terms there exists a valley proof between the normal forms with respect to the theory. This allows to normalize equational proofs by normalizing every term of the proof and replacing the rewrite steps by valley proofs. Then the terms in the normalized proof are bounded by the theory normal forms of the terms in the original proof. Later we will use this technique to show that equational proofs stay below a bound up to which transitivity holds.

Normalized rewriting was introduced by (Marché 1996). It gives rules from T priority over rules from S. For rewrite systems S and T we define *T-normalized rewriting with S* by $s \Rightarrow_{T!S} t$ if and only if $s \overset{*}{\Rightarrow}_T u$, u is irreducible with respect to T and $u \Rightarrow_S t$. We say that $T!S$ is *Church-Rosser modulo E* if $s \overset{*}{\Leftrightarrow}_{E \cup T \cup S} t$ implies that $s \Downarrow_{T!S} t$ for all terms s and t.

Lemma 4.29 *Let $T \cup S$ be terminating modulo E. Then $T \cup S$ is Church-Rosser modulo E if and only if $T!S$ is Church-Rosser modulo E.*

Proof: Since $(\Downarrow_{T!S}) \subseteq (\Downarrow_{T\cup S})$ the if-direction is trivial.

For the only-if-direction the proof is by induction on the following proof ordering: Let $\succ$ be $\overset{+}{\Rightarrow}_{(T\cup S)/E}$ extended by a new minimal element $\perp$. We order proof steps with respect to $\succ$ according to the complexity measure

$$\begin{aligned} c(s \Rightarrow_S t) = c(t \Leftarrow_S s) &= s \\ c(s \Rightarrow_T t) = c(t \Leftarrow_T s) &= \perp \\ c(s \Leftrightarrow_E t) &= \perp, \end{aligned}$$

i.e., only the larger sides of S-steps count. As the proof ordering we use the multiset extension of the ordering on proof steps.

Suppose that $T \cup S$ is Church-Rosser modulo E but $T!S$ is not, and consider the smallest proof $s \overset{*}{\Leftrightarrow}_{E\cup T\cup S} t$ such that $s \Downarrow_{T!S} t$ does not hold. Then $s \Downarrow_{T\cup S} t$ and there exists an S-step $u \Rightarrow_S v$ (or $v \Leftarrow_S u$) in $s \Downarrow_{T\cup S} t$ whose larger side u is reducible by T. For $u' \Leftarrow_T u \Rightarrow_S v$ there exists a proof $u' \Downarrow_{T\cup S} v$, and we may replace the subproof $u \Rightarrow_S v$ by the smaller subproof $u \Rightarrow_T u' \Downarrow_{T\cup S} v$. On the whole we obtain a smaller proof

$$s \Downarrow_{T\cup S} u' \Downarrow_{T\cup S} v \Downarrow_{T\cup S} t$$

and by induction hypothesis $s \Downarrow_{T!S} t$, a contradiction. □

This suggests that T-normalized rewriting with S can be interchanged freely with rewriting by $T \cup S$.

A relation $\Rightarrow_S$ is *semi-compatible* if $s \Rightarrow_S t$ implies $u[s] \Downarrow_S u[t]$ for all terms s and t and contexts u. Semi-compatibility was introduced by Bündgen (1991, 1996).

Lemma 4.30 *Let S be a set of ground rewrite rules which is strongly symmetrized with respect to T modulo E. Then T-normalized rewriting with S modulo E is semi-compatible.*

Proof: Consider a rewrite step $s \Rightarrow_{T!S} t$ and a context u. Then by putting the context $u[\,]$ around every term in the proof

$$s \Rightarrow_T \ldots \Rightarrow_T s' \Rightarrow_S t,$$

which represents the normalized rewrite step, one obtains a proof

$$u[s] \Rightarrow_T \ldots \Rightarrow_T u[s'] \Rightarrow_S u[t]$$

of $u[s] \approx u[t]$. Since this proof contains only one S-step all the rules used are in $T \cup S_i$ for some $i \in I$, where $(S_i)_{i\in I}$ is the partition of S in the definition of strong symmetrization. Since $T \cup S_i$ is convergent modulo E, by Lemma 4.29 there also exists a valley proof of $u[s] \approx u[t]$ by normalized rewriting. □

Without strong symmetrization it is not possible to remove all peaks between two rules from S, which are introduced when the symmetrization property is applied to remove several S/T-peaks consecutively. These S/S-peaks are not bounded by the normal form of a term in the original proof with added context. In contrast to this, strong symmetrization and in turn semi-compatibility of normalized rewriting lead to proofs which are bounded by the T-normal forms of the terms in the proof with added context (Figure 4.5).

$$\begin{tikzcd}
u[t_1] \arrow[d, "T"', "!"] & & & u[t_2] \arrow[d, "T"', "!"] & & & u[t_3] \arrow[d, "T"', "!"] \\
T(u[t_1]) \arrow[dr, "*", "T\cup S"'] & & & T(u[t_2]) \arrow[dl, "*"', "T\cup S"] \arrow[dr, "*", "T\cup S"'] & & & T(u[t_3]) \arrow[dl, "*"', "T\cup S"] \\
& \cdot \arrow[r, leftrightarrow, "*", "E"'] & \cdot & & \cdot \arrow[r, leftrightarrow, "*", "E"'] & \cdot &
\end{tikzcd}$$

Figure 4.5: With normalized rewriting

4.10 Merging paramodulation

Instead of equality factoring it is also possible to use Merging Paramodulation, by making the following modifications in the inference system and the completeness proof:

Replace Equality Factoring in Sup_T by the inference rules Factoring, Merging Paramodulation and Merging Theory Paramodulation, to obtain $\mathsf{Sup}_T^{\mathsf{MP}}$.

T-Factoring
$$\frac{s \approx t \ \vee \ s' \approx t' \ \vee \ C}{t \not\approx t' \ \vee \ s' \approx t' \ \vee \ C}$$

if (i) $s \approx t$ is in T-normal form and selected in $s \approx t \ \vee \ s' \approx t' \ \vee \ C$, (ii) $s =_E s'$ and $t =_E t'$.

T-Merging Theory Paramodulation
$$\frac{s \approx t[l''] \ \vee \ s' \approx t' \ \vee \ C}{s \approx t[r'] \ \vee \ s' \approx t' \ \vee \ C}$$

if (i) $l' \Rightarrow r' \in T$, (ii) $l' =_E l''$, (iii) $s \approx_E s'$, and (iv) $s \approx t[l']$ is T-normal form and selected in $s \approx t[l'] \ \vee \ C$.

T-Merging Paramodulation
$$\frac{l \approx r \ \vee \ D \qquad s \approx t[l''] \ \vee \ s' \approx t' \ \vee \ C}{s \approx t[r'] \ \vee \ s' \approx t' \ \vee \ C}$$

if (i) $l' \Rightarrow r' \in \mathcal{S}_T(l \approx r)$, (ii) $l' =_E l''$, (iii) $s \approx_E s'$, (iv) $s \approx t[l']$ is T-normal form and selected in $s \approx t[l'] \vee C$, and (v) $l \approx r$ is in T-normal form and selected in $l \approx r \vee D$.

Theorem 4.31 $\mathsf{Sup}_T^{\mathsf{MP}}$ *has the reduction property for counterexamples.*

Proof: The proof is similar to that of Theorem 4.17, with the following modified cases:

Case (ii.2): The maximal literal is positive and occurs more than once. That is, the minimal counterexample C has the form $s \approx t \vee s' \approx t' \vee C''$ where $s =_{\mathrm{AC}} s'$ and $t =_{\mathrm{AC}} t'$. Then the Factoring inference

$$\frac{s \approx t \ \vee \ s' \approx t' \ \vee \ C''}{s' \approx t' \ \vee \ C''}$$

reduces the minimal counterexample C.

Case (v): Suppose $C = s \approx t \ \vee \ C'$ and C is reductive for $s \Rightarrow t$, $s \approx t$ is in T-normal form, s is irreducible by S_C, but C' is true in $(T \cup S_C \cup \mathcal{S}_T(s \approx t))^{\Downarrow}$. The only way that this can happen is that there is another positive equation with maximal term $s' =_E s$ in C', that is, $C' = s' \approx t' \ \vee \ C''$, such that $t \Downarrow_{T \cup S_N} t'$ and $t \succ t'$. Then t is reducible by $T \cup S_C$.

(v.1) If t is reducible by T then $\mathsf{Sup}_T^{\mathsf{MP}}$ contains the Theory Merging Paramodulation inference

$$\frac{s \approx t[l'] \ \vee \ s' \approx t' \ \vee \ C''}{s \approx t[r] \ \vee \ s' \approx t' \ \vee \ C''}$$

where $l \Rightarrow r \in T$ and $l =_E l'$, which reduces C.

(v.2) Otherwise t is reducible by some rule $l' \Rightarrow r'$ in $\mathcal{S}_T(l \approx r)$ for some productive clause $D = l \approx r \ \vee \ D'$. Then the following Merging Paramodulation inference

$$\frac{l \approx r \ \vee \ D' \qquad s \approx t[l'] \ \vee \ s' \approx t' \ \vee \ C''}{s \approx t[r'] \ \vee \ s' \approx t' \ \vee \ C''}$$

reduces C. □

In this context it is not clear whether Merging Paramodulation or Equality Factoring is better. On the one hand Merging Paramodulation rewrites t and t' independently of each other, which may lead to duplicated effort. Equality Factoring would combine them in a negative literal, and the Simplification rule would use a more refined strategy to reach a T-normal form of this literal. On the other hand, it may be the case that there are no rewrite rules in T or clauses in N that can paramodulate into the greater of the right-hand sides. In this case Merging Paramodulation would do nothing, while Equality Factoring would produce a negative literal that would in turn lead to more inferences to do the simplification.

A theoretical advantage of Merging Paramodulation is that I_N is the perfect model of N when N is saturated with respect to $\mathsf{Sup}_T^{\mathsf{MP}}$ (Bachmair and Ganzinger 1991), which is not the case for Equality Factoring (Moser 1997). Whenever the model construction encounters a clause with more than one occurrence of the maximal term in positive and none in negative literals, it may choose which of the literals is made true. Since Merging Paramodulation rewrites these literals until the left-hand side of the largest one is irreducible, the model construction then chooses this literal, whose normal form with respect to $T \cup S_C$ is larger than the normal forms of the other literals. In contrast to this, Equality Factoring eliminates the maximal literal, which need not be the maximal one after reducing the left-hand sides with respect to $T \cup S_C$.

5

Abelian Groups

Let $\widehat{\mathrm{AG}}$ be the term rewriting system modulo AC for abelian groups of Peterson and Stickel (1981):

$$x + 0 \Rightarrow x \tag{AG.1}$$

$$x + (-x) \Rightarrow 0 \tag{AG.2}$$

$$x + y + (-y) \Rightarrow x \tag{AG.2e}$$

$$-0 \Rightarrow 0 \tag{AG.3}$$

$$-(-x) \Rightarrow x \tag{AG.4}$$

$$-(x + y) \Rightarrow (-x) + (-y) \tag{AG.5}$$

In this chapter we let $\mathrm{AC} = \mathrm{AC}(+)$ and $\mathrm{AG} = \mathrm{AC}\backslash gnd(\widehat{\mathrm{AG}})$. That is, AG is a ground term rewriting system that does rewriting with AC-matching with respect to $\widehat{\mathrm{AG}}$. Since AG contains all ground instances of equations in $\widehat{\mathrm{AG}}$, any Herbrand model that satisfies AG also satisfies $\widehat{\mathrm{AG}}$. Then $\mathrm{AG}_1 = \mathrm{AG} \cup \mathrm{AC} \cup \mathrm{Eq}$.

Proposition 5.1 AG *is convergent modulo* AC.

A term with $+$ as its root symbol is called a *proper sum*. We may consider any term t as a sum $t_1 + \cdots + t_n$ where $n \geq 1$ and t_i is not a proper sum for $i = 1, \ldots, n$. We call a $t_1, \ldots, t_n$ the *summands* of t. A term t is a *summand* of the literal $[\neg](p \approx q)$ if it is a summand of p or a summand of q. It is a *summand* of a clause C if it is a summand of some literal in C.

We will use the following notational conventions: r, s, t, u, v and w are used for arbitrary terms. For $n > 0$ we let nt denote a sum $t_1 + \cdots + t_n$ where $t_i =_{\mathrm{AC}} t$ for $i = 1, \ldots, n$. Also, $(-n)t$ denotes $n(-t)$ and $0t$ denotes 0. We will use $s - t$ as an abbreviation for $s + (-t)$. x, y and z denote variables. To simplify the presentation, our meta-level notation will be modulo ACU for $+$. That is, when we write $s = nt + s'$ then nt denotes n summands which are AC-equivalent to t and which occur somewhere in the sum, not necessarily in front. Moreover, neither nt nor s' need to be present. Thus s may also be of the form nt, s' or 0, depending on whether $s' = 0$, $n = 0$, or $s' = 0$ and $n = 0$, respectively.

5.1 Termination

For the term ordering we construct a theory path ordering. We let $F_I = F_{\mathrm{AC}} = \{+\}$ and use the precedence

$$\cdots \succ_p f_2 \succ_p f_1 \succ_p (-) \succ_p 0 \succ_p (+)$$

which is TPO-admissible for F_I. Then we let $\succeq_{\mathrm{AG}} = \succeq_{tpo}(\succeq_p, \succeq^{st}_{\mathrm{AC}(+)})$. This is the APO of section 3.4 for $f = (+)$.

Proposition 5.2 *$\succeq_{\mathrm{AG}}$ is an AC-compatible and AC-antisymmetric simplification quasi-ordering that is total on ground terms such that $\mathrm{AG} \subseteq (\succ_{\mathrm{AG}})$.*

For the remainder of this chapter we write $\succ$ for $\succ_{\mathrm{AG}}$.

The ordering on terms induces an ordering $\succ_{\mathbb{Z}}$ on integers, because $n_1\alpha \succ n_2\alpha$ is independent of the atomic term α, and depends only on n_1 and n_2. Formally, we let $n_1 \succ n_2$ if and only if $n_1 > n_2 \geq 0$, or $n_1 < 0$ and $n_1 < n_2$. Then we obtain the ordering

$$\cdots \succ_{\mathbb{Z}} -2 \succ_{\mathbb{Z}} -1 \succ_{\mathbb{Z}} \cdots \succ_{\mathbb{Z}} 2 \succ_{\mathbb{Z}} 1 \succ_{\mathbb{Z}} 0.$$

Proposition 5.3 *Let n_1 and n_2 be integers and α an AG-atomic ground term.*

1. *$n_1 \succ_{\mathbb{Z}} n_2$ if and only if $n_1\alpha \succ n_2\alpha$.*

2. *$n_1 \succ_{\mathbb{Z}} n_2$ implies $n_1\alpha \succ n_2\alpha + r$ for any ground term r such that $\alpha \succ r$.*

Proof: It suffices to note that $+$ has multiset status in the APO, and that $-\alpha$ is greater than $n\alpha$ for any $n \geq 0$. For the second case note that r is dominated by α in the multiset extension. □

5.2 Symmetrization

An equation $l \approx r$ is in *abelian group normal form* (or AG-*normal form*) if (i) $l = r = 0$, or (ii) $l = n\alpha$, $n \geq 1$, $\alpha \succ r$, and α is AG-atomic and irreducible with respect to AG. This implies in particular that there are no terms in both sides of the equation that can be cancelled. A literal is in AG-normal form if its equation is in AG-normal form. We denote the set of literals in AG-normal form by $\mathrm{Norm}_{\mathrm{AG}}$. For these normal forms we obtain the following symmetrization function:

$$\begin{aligned}
\mathcal{S}_{\mathrm{AG}}(0 \approx 0) &= \emptyset && \text{(AG.S0)}\\
\mathcal{S}_{\mathrm{AG}}(\alpha \approx r) &= \{\alpha \Rightarrow r\} && \text{(AG.S1)}\\
\mathcal{S}_{\mathrm{AG}}(n\alpha \approx r) &= \{n\alpha \Rightarrow r\} && \text{(AG.S2a)}\\
&\cup\, gnd(\{x + n\alpha \Rightarrow x + r\}) && \text{(AG.S2b)}\\
&\cup\, \{-\alpha \Rightarrow (n-1)\alpha - r\} \qquad \text{if } n \geq 2. && \text{(AG.S2c)}
\end{aligned}$$

Lemma 5.4 *$\mathcal{S}_{\mathrm{AG}}$ is a strong symmetrization function for AG modulo AC with respect to $\succeq_{\mathrm{AG}}$.*

Proof: By Proposition 5.3 and the multiset property we obtain that $\mathcal{S}_{\mathrm{AG}}(l \approx r)$ is included in $\succ_{\mathrm{AG}}$ for any equation $l \approx r$ in AG-normal form. Minimality of l among the left-hand sides of $\mathcal{S}_T(l \approx r)$ is obtained by inspection.

Next we consider soundness. That is, we show how the rules in the symmetrization can be derived from an AG-normal form and the rules in AG. The only interessting case is (AG.S2c), which can be obtained from the critical pair

$$(n-1)\alpha \overset{\mathrm{AG.2}}{\Leftarrow} (-\alpha) + n\alpha \overset{\mathrm{AG.S2a}}{\Rightarrow} (-\alpha) + r.$$

By adding $-r$ on both sides and applying the inverse law one obtains $-\alpha \approx (n-1)\alpha - r$.

It remains to consider convergence modulo AC of $\mathcal{S}_{\mathrm{AG}}(l \approx r)$. Cliffs with ground instances of AC converge, since $\mathcal{S}_{\mathrm{AG}}(n\alpha \approx r)$ contains all ground instances of AC-extensions, namely (AG.S2b). Peaks between rules in AG converge by convergence of AG. Since the left-hand side $n\alpha$ is irreducible with respect to AG, we need not consider overlaps of rules in AG into rules in $\mathcal{S}_{\mathrm{AG}}(l \approx r)$ below the root position. It remains to consider overlaps of rules in $\mathcal{S}_{\mathrm{AG}}(l \approx r)$ into rules in AG, and overlaps among rules in $\mathcal{S}_{\mathrm{AG}}(l \approx r)$. The only nontrivial case is (AG.S3) where the rule is of the form $n\alpha \approx r$ with $n \geq 2$. Since α is an atomic ground term that is irreducible with respect to AG, it can only overlap with itself. Hence it suffices to consider rules of the form $nc \approx r$ where c is a free constant and $n \geq 2$. We have checked the general case of critical pairs of rules in $\mathcal{S}_{\mathrm{AG}}(n\alpha \approx r)$ into rules in AG by hand and verified this against a machine-generated list of critical pairs for $n = 3$. E.g., for the critical pair above one rewrites the right-hand side to $(n-1)\alpha - r + r$ and normalizes with respect to AG, obtaining $(n-1)\alpha$, the left-hand side.

For critical pairs of rules in $\mathcal{S}_{\mathrm{AG}}(n\alpha \approx r)$ into rules in $\mathcal{S}_{\mathrm{AG}}(n\alpha \approx r)$ the only interesting rule is (AG.S2b). Suppose the two rules are $x + nc \Rightarrow x + r$ and $x' + nc \Rightarrow x + r$. Then the AC-unifier $\{x/x'\}$ of the left-hand sides subsumes all other AC-unifiers of the left hand sides and we obtain the trivial critical pair $x' + r \approx x' + r$. $\square$

5.3 Critical extension peaks and transitivity

Let us now consider extension peaks.

Theorem 5.5 *There is no extension peak with respect to* AG.

Proof: Extended rules are of the form AG.S2b or AG.S2c, so suppose $n_1 t_1 \Rightarrow r_1$ and $n_2 t_2 \Rightarrow r_2$ are rules whose symmetrizations $S_i = \mathcal{S}_{\mathrm{AG}}(l_i \Rightarrow r_i)$ form an extension peak. This peak can have two forms:

$$(n - n_1)\alpha + r_1 \Leftarrow_{S_1} n\alpha \Rightarrow_{S_2} (n - n_2)\alpha + r_2$$
$$(n_1 - 1)\alpha - r_1 \Leftarrow_{S_1} -\alpha \Rightarrow_{S_2} (n_2 - 1)\alpha - r_2$$

where $n \geq n_1, n_2$ and without loss of generality $n_1 \geq n_2$. But then $n_1\alpha \Rightarrow_{S_2} (n_1 - n_2)\alpha + r_2$, that is l_1 is reducible by S_2, and thus the peak is not an extension peak. $\square$

There are no critical terms either for this theory, hence $\mathrm{cc}_{\mathrm{AG}}(t)$ is the set of all terms and $\mathrm{Trans}_{\mathrm{cc}_{\mathrm{AG}}(C)} = \mathit{gnd}(\mathrm{Trans})$ for any ground term t and ground clause C.

Corollary 5.6 *Transitivity holds in I_N for all sets of ground clauses N.*

Proof: Let C be the empty clause. Then $\mathrm{Trans}_C = \emptyset$ and $\mathrm{Trans}_{\mathrm{cc}_{\mathrm{AG}}(C)} = \mathit{gnd}(\mathrm{Trans})$, and Lemma 4.27 becomes $\emptyset \models_I \mathit{gnd}(\mathrm{Trans})$. $\square$

5.4 Simplification

We first present AG-Isolation and show that it is a simplification rule. We show that it is a simplification even if the maximal term is not irreducible with respect to AG, as this will make it easier to lift Isolation to nonground clauses. Later we will use a restricted version in the simplification function for abelian groups.

AG-*Isolation*
$$\frac{[\neg](n_1 s_1 + r_1 \approx n_2 s_2 + r_2)}{[\neg]((n_1 - n_2)s_1 \approx r_2 - r_1)}\rightarrow$$

if (i) $s_1 =_{\mathrm{AC}} s_2$, (ii) s_1 is not a proper sum, (iii) $n_1 \geq n_2$, (iv) $n_2 \neq 0$ or $r_1 \neq 0$, and (v) $s_1 \succ r_1$ and $s_2 \succ r_2$.

Lemma 5.7 AG-*Isolation is a simplification rule.*

Proof: Since transitivity is universally valid in any model I_N in the abelian group case, it suffices to note that premise and conclusion are AG-equivalent, and that the conclusion is strictly smaller than the premise. To see the latter we show that at least one side of the premise is greater than the conclusion. Since $s_1 =_{\mathrm{AC}} s_2$ dominate the ordering, it suffices to consider n_1 and n_2. If n_1 or n_2 is negative then that side of the premise dominates the conclusion. If both n_1 and n_2 are strictly positive then n_1 is strictly greater than $(n_1 - n_2)s_1$. In the remaining case where $n_2 = 0$ and $r_1 \neq 0$ we have that $n_1 s_1 + r_1 \succ n_1 s_1$. □

Now we consider which simplification rules are needed to achieve AG-normal form. We further restrict AG-Rewriting and AG-Isolation in order to keep $\mathrm{Simp}_{\mathrm{AG}}$ small, since every simplification will become an inference when lifted. In $\mathrm{Simp}_{\mathrm{AG}}$-Rewriting it is not necessary to use rewriting with AG eagerly and to go all the way to a normal form on both sides of the equation. We need to reduce only summands which are maximal with respect to $\succ$, since these maximal summands will ultimately become the left-hand side. The nonmaximal summands will end up in r. Hence it is not necessary to reduce them, once they have been separated from the maximal terms. Consider an example of simplification where $s \succ t$:

$$s + s + t \approx -((-s) + (t + 0)) \tag{5.1}$$
$$s + s + t \approx (-(-s)) + (-(t + 0)) \tag{5.2}$$
$$s + s + t \approx s + (-(t + 0)) \tag{5.3}$$
$$s \approx (-(t + 0)) + (-t) \tag{5.4}$$

The maximal summand of (5.1) is $-((-s) + (t + 0))$, which includes both s and t. It is rewritten to the top-level sum $(-(-s)) + (-(t + 0))$ where $(-(-s))$ is the maximal summand of (5.2), and this in turn is rewritten to s. The last step from (5.3) to (5.4) isolates s on the left-hand side. Note that $t + 0$ is not reduced, since $s \succ -(t + 0)$. Of course, if the term $t + 0$ were present also on the nonground level, it would be a good idea to reduce it to t. But in general this is not the case, as there could be a variable instead of 0. We formalize this intuition in the simplification rules Sum Contraction and Theory Superposition. Sum Contraction handles the cancellation of several maximal terms, while Theory Superposition does rewriting inside a maximal term.

$\mathrm{Simp}_{\mathrm{AG}}$-*Sum Contraction*
$$\frac{[\neg](s - s' + p \approx q)}{[\neg](p \approx q)}\rightarrow$$

if (i) $s =_{\mathrm{AC}} s'$, (ii) $-s'$ is a maximal summand in $s - s' + p$, and (iii) $s - s' + p \succeq q$.

Simp_{AG}-*Summand Rewriting* $\quad \dfrac{[\neg](u[l'] + p \approx q)}{[\neg](u[r] + p \approx q)} \rightarrow$

if (i) $l \Rightarrow r$ is a rule in AG, (ii) $l' =_{\text{AC}} l$, (iii) $u[l']$ is a maximal summand in $u[l'] + p$, and (iv) $u[l'] + p \succeq q$.

Simp_{AG}-*Isolation* $\quad \dfrac{[\neg](n_1\alpha_1 + r_1 \approx n_2\alpha_2 + r_2)}{[\neg]((n_1 - n_2)\alpha_1 \approx r_2 - r_1)} \rightarrow$

if (i) $\alpha_1 =_{\text{AC}} \alpha_2$, (ii) α_1 is AG-atomic, (iii) α_1 is irreducible with respect to AG, (iv) $n_1 \geq n_2$, (v) $n_2 \neq 0$ or $r_1 \neq 0$, and (vi) $\alpha_1 \succ r_1$ and $\alpha_2 \succ r_2$.

We let $\text{Simp}_{\text{AG}}(L)$ consist of all literals L' such that there exists a simplification by AG-Sum Contraction, AG-Atom Rewriting or AG-Isolation with premise L and conclusion L'.

Lemma 5.8 Simp_{AG} *is a simplification function that is admissible with respect to* $\mathcal{S}_{\text{AG}}$.

Proof: By Lemmas 4.10 and 5.7 the rules are simplification rules.

It remains to show that any literal not in Norm_{AG} can be simplified by Simp_{AG}. Let $L = [\neg](p \approx q)$ and assume without loss of generality $p \succ q$. Suppose s is the maximal summand in the sum p. Then $p = n_1 s + p'$ and $q = n_2 s + q'$, $n_1 \geq 0$, $n_2 \geq 0$, $s \succ p'$ and $s \succ q'$ by the multiset property of $\succ$.

(1) Suppose $s = -s'$, and $p' =_{\text{AC}} s' + p''$. Then Sum Contraction applies. It cannot be the case that s' occurs in q' but not in p', since this contradicts $p \succeq q$.

(2) Suppose this is not the case, and s is reducible by AG. Then AG-Theory Superposition applies.

(3) Otherwise p is irreducible with respect to $\text{AG} \setminus \{(\text{AG}.1)\}$, and s and $-s$ do not occur together on one side of the literal. We consider applying AG-Isolation. s is of one of the two forms α or $-\alpha$, and we can write L as $[\neg](m_1\alpha + r_1 \approx m_2\alpha + r_2)$ such that m_1 and m_2 are integers which are not both zero, $m_1 \geq m_2$, $\alpha \succ r_1$ and $\alpha \succ r_2$. Isolation is applicable if $m_2 \neq 0$ or $r_1 \neq 0$. If this is not the case then L is of the form $[\neg](m_1\alpha \approx r_2)$ where $\alpha \succ r_2$. That is, L is in AG-normal form. □

5.5 The inference system

We obtain the ground inference system below. We do not mention selection as it is identical to the general case.

AG-*Sum Contraction* $\quad \dfrac{[\neg](s - s' + p \approx q) \ \vee\ C}{[\neg](p \approx q) \ \vee\ C}$

if (i) $s =_{\text{AC}} s'$, (ii) $-s'$ is a maximal summand in $s - s' + p$, and (iii) $s - s' + p \succeq q$.

AG-*Summand Rewriting* $\quad \dfrac{[\neg](u[l'] + p \approx q) \ \vee\ C}{[\neg](u[r] + p \approx q) \ \vee\ C}$

if (i) $l \Rightarrow r$ is a rule in AG, (ii) $l' =_{\text{AC}} l$, (iii) $u[l']$ is a maximal summand in $u[l'] + p$, and (iv) $u[l'] + p \succeq q$.

AG-*Isolation*
$$\frac{[\neg](n_1\alpha_1 + r_1 \approx n_2\alpha_2 + r_2) \lor C}{[\neg]((n_1 - n_2)\alpha_1 \approx r_2 - r_1) \lor C}$$

if (i) $\alpha_1 =_{\mathrm{AC}} \alpha_2$, (ii) α_1 is AG-atomic, (iii) α_1 is irreducible with respect to AG, (iv) $n_1 \geq n_2$, (v) $n_2 \neq 0$ or $r_1 \neq 0$, and (vi) $\alpha_1 \succ r_1$ and $\alpha_2 \succ r_2$.

AG-*Superposition A*
$$\frac{n\alpha \approx r \lor D \quad [\neg]((m\beta)[n\alpha'] \approx q) \lor C}{[\neg]((m\beta)[r] \approx q) \lor C \lor D}$$

if (i) $\alpha =_{\mathrm{AC}} \alpha'$, (ii) $m, n \geq 1$, (iii) α and β are AG-atomic, (iii) α and β are irreducible with respect to AG, (v) $\alpha \succ r$, and (vi) $\beta \succ q$.

AG-*Superposition B*
$$\frac{n\alpha \approx r \lor D \quad [\neg]((m\beta)[l'] \approx q) \lor C}{[\neg]((m\beta)[r + t] \approx q) \lor C \lor D}$$

if (i) $l' =_{\mathrm{AC}} n\alpha + t$, (ii) $m, n \geq 1$, (iii) α and β are AG-atomic, (iv) α and β are irreducible with respect to AG, (v) $\alpha \succ r$, and (vi) $\beta \succ q$.

AG-*Superposition C*
$$\frac{n\alpha \approx r \lor D \quad [\neg]((m\beta)[-\alpha'] \approx q) \lor C}{[\neg]((m\beta)[(n-1)\alpha - r] \approx q) \lor C \lor D}$$

if (i) $\alpha =_{\mathrm{AC}} \alpha'$, (ii) $m \geq 1$, $n \geq 2$, (iii) α and β are AG-atomic, (iv) α and β are irreducible with respect to AG, (v) $\alpha \succ r$, and (vi) $\beta \succ q$.

Superposition A combines (AG.S1) and (AG.S2a), Superposition B uses the AC-extension (AG.S2b), and Superposition C superposes with (AG.S2c).

AG-*Reflexivity Resolution*
$$\frac{0 \not\approx 0 \lor C}{C}$$

In practice one might want to replace the literal $0 \not\approx 0$ in AG-normal form by $p \not\approx q$ and $p =_{\mathrm{AC}} q$ in order to find proofs more quickly. Then the corresponding ordering restrictions on AG-Sum Contraction and AG-Summand Rewriting may be strengthened from $\succeq$ to $\succ$.

AG-*Equality Factoring*
$$\frac{n\alpha \approx r \lor n\alpha' \approx r' \lor C}{r \not\approx r' \lor n\alpha' \approx r' \lor C}$$

if (i) $\alpha =_{\mathrm{AC}} \alpha'$, (ii) $n \geq 1$, (iii) α is AG-atomic, (iv) α is irreducible with respect to AG, (v) $\alpha \succ r$, (vi) $\alpha \succ r'$, (vii) $r \succeq r'$.

We let $\mathsf{Sup}_{\mathrm{AG}}$ be the set of these inferences, where for each inference the same restrictions by selection as in the general case apply.

Theorem 5.9 $\mathsf{Sup}_{\mathrm{AG}}$ *is refutationally complete for* AG_1.

Proof: This follow from Theorem 4.20 in combination with Propositions 5.1 and 5.2, and Lemmas 5.4 and 5.8. Note that due to the absence of extension peaks there are no Extension Superposition inferences. □

6

Commutative Rings

Commutative rings (with a unit element) extend abelian groups by a commutative operation $\cdot$ with a unit element 1, such that $\cdot$ distributes over $+$. We will only consider rings with a unit element. Since $\cdot$ is associative and commutative we let $\mathrm{AC} = \mathrm{AC}(+) \cup \mathrm{AC}(\cdot)$ in this section. The following term rewriting system $\widehat{\mathrm{CR}}$ modulo AC is again by Peterson and Stickel (1981).

$$x + 0 \Rightarrow x \tag{CR.1}$$
$$x + (-x) \Rightarrow 0 \tag{CR.2}$$
$$x + y + (-y) \Rightarrow x \tag{CR.2e}$$
$$-0 \Rightarrow 0 \tag{CR.3}$$
$$-(-x) \Rightarrow x \tag{CR.4}$$
$$-(x + y) \Rightarrow (-x) + (-y) \tag{CR.5}$$
$$x \cdot 0 \Rightarrow 0 \tag{CR.6}$$
$$x \cdot 1 \Rightarrow x \tag{CR.7}$$
$$x \cdot (y + z) \Rightarrow (x \cdot y) + (x \cdot z) \tag{CR.8}$$
$$x \cdot (-y) \Rightarrow -(x \cdot y) \tag{CR.9}$$

Rules (CR.1)–(CR.5) are the rules for abelian groups. Again, $\widehat{\mathrm{CR}}$ contains the necessary AC-extensions, which in this case is only CR.2e. We let $\mathrm{CR} = \mathrm{AC}\backslash gnd(\widehat{\mathrm{CR}})$.

Proposition 6.1 CR *is convergent modulo* AC.

Proof: This follows from the convergence modulo AC of $\widehat{\mathrm{CR}}$ (Peterson and Stickel 1981). □

The set of interpreted function symbols F_{CR} is $\{+, \cdot, -, 0, 1\}$. In addition to the notational conventions for abelian groups we let ϕ and ψ denote terms of the form $\alpha_1 \cdots \alpha_n$ where α_i is atomic for $i = 1, \ldots, n$. We will call such a term a *product*. The product is *proper* if $n \geq 2$. For $n = 0$ we obtain the empty product, which we identify with the constant 1. More generally, our meta-level notation will now be modulo ACU for $\cdot$ as well. That is, when we write $\phi = \phi_1\phi_2$ then ϕ_1 or ϕ_2 may be missing, and ϕ can have one of the forms ϕ_1, ϕ_2 or 1, where $\phi_2 = 1$, $\phi_1 = 1$, or $\phi_1 = \phi_2 = 1$, respectively. We say that ϕ *divides* ψ and write $\phi \mid \psi$ if there exists some ϕ' such that $\phi\phi' =_{\mathrm{AC}} \psi$. Division is a quasi-ordering with equivalence kernel AC. By factoring one obtains a partial ordering, which is isomorphic to the submultiset ordering on finite multisets over atomic terms.

Hence it is easily seen to be a lattice, and we may write $\mathrm{lcm}(\phi_1, \phi_2)$ for the *least common multiple* and $\gcd(\phi_1, \phi_2)$ for the *greatest common divisor* of ϕ_1 and ϕ_2. lcm and gcd are only determined up to AC. Terms of the form $n\phi$ will be called *monomials*.

6.1 Termination

In this case we need a more complicated term ordering. We use the lexicographic combination of a quasi-ordering that is essentially the modified associative path ordering (MAPO) (Delor and Puel 1993), an ordering by polynomial interpretation (Peterson and Stickel 1981), and the AC-RPO (Rubio and Nieuwenhuis 1995).

We define the first ordering as a TPO for $F_I = \{+, \cdot, -, 0, 1\}$. We assume that a total precedence $\succeq_p$ on $F \setminus F_I$ is given, and let $\succeq_p$ also denote the TPO-admissible extension of $\succeq_p$ to F. For the TPO-status we use the method of Section 3.3. That is, we are given a set of constants F_C and a quasi-ordering $\succeq_c$ on F_C, and we need to define an ordering $\succeq_t$ on terms over $F_I \cup F_C$ that extends $\succeq_c$ and satisfies the conditions of Theorem 3.13.

That ordering is defined by assigning a certain complexity to any term in normal form with respect to the term rewriting system $\mathrm{D_{CR}} = \mathrm{AC} \backslash gnd(\widehat{\mathrm{D_{CR}}})$ modulo AC, where $\widehat{\mathrm{D_{CR}}}$ consists of the rules

$$-(x+y) \Rightarrow (-x) + (-y) \tag{CR.5}$$

$$x \cdot (y+z) \Rightarrow (x \cdot y) + (x \cdot z) \tag{CR.8}$$

$$x \cdot (-y) \Rightarrow -(x \cdot y) \tag{CR.9}$$

from CR. Hence $\mathrm{D_{CR}}$ is a subset of CR. This term rewriting system is convergent modulo AC. It distributes $\cdot$ and $-$ over $+$, and $\cdot$ over $-$. We denote the normal form of t with respect to $\mathrm{D_{CR}}$ by $\mathrm{D_{CR}}(t)$. A $\mathrm{D_{CR}}$-normal form of a term over $F_I \cup F_C$ is made up of possibly empty layers of function symbols, with $+$ symbols above $-$ symbols above $\cdot$ symbols and constants at the bottom. That is,

$$\mathrm{D_{CR}}(t) = -^{m_1}(\phi_1) + \cdots + -^{m_n}(\phi_n)$$

where $n \geq 1$, $m_i \geq 0$, $\phi_i = c_{i1} \cdots c_{ik_i}$, $k_i \geq 1$, and $c_{ij} \in F_C \cup \{0, 1\}$ for $1 \leq i \leq n$ and $1 \leq j \leq k_i$. To such a term we assign the complexity

$$\kappa(\mathrm{D_{CR}}(t)) = \{\langle \kappa_{\phi_1}, m_1 \rangle, \ldots, \langle \kappa_{\phi_n}, m_n \rangle\},$$

where $\kappa_\phi = \{c_1, \ldots, c_k\}$ for any product $\phi = c_1 \cdots c_k$. To order complexities we first extend $\succeq_c$ to $F_C \cup \{0, 1\}$ such that $c \succ_c 1 \succ_c 0$ for any constant c in F_C. The inner multisets κ_ϕ are ordered according to the multiset extension of $\succeq_c$, the pairs $\langle \kappa_\phi, m \rangle$ are ordered by the lexicographic product of the ordering on the inner multisets and $\geq$, and the complexities $\kappa(\mathrm{D_{CR}}(t))$ are in turn ordered by the multiset extension of the ordering on pairs. We denote the ordering on complexities by $\succeq_\kappa$. Then we define the ordering $\succeq_t$ on terms over $F_I \cup F_C$ by $s \succeq_t t$ if and only if $\kappa(\mathrm{D_{CR}}(s)) \succeq_\kappa \kappa(\mathrm{D_{CR}}(t))$ where s and t are terms over $F_I \cup F_C$. Finally we get the TPO-status $\succeq_1^{st}$ as the status derived from $\succeq_t$, and let

$$\succeq_1(\succeq_p) = \succeq_{tpo}(\succeq_p, \succeq_1^{st}).$$

Lemma 6.2 *Let $\succeq_p$ be a well-ordering on $F \setminus F_{\mathrm{CR}}$. Then $\succeq_1(\succeq_p)$ is a total* $\mathrm{AC} \cup \mathrm{D_{CR}}$*-compatible and* $\mathrm{AC} \cup \mathrm{D_{CR}}$*-antisymmetric simplification quasi-ordering on ground terms that contains* $\mathrm{CR} \setminus \mathrm{D_{CR}}$.

Proof: We first show that $\succeq_1^{st}$ is a TPO-status by Lemma 3.9. Clearly $\succeq_t(\succeq_c)$ is a quasi-ordering that extends $\succeq_c$.

(*Strictly compatible with contexts*) Let $f_{u[]}$ be the function that maps any complexity $\kappa(\mathrm{D_{CR}}(t))$ to the complexity $\kappa(\mathrm{D_{CR}}(u[t]))$. To see that $f_{u[]}$ is well-defined note that there is a one-to-one correspondence between AC-equivalence classes of terms in $\mathrm{D_{CR}}$-normal form and complexities. Hence $\mathrm{D_{CR}}(t)$ can be reconstructed from a given complexity. We have to show that $f_{u[]}$ is strictly monotonic for any context $u[]$. Since strict monotonicity is preserved by composition, it suffices to consider contexts of depth one.

(1) Consider $u = s + []$. Then $f_{u[]}$ maps $\kappa(\mathrm{D_{CR}}(t))$ to $\kappa(\mathrm{D_{CR}}(s)) \cup \kappa(\mathrm{D_{CR}}(t))$, which is strictly monotonic.

(2) Consider $u = -[]$. Then

$$f_{u[]}(\{\langle\kappa_{\phi_1}, m_1\rangle, \ldots, \langle\kappa_{\phi_k}, m_k\rangle\}) = \{\langle\kappa_{\phi_1}, m_1 + 1\rangle, \ldots, \langle\kappa_{\phi_k}, m_k + 1\rangle\}.$$

$f_{u[]}$ is strictly monotonic, since it is the multiset extension of the pairing of the identity and the strictly monotonic function $x \mapsto x + 1$.

(3) Consider $u = s \cdot []$. For $s = -^n(\psi)$ we get

$$f_{u[]}(\{\langle\kappa_{\phi_1}, m_1\rangle, \ldots, \langle\kappa_{\phi_k}, m_k\rangle\}) = \{\langle\kappa_{\phi_1} \cup \kappa_\psi, m_1 + n\rangle, \ldots, \langle\kappa_{\phi_k} \cup \kappa_\psi, m_k + n\rangle\}.$$

In this case $f_{u[]}$ is the multiset extension of $\langle\kappa_\phi, m\rangle \mapsto \langle\kappa_\phi \cup \kappa_\psi, m + n\rangle$, which is strictly monotonic, since $\kappa_\phi \mapsto \kappa_\phi \cup \kappa_\psi$ and $m \mapsto m + n$ are. Hence $f_{u[]}$ is strictly monotonic for $s = -^n(\psi)$. For $s = -^{n_1}(\psi_1) + \cdots + -^{n_l}(\psi_l)$ we see that $f_{u[]} = (f_{u_1[]} \cup \cdots \cup f_{u_l[]})$ where $u_i[] = -^{n_i}(\psi_i) \cdot []$ for $i = 1, \ldots, l$. As a finite union of strictly monotonic functions $f_{u[]}$ is strictly monotonic.

(*Subterm property*) It suffices to consider contexts of depth one, then the subterm property follows by structural induction. We have to show (1) $s + t \succeq_t t$, (2) $s \cdot t \succeq_t t$, and (3) $-t \succeq_t t$. For (1) we observe that $\kappa(\mathrm{D_{CR}}(t))$ is a proper submultiset of $\kappa(\mathrm{D_{CR}}(s + t))$. For (2) we observe that for every element of $\kappa(\mathrm{D_{CR}}(t))$ there exists at least one element of $\kappa(\mathrm{D_{CR}}(s \cdot t))$ with at least one additional factor. For (3) the m_i increase by one. Since nothing else changes, $u[t]$ is in each case greater than t.

(*Decreases infinite derivations*) Suppose there is some infinite descending chain

$$\kappa(\mathrm{D_{CR}}(t_1)) \succ_\kappa \kappa(\mathrm{D_{CR}}(t_2)) \succ_\kappa \cdots.$$

Then there exists an infinite descending chain of pairs

$$\langle\kappa_{\phi_1}, m_1\rangle \succ \langle\kappa_{\phi_2}, m_2\rangle \succ \ldots,$$

and since $>$ on natural numbers is well-founded an infinite descending chain

$$\kappa_{\phi_{i_1}} \succ_{mul} \kappa_{\phi_{i_2}} \succ_{mul} \cdots.$$

Then in turn there exists an infinite descending chain of constants

$$c_1 \succ_c c_2 \succ_c \ldots,$$

where c_i occurs in some t_{j_i} for $1 \le j_1 < j_2 < \ldots$. The constants $c_1, c_2, \ldots$ are from F_C, since there is no infinite descending chain starting in 0 or 1.

(*Constant dominance condition*) By inspection.

(AC $\cup$ D$_{\text{CR}}$-*compatible*) We have to show that $\succeq_t$ is AC $\cup$ D$_{\text{CR}}$-compatible. Suppose $s \Leftrightarrow_{\text{AC}\cup\text{D}_{\text{CR}}} s' \succeq_t t' \Leftrightarrow_{\text{AC}\cup\text{D}_{\text{CR}}} t$. Then

$$\kappa(\text{D}_{\text{CR}}(s)) = \kappa(\text{D}_{\text{CR}}(s')) \succeq_\kappa \kappa(\text{D}_{\text{CR}}(t)) = \kappa(\text{D}_{\text{CR}}(t'))$$

implies $s \succeq_t t$.

(AC $\cup$ D$_{\text{CR}}$-*antisymmetric*) We have to show that the quasi-ordering $\succeq_t(\succeq_c)$ is (AC $\cup$ D$_{\text{CR}} \cup \sim_c$)-antisymmetric. Suppose $s \sim_1 t$ and

$$\begin{aligned} \text{D}_{\text{CR}}(s) &= -^{m_1}(\phi_1) + \cdots + -^{m_k}(\phi_k) \text{ and} \\ \text{D}_{\text{CR}}(t) &= -^{n_1}(\psi_1) + \cdots + -^{n_l}(\psi_l). \end{aligned}$$

Then $\kappa(\text{D}_{\text{CR}}(s)) \sim_\kappa \kappa(\text{D}_{\text{CR}}(t))$ and there exists a permutation π such that $\langle \kappa_{\phi_i}, m_i \rangle \sim_{lex} \langle \kappa_{\psi_{\pi(i)}}, n_{\pi(i)} \rangle$, and hence $\kappa_{\phi_i} \sim_{mul(\succeq_c)} \kappa_{\psi_{\pi(i)}}$. Now suppose $\phi_i = c_{i1} \cdots c_{ik_i}$ and $\psi_i = d_{i1} \cdots d_{il_i}$. Then there again exists permutations π_i such that $c_{ij} \sim d_{\pi(i)\pi_{\pi(i)}(j)}$. Based on these permutations we show that $s \overset{*}{\Leftrightarrow}_{\text{AC}\cup\text{D}_{\text{CR}}\cup\sim_c} t$. First we normalize s and t with respect to D$_{\text{CR}}$. Next we pick a representative for each $\sim_c$-equivalence class and replace the constants c_{ij} and d_{ij} by their representatives c'_{ij} and d'_{ij}, respectively. Then $c'_{ij} = d'_{\pi(i)\pi_{\pi(i)}(j)}$ for $1 \leq i \leq k$ and $1 \leq j \leq k_i$, which implies that $\phi'_i =_{\text{AC}} \psi'_{\pi(i)}$ where $\phi'_i = c'_{i1} \cdots c'_{ik_i}$ and $\psi'_{\pi(i)} = d'_{\pi(i)1} \cdots d'_{\pi(i)l_{\pi(i)}}$. Since also $m_i = n_{\pi(i)}$ for $i = 1, \ldots, k$, we finally obtain $\text{D}_{\text{CR}}(s) =_{\text{AC}} \text{D}_{\text{CR}}(t)$. Altogether we have shown $s =_{\text{AC}\cup\text{D}_{\text{CR}}\cup\sim_c} t$.

(*Total*) Since $\succeq_c$ is assumed total, $\geq$ on natural numbers is total, and since multiset extension and lexicographic product preserve totality, $\succeq_t(\succeq_c)$ is total. Hence $\succeq_1$ is total by Lemma 3.10.

(*Orients rules from left to right*) (CR.2) is oriented from left to right by $\succeq_1(\succeq)$, because 0 is the minimal constant. For the other rules in CR$\setminus$D$_{\text{CR}}$ the right-hand side is a subterm of the left-hand side, hence by the subterm property of $\succeq_1(\succeq)$ they are oriented from left to right. $\square$

Let $\succeq^p_{\text{CR}}$ be the ordering by polynomial interpretation induced by the following interpretation:

$$\begin{aligned} p_0^{\text{CR}} &= 2 \\ p_1^{\text{CR}} &= 2 \\ p_+^{\text{CR}}(x,y) &= x + y + 5 \\ p_\cdot^{\text{CR}}(x,y) &= x \cdot y \\ p_-^{\text{CR}}(x) &= 2 \cdot x + 2 \\ p_f^{\text{CR}}(x_1, \ldots, x_n) &= x_1 + \cdots + x_n + 2 \qquad \text{for } f \text{ free.} \end{aligned}$$

This is the ordering used by Peterson and Stickel (1981) to prove termination of CR modulo AC, extended to terms containing free function symbols.

Lemma 6.3 *$\succeq^p_{\text{CR}}$ is a total* AC*-compatible simplification quasi-ordering on ground terms that orients the rules in* D$_{\text{CR}}$ *from left to right.*

Proof: $\succeq^p_{\text{CR}}$ is a total simplification quasi-ordering on ground terms by Proposition 2.10 and AC-compatible with respect to $+$ and $\cdot$ by Proposition 2.11. By computing the polynomials one sees that the rules in $\mathrm{D_{CR}}$ are oriented from left to right:

$$\begin{aligned}
p^{\mathrm{CR}}(-(x+y)) &= 2p^{\mathrm{CR}}(x+y)+2 = 2(x+y+5)+2 = 2x+2y+12\\
p^{\mathrm{CR}}((-x)+(-y)) &= p^{\mathrm{CR}}(-x)+p^{\mathrm{CR}}(-y)+5 = 2x+2+2y+2+5 = 2x+2y+9\\
p^{\mathrm{CR}}(x\cdot(y+z)) &= xp^{\mathrm{CR}}(y+z) = x(y+z+5) = xy+xz+5x\\
p^{\mathrm{CR}}(x\cdot y + x\cdot z) &= xy+xz+5\\
p^{\mathrm{CR}}(x\cdot(-y)) &= xp^{\mathrm{CR}}(-y) = x(2y+2) = 2xy+2x\\
p^{\mathrm{CR}}(-(x\cdot y)) &= 2xy+2
\end{aligned}$$

Note that a polynomial nx is greater than n, since variables are instantiated by numbers ≥ 2. □

We let $\succeq_{\mathrm{CR}}$ be the lexicographic combination of $\succeq_1$, $\succeq^p_{\mathrm{CR}}$ and $\succeq_{acrpo}$ over an arbitrary precedence.

Proposition 6.4 *Let $\succ_p$ be a total precedence on $F \setminus F_{\mathrm{CR}}$. Then $\succeq_{\mathrm{CR}}(\succeq_p)$ is a total* AC*-antisymmetric and* AC*-compatible simplification quasi-ordering on ground terms that contains* CR.

Proof: The lexicographic combination preserves total AC-compatible simplification quasi-orderings. $\succeq_{\mathrm{CR}}(\succeq_p)$ is AC-antisymmetric because $\succeq_{acrpo}$ is. The rules in CR are oriented from left to right, because $\succeq_1$ orients the rules in $\mathrm{CR} \setminus \mathrm{D_{CR}}$ from left to right and satisfies $l \sim_1 r$ for the rules in $\mathrm{D_{CR}}$, and because $\succeq^p_{\mathrm{CR}}$ orients the rules in $\mathrm{D_{CR}}$ from left to right. □

From now on we assume an arbitrary precedence $\succeq_p$ and write $\succeq_{\mathrm{CR}}$ or even $\succeq$ for $\succeq_{\mathrm{CR}}(\succeq_p)$.

6.2 Symmetrization

We say a ground equation is in CR-*normal form* if it is of one of the forms (i) $0 \approx 0$; or (ii) $n\phi \approx r$ where $n \geq 1$, $\phi = \alpha_1 \ldots \alpha_k$, $k \geq 0$, $\alpha_1, \ldots, \alpha_k$ are CR-atomic, and $\phi \succ r$. We let $\mathrm{Norm_{CR}}$ be the set of literals whose equation is in CR-normal form. We can now define the symmetrization function for commutative rings as follows:

$$\begin{aligned}
\mathcal{S}_{\mathrm{CR}}(0 \approx 0) &= \emptyset && \text{(CR.S1)}\\
\mathcal{S}_{\mathrm{CR}}(\alpha \approx r) &= \{\alpha \Rightarrow r\} && \text{(CR.S2)}\\
\mathcal{S}_{\mathrm{CR}}(\phi \approx r) &= \{\phi \Rightarrow r\} && \text{(CR.S3a)}\\
&\quad \cup\, gnd(\{y\cdot\phi \Rightarrow y\cdot r\}) \quad \text{if } \phi \text{ is a proper product;} && \text{(CR.S3b)}\\
\mathcal{S}_{\mathrm{CR}}(n\phi \approx r) &= \{n\phi \Rightarrow r\} && \text{(CR.S4a)}\\
&\quad \cup\, gnd(\{x + n\phi \Rightarrow x + r\}) && \text{(CR.S4b)}\\
&\quad \cup\, gnd(\{n(y\cdot\phi) \Rightarrow y\cdot r\}) && \text{(CR.S4c)}\\
&\quad \cup\, gnd(\{x + n(y\cdot\phi) \Rightarrow x + y\cdot r\}) && \text{(CR.S4d)}\\
&\quad \cup\, \{-\phi \Rightarrow (n-1)\phi - r\} && \text{(CR.S4e)}\\
&\quad \cup\, gnd(\{-(y\cdot\phi) \Rightarrow (n-1)(y\cdot\phi) - (y\cdot r)\}) \quad \text{if } n \geq 2. && \text{(CR.S4f)}
\end{aligned}$$

Note that ϕ need not be a proper product for (CR.S4a)–(CR.S4f); this case applies even to equations like $1+1 \approx 0$.

Lemma 6.5 *Let ϕ be a ground product in* CR*-normal form and r a ground term such that $\phi \succ_{\mathrm{CR}} r$. Then $-\phi \succ_{\mathrm{CR}} n\phi - r$ and $-(s \cdot \phi) \succ_{\mathrm{CR}} n(s \cdot \phi) - (s \cdot r)$.*

Proof: Since ϕ is in CR-normal form it is also in $\mathrm{D_{CR}}$-normal form, and hence a minimal element of its ACD-equivalence class. But then $\phi \succ_{\mathrm{CR}} r$ implies that ϕ and r are ACD-distinct, and since $\succeq_1$ is ACD-antisymmetric and total $\phi \succ_1 r$. We have $\mathrm{D_{CR}}(\phi) = \phi$ and $\kappa(\phi) = \{\langle \kappa_\phi, 0\rangle\}$. Suppose

$$\begin{aligned}\mathrm{D_{CR}}(s) &= -^{m_1}(\phi_1) + \cdots + -^{m_k}(\phi_k) \text{ and}\\ \mathrm{D_{CR}}(r) &= -^{n_1}(\psi_1) + \cdots + -^{n_l}(\psi_l).\end{aligned}$$

Then $\phi \succ_1 r$ implies that $\kappa_\phi \succ_{mul(\succeq_c)} \kappa_{\psi_j}$ for any $j = 1, \ldots, l$. We have

$$\begin{aligned}\kappa(\mathrm{D_{CR}}(s \cdot \phi)) &= \{\langle \kappa_{\phi_1} \cup \kappa_\phi, m_1\rangle, \ldots, \langle \kappa_{\phi_k} \cup \kappa_\phi, m_k\rangle\} \text{ and}\\ \kappa(\mathrm{D_{CR}}(-(s \cdot r))) &= \{\langle \kappa_{\phi_1} \cup \kappa_{\psi_1}, m_1 + n_1 + 1\rangle, \ldots, \langle \kappa_{\phi_k} \cup \kappa_{\psi_1}, m_k + n_1 + 1\rangle,\\ &\quad \ldots,\\ &\quad \langle \kappa_{\phi_1} \cup \kappa_{\psi_l}, m_1 + n_l + 1\rangle, \ldots, \langle \kappa_{\phi_k} \cup \kappa_{\psi_l}, m_k + n_l + 1\rangle\},\end{aligned}$$

and we see that any pair $\langle \kappa_{\phi_i} \cup \kappa_\phi, m_i\rangle$ in $\kappa(n(s \cdot \phi))$ and any pair $\langle \kappa_{\phi_i} \cup \kappa_{\psi_j}, m_i + n_j + 1\rangle$ in $\kappa(\mathrm{D_{CR}}(-(s \cdot r)))$ is strictly smaller than $\langle \kappa_{\phi_i} \cup \kappa_\phi, m_i + 1\rangle$ in $\kappa(\mathrm{D_{CR}}(-(s \cdot \phi)))$. Hence $-(s \cdot \phi) \succ_1 n(s \cdot \phi) - (s \cdot r)$ and $-(s \cdot \phi) \succ_{\mathrm{CR}} n(s \cdot \phi) - (s \cdot r)$. The case $-\phi \succ_1 n\phi - r$ can be obtained in the same way by stipulating $\kappa(\mathrm{D_{CR}}(s)) = \{\langle \emptyset, 0\rangle\}$. □

Lemma 6.6 *$\mathcal{S}_{\mathrm{CR}}$ is a strong symmetrization function for* CR.

Proof: To see that the rules in the symmetrization are oriented from left to right suppose $\phi \succ_{\mathrm{CR}} r$. Then $s \cdot \phi \succ_{\mathrm{CR}} s \cdot r$ for any ground term s by compatibility with contexts, and $n\phi \succ_{\mathrm{CR}} r$ and $n(s \cdot \phi) \succ_{\mathrm{CR}} s \cdot r$ by the subterm property. Applying monotonicity again, we get $x + n\phi \succ_{\mathrm{CR}} x + r$ and $x + n(s \cdot \phi) \succ_{\mathrm{CR}} x + s \cdot r$. This covers (CR.S2)–(CR.S4d). (CR.S4e) and (CR.S4f) are oriented from left to right by Lemma 6.5.

By inspection we see that $n\phi$ is minimal among the left-hand sides in $\mathcal{S}_T(n\phi \approx r)$.

Next, let us discuss how the rules in the symmetrization can be derived from an equation $l \approx r$ in CR-normal form. (CR.S4c) is obtained from the critical peak

$$n(t \cdot \phi) \overset{\mathrm{CR.8}}{\Leftarrow} t \cdot (n_1\phi + n_2\phi) \overset{\mathrm{CR.S4a}}{\Rightarrow} t \cdot r.$$

(CR.S4e) and (CR.S4f) can be obtained from a critical peak with (CR.2e) as in the case of abelian groups. The other rules are AC-extensions, which follow by compatibility with contexts of equality.

Finally we have to show convergence of $\mathrm{CR} \cup \mathcal{S}_{\mathrm{CR}}(l \approx r)$ modulo AC for any equation $l \approx r$ in CR-normal form. To see that the peaks and cliffs converge it suffices like for abelian groups to consider overlaps of rules in $\mathcal{S}_{\mathrm{CR}}(l \approx r)$ into rules in CR below the root position, and overlaps among rules in $\mathcal{S}_{\mathrm{CR}}(l \approx r)$. The only nontrivial cases are (CR.S3) and (CR.S4). We observe that no left-hand side of a rule in $\widehat{\mathrm{CR}}$ has $\cdot$ immediately above two free variables. Hence the product ϕ in a rule of the symmetrization cannot be split into parts when a critical pair is formed, and we may treat it as a constant. We have checked the general case of critical pairs of rules in $\mathcal{S}_{\mathrm{CR}}(n\phi \approx r)$ into rules in CR by hand and verified this against a machine-generated list of critical pairs for $n = 3$ and $\phi = c$. For instance, the critical pair between (CR.8) and (CR.S4a) above converges by applying (CR.S4c).

$$\begin{tikzcd}
& n_1\phi_1\phi_2 \arrow[dl, "S_1"] \arrow[dr, "S_2"'] & \\
r_1\phi_2 \arrow[dr, dashed, "S_2"'] & & n_1\phi_1 r_2 \arrow[dl, dashed, "S_1"] \\
& r_1 r_2 &
\end{tikzcd}$$

Figure 6.1: Convergence of an extended peak

We now come to overlaps of $\mathcal{S}_{\mathrm{CR}}(l \approx r)$ with itself. The AC-unifiers of $x \cdot \phi$ and $y \cdot \phi$ are $\{x/y\}$, $\{x/\alpha_{i_1} \cdots \alpha_{i_l}, y/\alpha_{i_1} \cdots \alpha_{i_l}\}$, and $\{x/z \cdot \alpha_{i_1} \cdots \alpha_{i_l}, y/z \cdot \alpha_{i_1} \cdots \alpha_{i_l}\}$, of which only the first is most general. Hence (CR.S3) leads only to trivial critical peaks. Similar arguments apply to addition and multiplication for (CR.S4). As a consequence the peaks are either trivial, or result from overlaps of AC-extensions, where rewriting takes place in parallel subterms. □

6.3 Critical extension peaks and transitivity

The symmetrizations of two rules $n_1\phi_1 \Rightarrow r_1$ and $n_2\phi_2 \Rightarrow r_2$ always form a peak. For if without loss of generality $n_1 \geq n_2$ then there is a peak

$$r_1\psi_1 \Leftarrow n_1\phi \Rightarrow (n_1 - n_2)\phi + r_2\psi_2$$

where $\phi =_{\mathrm{AC}} \mathrm{lcm}(\phi_1, \phi_2) =_{\mathrm{AC}} \phi_1\psi_1 =_{\mathrm{AC}} \phi_2\psi_2$. Note that this peak is minimal, there is no peak with a smaller middle term. This is an extension peak if and only if $n_1\phi$ properly contains both $n_1\phi_1$ and $n_2\phi_2$. That is, $\psi_1 \neq 1$ and $n_1 > n_2$ or $\psi_2 \neq 1$.

Theorem 6.7 *Let $n_i\phi_i \Rightarrow r_i$ be a rewrite rule in* $\mathrm{Norm}_{\mathrm{CR}}$ *for $i = 1, 2$, and assume without loss of generality $n_1 \geq n_2$.*

These two rules have the single critical extension peak

$$r_1\psi_1 \Leftarrow n_1\phi \Rightarrow (n_1 - n_2)\phi + r_2\psi_2$$

where $\phi =_{\mathrm{AC}} \mathrm{lcm}(\phi_1, \phi_2) =_{\mathrm{AC}} \phi_1\psi_1 =_{\mathrm{AC}} \phi_2\psi_2$ if (1) $\psi_1 \neq 1$, (2) $n_1 > n_2$ or $\psi_2 \neq 1$, and (3) either (a) $n_1, n_2 \geq 2$, (b) $n_1 \geq 2$, $n_2 = 1$ and ϕ_2 is a proper product with $\mathrm{gcd}(\phi_1, \phi_2) \neq 1$, *or (c) $n_1 = n_2 = 1$ and ϕ_1 and ϕ_2 are proper products with* $\mathrm{gcd}(\phi_1, \phi_2) \neq 1$.

Otherwise there is no critical extension peak between these two rules.

Proof: First we note that whenever $n_i = 1$ and ϕ_i is not a proper product for either $i = 1$ or $i = 2$ then there is no extension peak. Otherwise, if $n_2 = 1$ and $\mathrm{gcd}(\phi_1, \phi_2) = 1$ then the peak converges, as indicated in Figure 6.1. If this is not the case then either $n_1, n_2 \geq 2$, or $n_1 \geq 2$, $n_2 = 1$ and ϕ_2 is a proper product with $\mathrm{gcd}(\phi_1, \phi_2) \neq 1$, or $n_1 = n_2 = 1$ and ϕ_1 and ϕ_2 are proper products in with $\mathrm{gcd}(\phi_1, \phi_2) \neq 1$, which is the condition for the existence of a critical extension peak in the theorem. For this peak we now show that it is the only one.

By Lemma 4.24 it suffices to show that any Extension Superposition inference

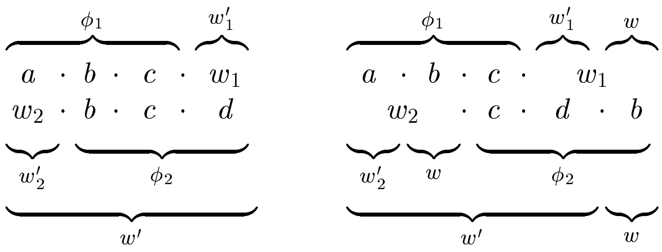
$$\frac{n_1\phi_1 \approx r_1 \qquad n_2\phi_2 \approx r_2}{t_1 \approx t_2}$$

(a) A trivial overlap: $\begin{array}{l} abcw_1 + abcw_1 + \qquad\qquad v_1 \\ \qquad v_2 \qquad\quad + bcdw_2 + bcdw_2 + bcdw_2 \end{array}$

(b) A nontrivial overlap: $\begin{array}{l} abcw_1 + abcw_1 + \qquad v_1 \\ \quad v_2 \quad + bcdw_2 + bcdw_2 + bcdw_2 \end{array}$

(c) A minimal overlap: $\begin{array}{l} abcw_1 + abcw_1 + \quad v_1 \\ bcdw_2 + bcdw_2 + bcdw_2 \end{array}$

Figure 6.2: Overlaps of sums



(a) A minimal overlap (b) A nonminimal overlap

Figure 6.3: Overlaps of products

with main premise

$$C = t_1 \not\approx t \vee t \not\approx t_2 \vee t_1 \approx t_2$$

is redundant, if the extension peak $t_1 \Leftarrow t \Rightarrow t_2$ between $n_1\phi_1 \Rightarrow r_1$ and $n_2\phi_2 \Rightarrow r_2$ is not of the form above. That is, we have to show

$$\mathrm{Trans}_C \cup \{n_1\phi_1 \Rightarrow r_1, n_2\phi_2 \Rightarrow r_2\} \models_I t_1 \approx t_2.$$

Consider some interpretation I_N that satisfies $\mathrm{Trans}_C \cup \{n_1\phi_1 \Rightarrow r_1, n_2\phi_2 \Rightarrow r_2\}$. We have to show that $t_1 \approx t_2$ is true in I_N. Then R_N contains these two rules, $S_N = \mathcal{S}_{\mathrm{CR}}(R_N)$, and $\mathrm{CR} \cup S_N$ is Church-Rosser on terms below t. Let $S_i = \mathcal{S}_{\mathrm{CR}}(l_i \approx r_i)$ for $i = 1, 2$. The rules from S_i can be grouped into the following two most general cases:

$$v_i + n_i(w_i\phi_i) \Rightarrow v_i + w_ir_i \tag{6.1}$$

$$-(w_i\phi_i) \Rightarrow (n_i - 1)(w_i\phi_i) - (w_ir_i) \tag{6.2}$$

Other forms may be seen as special cases, where v_i or w_i are omitted, where $n_i = 1$, or where ϕ_i is missing. The analysis of these cases is essentially the same. Without loss of generality we assume $n_1 \geq n_2$.

(1) We start by considering extended peaks of two ground rules $l_i \Rightarrow r_i$ of the form (6.1). Since AC-contexts can be moved into v_i or w_i, we need to consider only overlaps at the root in which the term at the top of the peak is equal to both left-hand sides.

(a) If the sums $n_iw_i\phi_i$ don't overlap in some product, the rewriting steps are independent of each other and the peak is trivially redundant (Figure 6.2a).

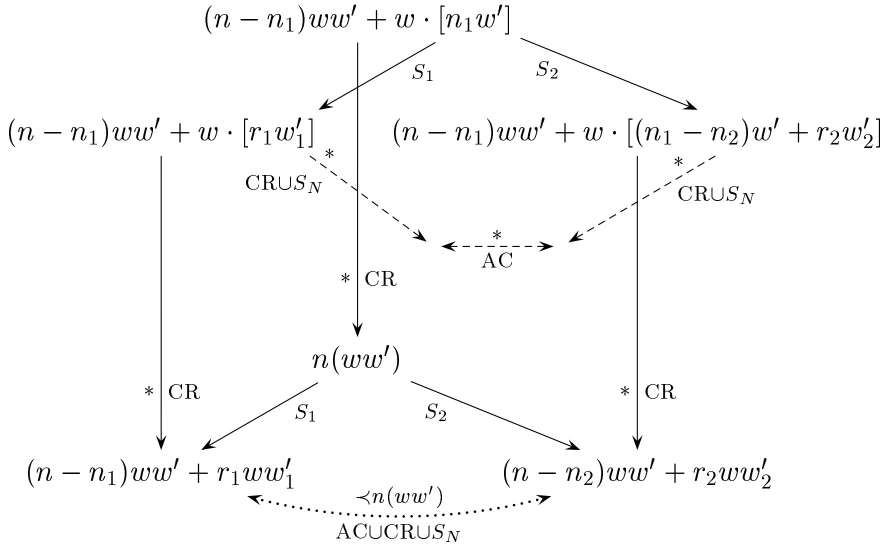


Figure 6.4: Redundancy of an extended peak

(b) If at least one product overlaps then $w_1\phi_1 =_{\mathrm{AC}} w_2\phi_2$ (Figure 6.2b). Let w be the greatest common divisor of w_1 and w_2. Then $w_i = ww_i'$ for $i = 1, 2$ (Figure 6.3b). Next, let $w' = w_1'\phi_1 =_{\mathrm{AC}} w_2'\phi_2$. We get the most general nontrivial overlap $v_1 + n_1 ww' =_{\mathrm{AC}} v_2 + n_2 ww'$ where $v_i = v + (n - n_i)ww'$ for some v and n. Then $v + (n - n_1)ww'$ is the part that v_1 and v_2 have in common.

If v is not empty, we can show redundancy by taking the smaller proof without v and putting the context $v + []$ around every term in the proof. Hence we may assume that v is empty.

Next we consider the case $w \neq 1$ or $n > n_1$. Figure 6.4 shows why such a peak is redundant. The nonempty context implies that the peak consisting of the boxed formulas is smaller than $n(ww')$, hence we may assume the existence of the dashed proof. From this we get the dotted proof by putting the context $(n - n_1)ww' + w \cdot []$ around its terms and normalizing it. The terms in the dashed proof have strictly less occurrences of w' than $n_1 w'$, and normalizing transforms them into the same number of occurrences of ww'. Thus, the terms in the dotted proof stay below the bound $n(ww')$. From the Church-Rosser property below $n(ww')$ we obtain that $(n-n_1)ww' + r_1 ww_1' \Downarrow_{\mathrm{CR} \cup S_N} (n-n_2)ww' + r_2 ww_2'$, i.e., the critical peak converges. Note that we get the critical extended peak if the context is empty.

(c) The only other overlap of two extended rules at the root position occurs between rules of the form (6.2), that is $-(w_i\phi_i) \Rightarrow (n_i - 1)(w_i\phi_i) - (w_i r_i)$. Then the overlapping term is $-w =_{\mathrm{AC}} -(w_1\phi_1) =_{\mathrm{AC}} -(w_2\phi_2)$ where we let $w =_{\mathrm{AC}} w_1\phi_1 =_{\mathrm{AC}} w_2\phi_2$. Figure 6.5 illustrates why this peak is also redundant. We get the dotted proof from the dashed proof by putting the context $(n_1 - 1)w - []$ around it and using the same construction as in case (b). The dotted proof stays below the bound $-w$, because the normal forms in the transformed proof contain no negated occurrence of the maximal product w, and the ordering is constructed in such a way that any number of positive occurrences is always smaller than one negative occurrence. □

Corollary 6.8 *Any critical ground term t is of the form $n\phi$ where $n > 0$.*

$$\begin{tikzcd}
& & (n_1-1)w-[n_1w] \arrow[dll, "S_1"'] \arrow[drr, "S_2"] \arrow[dd, "S_1"] & & \\
(n_1-1)w-[r_1w_1] \arrow[dr, dashed, "*", "\mathrm{CR}\cup S_N"'] \arrow[ddd, "*\;\mathrm{CR}"] & & & & -[(n_1-n_2)w+r_2w_2]+(n_1-1)w \arrow[dl, dashed, "*"', "\mathrm{CR}\cup S_N"] \arrow[ddd, "*\;\mathrm{CR}"] \\
& {} \arrow[rr, dashed, leftrightarrow, "*", "\mathrm{AC}"'] & {} \arrow[d, "*\;\mathrm{CR}"'] & {} & \\
& & -w \arrow[dll, "S_1"] \arrow[drr, "S_2"] & & \\
(n_1-1)w-r_1w_1 \arrow[rrrr, dotted, leftrightarrow, "\prec -w", "\mathrm{AC}\cup\mathrm{CR}\cup S_N"'] & & & & (n_2-1)w-r_2w_2
\end{tikzcd}$$

Figure 6.5: Redundancy of an extended peak

We will now use this to give a sufficient criterion for a term t being in the critical closure of a term s.

Lemma 6.9 *Let s and t be* CR*-irreducible ground terms, and let $n\phi$ be the maximal monomial of t. If $n\phi \preceq s$ then $t \in \mathrm{cc}_{\mathrm{CR}}(s)$.*

Proof: Since t is irreducible with respect to CR, and has the maximal monomial $n\phi$, it has the form $n\phi + p$ where $\phi \succ p$ and $n \succeq_{\mathbb{Z}} 1$. Since $s \succeq n\phi$, we have $s \succ p$. Suppose t were not in $\mathrm{cc}_{\mathrm{CR}}(s)$. Then there would be a critical term $m\psi$ such that $t \succeq m\psi \succ s$, so in particular $n\phi \succeq m\psi \succ s$, a contradiction to $n\phi \preceq s$. □

We will need this to show that CR-Migration is a simplification.

6.4 Simplification

In this section we present simplification rules and an admissible simplification function for commutative rings.

CR-*Cancellation* $\qquad \dfrac{[\neg](s_1 + p \approx s_2 + q)}{[\neg](p \approx q)}\rightarrow$

if (i) $s_1 =_{\mathrm{AC}} s_2$.

CR-*Migration* $\qquad \dfrac{[\neg](-s + p \approx q)}{[\neg](p \approx s + q)}\rightarrow$

Lemma 6.10 CR-*Cancellation and* CR-*Migration are simplification rules.*

Proof: (CR-Cancellation) In the proof we don't need to distinguish the AC-equivalent terms s_1 and s_2, instead we always write s. Let

$$\begin{aligned} L &= [\neg](s + p \approx s + q) \text{ and} \\ L' &= [\neg](p \approx q). \end{aligned}$$

Clearly $s + p \approx s + q$ is CR_1-equivalent to $p \approx q$, hence $\{L\} \cup \mathrm{CR}_1 \models L'$. It remains to show

$$\{[\neg](p \approx q)\} \cup \mathrm{Trans}_L \models_I [\neg](s + p \approx s + q).$$

If the literals are positive this is immediate by monotonicity, which holds in any candidate model I_N. For negative literals we have to show

$$\{s + p \approx s + q\} \cup \mathrm{Trans}_L \models_I p \approx q,$$

where we must be careful not to exceed the bound on transitivity. Let I_N be some candidate model that satisfies Trans_L and $s + p \approx s + q$. That is, there exists a valley proof $s + p \Downarrow_{\mathrm{CR} \cup S_N} s + q$.

(1) Suppose one of s, p or q is not in CR-normal form. Then we have the following situation:

$$\begin{tikzcd}
s+p \arrow[d, "\mathrm{CR}"', "*"] \arrow[dr, "*", "\mathrm{CR}\cup S"'] & & & s+q \arrow[d, "*"', "\mathrm{CR}"] \arrow[dl, "*"', "\mathrm{CR}\cup S"] \\
s'+p' \arrow[dr, dashed, "*", "\mathrm{CR}\cup S"'] & {} \arrow[r, leftrightarrow, "*", "E"'] & {} & s'+q' \arrow[dl, dashed, "*"', "\mathrm{CR}\cup S"] \\
& {} \arrow[r, dashed, leftrightarrow, "*", "E"'] & {} &
\end{tikzcd}$$

If we can cancel s' from $s' + p' \approx s' + q'$ then we obtain $p' \Downarrow q'$, which implies $p \Downarrow q$. So assume from now on that s, p and q are in CR-normal form.

(2) If s is a proper sum then it suffices to cancel one summand at a time. So assume $s = \phi$ or $s = -\phi$.

(3) If $s = -\phi$ then we add the context $\phi + [\,]$ around every term in the proof and normalize. Since ϕ is smaller than $-\phi$ which is bounded by the maximal term, transitivity holds for all terms in this proof, and we obtain $p \Downarrow q$.

(4) Otherwise $s = \phi$. If ϕ is not a maximal summand in $\phi + p$ or $\phi + q$, then we may add the context $-\phi + [\,]$ around every term in the proof and proceed as in (3).

(5) It remains to consider the case where $s = \phi$ is a maximal summand. We write S_ϕ for the set of rewrite rules of the form $u + n\phi \Rightarrow u + r$ in S_N. We let $p = (m-1)\phi + p'$, $q = (n-1)\phi + q'$, and move steps with S_ϕ to the front of the rewrite sequences. Assume without loss of generality $m \geq n$.

$$\begin{tikzcd}
m\phi+p' \arrow[d, "S_\phi"', "*"] \arrow[dr, "*", "\mathrm{CR}\cup S"'] & & & n\phi+q' \arrow[d, "*"', "S_\phi"] \arrow[dl, "*"', "\mathrm{CR}\cup S"] \\
n_0\phi+r_1+p' \arrow[dr, dashed, "*", "\mathrm{CR}\cup S"'] & {} \arrow[r, leftrightarrow, "*", "E"'] \arrow[dr, phantom, "(1)"] & {} & n_0\phi+r_2+q' \arrow[dl, dashed, "*"', "\mathrm{CR}\cup S"] \\
& {} \arrow[r, dashed, leftrightarrow, "*", "E"'] & {} &
\end{tikzcd}$$

Observe that $n_0\phi$ is not reduced at all, and that p' and q' are only reduced in the valley proof at the bottom. By taking the valley proof apart we obtain rewrite sequences $(m - n_0)\phi \overset{*}{\Rightarrow} r_1$ and $(n - n_0)\phi \overset{*}{\Rightarrow} r_2$, and a valley proof $r_1 + p' \Downarrow r_2 + q'$. We will construct the desired small proof from these parts.

We first consider the peak formed by the S_ϕ-reductions. Since $(m - n_0)\phi \preceq_{\mathrm{CR}} m\phi$ and transitivity holds up to $m\phi$, this peak converges, even if one of the rewrite sequences is empty. We obtain the dashed proof in the following diagram:

$$
\begin{tikzcd}
& & (m-n_0)\phi \arrow[dll, "*", "S_\phi"'] \arrow[drr, "*"', "S_\phi"] & & \\
r_1 \arrow[dr, dashed, "*", "\mathrm{CR}\cup S"'] & & (2) & & (m-n)\phi + r_2 \arrow[dl, dashed, "*"', "\mathrm{CR}\cup S"] \\
& {} \arrow[rr, dashed, leftrightarrow, "*", "E"'] & & {} &
\end{tikzcd}
$$

Finally, in the following proof we combine the valley proof (1) from above with this proof (2), under contexts $-r_1 + (m-1)\phi + [\,]$ and $-[\,] + (m-1)\phi + r_2 + q'$, respectively:

$$
\begin{tikzcd}[column sep=small]
-r_1 + (m-1)\phi + r_1 + p' \arrow[dd, "*", "\mathrm{CR}"'] \arrow[ddr, "*", "\mathrm{CR}\cup S"'] & & & & & & -((m-n)\phi + r_2) + (m-1)\phi + r_2 + q' \arrow[ddl, dashed, "*", "\mathrm{CR}\cup S"'] \arrow[dd, "*"', "\mathrm{CR}"] \\
& (1) & & -r_1 + (m-1)\phi + r_2 + q' \arrow[dl, "*", "\mathrm{CR}\cup S"'] \arrow[dr, "*"', "\mathrm{CR}\cup S"] & & (2) & \\
(m-1)\phi + p' & {} \arrow[r, leftrightarrow, "*", "E"'] & {} & & {} \arrow[r, leftrightarrow, "*", "E"'] & {} & (n-1)\phi + q'
\end{tikzcd}
$$

By normalizing this proof we obtain a proof that contains $(m - 1)\phi$ as its maximal monomial, hence transitivity holds for all terms in the proof and we obtain a valley proof $p = (m - 1)\phi + p' \Downarrow (n - 1)\phi + q'$. Note that the normalization removes all occurrences of $-\phi$ in the dashed part of the proof.

(CR-Migration) Let

$$
\begin{aligned}
C &= [\neg](-s + p \approx q) \text{ and} \\
D &= [\neg](p \approx s + q).
\end{aligned}
$$

Clearly $\{C\} \cup \mathrm{CR}_1 \models D$. It remains to show

$$\{[\neg](-s + p \approx q)\} \cup \mathrm{Trans}_C \models_I [\neg](p \approx s + q).$$

(1) Suppose the literals are positive, and consider some candidate model I_N such that $-s + p \approx q$ and Trans_C hold in I_N. That is, there exists a valley proof $-s + p \Downarrow_{\mathrm{CR}\cup S_N} q$. We place the context $s + [\,]$ around every term in the proof and normalize with respect to CR, and obtain a proof $p \overset{*}{\Leftrightarrow}_{\mathrm{AC}\cup\mathrm{CR}\cup S_N} s + q$ where each term is smaller than $-s + p$. Hence $p \approx s + q$ is true in I_N by transitivity.

(2) Otherwise the literals are negative. Then we have to show

$$\{p \approx s + q\} \cup \mathrm{Trans}_C \models_I -s + p \approx q.$$

Let I_N be a candidate model such that $p \approx s + q$ and Trans_C hold in I_N. Then there is a valley proof $p \Downarrow_{\mathrm{CR} \cup S_N} s + q$. We add the context $-s + [\,]$ around every term in the proof, normalize with respect to CR, and obtain an equational proof $-s+p \overset{*}{\Leftrightarrow}_{\mathrm{AC} \cup \mathrm{CR} \cup S_N} q$. Since any monomial in this proof is bounded by a monomial in $-s + p$ or q, all the terms of the proof are in $\mathrm{cc}_{\mathrm{CR}}(-s + p)$ or $\mathrm{cc}_{\mathrm{CR}}(q)$, whichever is larger. Then transitivity yields $-s + p \Downarrow q$. □

Next we present the simplification rules $\mathrm{Simp}_{\mathrm{CR}}$-Sum Contraction, $\mathrm{Simp}_{\mathrm{CR}}$-Summand Rewriting and $\mathrm{Simp}_{\mathrm{CR}}$-Isolation, which are sufficient to obtain CR-normal forms of literals.

$\mathrm{Simp}_{\mathrm{CR}}$-*Sum Contraction* $\qquad \dfrac{[\neg](s - s' + p \approx q)}{[\neg](p \approx q)} \rightarrow$

if (i) $s =_{\mathrm{AC}} s'$, (ii) $-s'$ is a maximal summand in $s - s' + p$, (iii) $s - s' + p \succeq q$.

$\mathrm{Simp}_{\mathrm{CR}}$-*Summand Rewriting* $\qquad \dfrac{[\neg](u[l] + p \approx q)}{[\neg](u[r] + p \approx q)} \rightarrow$

if (i) $l \Rightarrow r$ is a rule in CR, (ii) $u[l]$ is a maximal summand in $u[l'] + p$, and (iii) $u[l] + p \succeq q$.

$\mathrm{Simp}_{\mathrm{CR}}$-*Isolation* $\qquad \dfrac{[\neg](n_1\phi_1 + r_1 \approx n_2\phi_2 + r_2)}{[\neg]((n_1 - n_2)\phi_1 \approx r_2 - r_1)} \rightarrow$

if (i) $\phi_1 =_{\mathrm{AC}} \phi_2$, (ii) ϕ_1 is a product, (iii) ϕ_1 is irreducible with respect to CR, (iv) $n_1 \geq n_2$, (v) $n_2 \neq 0$ or $r_1 \neq 0$, and (vi) $\phi_1 \succ r_1$ and $\phi_2 \succ r_2$.

We let $\mathrm{Simp}_{\mathrm{CR}}(L)$ consist of all literals L' such that there exists a simplification by $\mathrm{Simp}_{\mathrm{CR}}$-Sum Contraction, $\mathrm{Simp}_{\mathrm{CR}}$-Summand Rewriting or $\mathrm{Simp}_{\mathrm{CR}}$-Isolation with premise L and conclusion L'.

Lemma 6.11 $\mathrm{Simp}_{\mathrm{CR}}$ *is an admissible simplification function for* $\mathrm{Norm}_{\mathrm{CR}}$.

Proof: $\mathrm{Simp}_{\mathrm{CR}}$-Sum Contraction and $\mathrm{Simp}_{\mathrm{CR}}$-Summand Rewriting are instances of CR-Rewriting, and $\mathrm{Simp}_{\mathrm{CR}}$-Isolation can be obtained as a sequence of CR-Cancellation and CR-Migration simplifications. Hence they are simplification rules.

The proof of admissibility is strictly analogous to the one for abelian groups. □

6.5 The inference system

We obtain the following inference systems. Note that the inference rules except for CR-Superposition are essentially the same as their counterparts for abelian groups.

CR-*Sum Contraction* $\qquad \dfrac{[\neg](s - s' + p \approx q) \ \vee\ C}{[\neg](p \approx q) \ \vee\ C}$

if (i) $s =_{\mathrm{AC}} s'$, (ii) $-s'$ is a maximal summand in $s - s' + p$, and (iii) $s - s' + p \succeq q$.

CR-*Summand Rewriting* $$\frac{[\neg](u[l'] + p \approx q) \vee C}{[\neg](u[r] + p \approx q) \vee C}$$

if (i) $l \Rightarrow r$ is a rule in CR, (ii) $l =_{\mathrm{AC}} l'$, (iii) $u[l']$ is a maximal summand in $u[l'] + p$, and (iv) $u[l'] + p \succeq q$.

CR-*Isolation* $$\frac{[\neg](n_1\phi_1 + r_1 \approx n_2\phi_2 + r_2) \vee C}{[\neg]((n_1 - n_2)\phi_1 \approx r_2 - r_1) \vee C}$$

if (i) $\phi_1 =_{\mathrm{AC}} \phi_2$, (ii) ϕ_1 is a product, (iii) ϕ_1 is irreducible with respect to CR, (iv) $n_1 \geq n_2$, (v) $n_2 \neq 0$ or $r_1 \neq 0$, and (vi) $\phi_1 \succ r_1$ and $\phi_1 \succ r_2$.

CR-*Extension Superposition* $$\frac{n_1\phi_1 \approx r_1 \vee C_1 \qquad n_2\phi_2 \approx r_2 \vee C_2}{r_1\psi_1 \approx (n_1 - n_2)\phi + r_2\psi_2 \vee C_1 \vee C_2}$$

if (i) ϕ_1 and ϕ_2 are products, (ii) $\phi =_{\mathrm{AC}} \mathrm{lcm}(\phi_1, \phi_2) =_{\mathrm{AC}} \phi_1\psi_1 =_{\mathrm{AC}} \phi_2\psi_2$, (iii) ϕ_1 and ϕ_2 are irreducible with respect to CR, (iv) $\psi_1 \neq 1$, (v) $n_1 > n_2$ or $\psi_2 \neq 1$, (vi) (a) $n_1 \geq n_2 \geq 2$, or (b) $n_1 \geq 2$, $n_2 = 1$, ϕ_2 is a proper product with $\gcd(\phi_1, \phi_2) \neq 1$, or (c) $n_1 = n_2 = 1$ and ϕ_1 and ϕ_2 are proper products with $\gcd(\phi_1, \phi_2) \neq 1$, (vii) $\phi_1 \succ r_1$ and $\phi_2 \succ r_2$.

The main premise of this inference is

$$r_1\psi_1 \not\approx n_1\phi \vee n_1\phi \not\approx (n_1 - n_2)\phi + r_2\psi_2 \vee r_1\psi_1 \approx (n_1 - n_2)\phi + r_2\psi_2.$$

CR-*Superposition A* $$\frac{n\phi \approx r \vee D \quad [\neg]((m\psi)[n\phi'] \approx q) \vee C}{[\neg]((m\psi)[r] \approx q) \vee C \vee D}$$

if (i) $\phi =_{\mathrm{AC}} \phi'$, (ii) $m, n \geq 1$, (iii) ϕ and ψ are products, (iv) ϕ and ψ are irreducible with respect to CR, and (v) $\phi \succ r$ and $\psi \succ q$.

CR-Superposition A combines cases (CR.S2), (CR.S3a) and (CR.S4a) of the symmetrization.

CR-*Superposition B* $$\frac{n\phi \approx r \vee D \quad [\neg]((m\psi)[l'] \approx q) \vee C}{[\neg]((m\psi)[r + t] \approx q) \vee C \vee D}$$

if (i) $l' =_{\mathrm{AC}} n\phi + t$, (ii) $m, n \geq 1$, (iii) ϕ and ψ are products, (iv) either $n \geq 2$ or ϕ is a proper product, (v) ϕ and ψ are irreducible with respect to CR, (vi) $\phi \succ r$ and $\psi \succ q$.

Superposition B combines cases (CR.S3b) and (CR.S4b).

CR-*Superposition C* $$\frac{n\phi \approx r \vee D \quad [\neg]((m\psi)[l'] \approx q) \vee C}{[\neg]((m\psi)[s \cdot r] \approx q) \vee C \vee D}$$

if (i) $l' =_{\mathrm{AC}} n(s \cdot \phi)$, (ii) $m \geq 1$, $n \geq 2$, (iii) ϕ and ψ are products, (iv) ϕ and ψ are irreducible with respect to CR, (v) $\phi \succ r$ and $\psi \succ q$.

Superposition C is for case (CR.S4c).

CR-*Superposition D* $$\frac{n\phi \approx r \,\vee\, D \quad [\neg]((m\psi)[l'] \approx q) \,\vee\, C}{[\neg]((m\psi)[s \cdot r] \approx q) \,\vee\, C \,\vee\, D}$$

if (i) $l' =_{\mathrm{AC}} n(s \cdot \phi)$, (ii) $m \geq 1$, $n \geq 2$, (iii) ϕ and ψ are products, (iv) ϕ and ψ are irreducible with respect to CR, (v) $\phi \succ r$ and $\psi \succ q$.

CR-Superposition D is for case (CR.S4d).

CR-*Superposition E* $$\frac{n\phi \approx r \,\vee\, D \quad [\neg]((m\psi)[-\phi'] \approx q) \,\vee\, C}{[\neg]((m\psi)[(n-1)\phi - r] \approx q) \,\vee\, C \,\vee\, D}$$

if (i) $\phi =_{\mathrm{AC}} \phi'$, (ii) $m \geq 1$, $n \geq 2$, (iii) ϕ and ψ are products, (iv) ϕ and ψ are irreducible with respect to CR, (v) $\phi \succ r$ and $\psi \succ q$.

CR-Superposition E is for case (CR.S4e)

CR-*Superposition F* $$\frac{n\phi \approx r \,\vee\, D \quad [\neg]((m\psi)[l'] \approx q) \,\vee\, C}{[\neg]((m\psi)[(n-1)\phi - r] \approx q) \,\vee\, C \,\vee\, D}$$

if (i) $l' =_{\mathrm{AC}} -(s \cdot \phi)$, (ii) $m \geq 1$, $n \geq 2$, (iii) ϕ and ψ are products, (iv) ϕ and ψ are irreducible with respect to CR, (v) $\phi \succ r$ and $\psi \succ q$.

CR-Superposition F is for case (CR.S4f).

CR-*Reflexivity Resolution* $$\frac{0 \not\approx 0 \,\vee\, C}{C}$$

CR-*Equality Factoring* $$\frac{n\phi \approx r \,\vee\, n\phi' \approx r' \,\vee\, C}{r \not\approx r' \,\vee\, n\phi \approx r' \,\vee\, C}$$

if (i) $\phi =_{\mathrm{AC}} \phi'$, (ii) $n \geq 1$, (iii) ϕ is CR-atomic, (iv) ϕ is irreducible with respect to CR, (v) $\phi \succ r$ and $\phi' \succ r'$, (vi) $r \succeq r'$.

We let $\mathsf{Sup}_{\mathrm{CR}}$ be the set of these inferences, where for each inference the same restrictions by selection as in the general case apply.

Theorem 6.12 $\mathsf{Sup}_{\mathrm{CR}}$ *is refutationally complete for* CR_1.

Proof: Again this follow from Theorem 4.20 in combination with Propositions 6.1 and 6.4, and Lemmas 6.6 and 6.11. Note that by using Theorem 6.7 we restrict CR-Extension Superposition to critical extension peaks. □

7

Modules

Consider a ring $\langle R, +_R, \cdot_R, -_R, 0_R, 1_R \rangle$. A *(left) module over* R is an abelian group $\langle M, +_M, -_M, 0_M \rangle$ together with an operation $* : R \times M \to M$ of R on M such that

$$1 * s = s \tag{7.1}$$
$$(a \cdot_R b) * s = a * (b * s) \tag{7.2}$$
$$(a +_R b) * x = a * x +_M b * x \tag{7.3}$$
$$a * (x +_M y) = a * x +_M a * y. \tag{7.4}$$

Note that by (7.1) we only consider unitary modules. A module over a field is called a *vector space*.

We will develop a superposition calculus for proving validity in the class of all R-modules over some fixed ring R. The ring R must be equipped with a well-ordering $\succ_R$ such that $0 \prec_R 1$ and $1 \prec_R r$ for any $r \in R \setminus \{0, 1\}$. We restrict R further to be either the ring of integers with respect to the ordering

$$0 \prec_R 1 \prec_R 2 \prec_R \ldots \prec_R -1 \prec_R -2 \prec_R \ldots$$

or a field. Note that these are integral domains. We will use that integral domains have no zero divisors and that they obey the cancellation law for multiplication.

The equation $-_M(x) = (-1) * x$ is valid in any module and allows to eliminate $-_M$ in a preprocessing step.

Assumption 7.1 *We assume from now on that $-_M$ does not occur.*

We obtain the following signature for modules:

$$\begin{aligned} +_M &: \mathsf{M} \times \mathsf{M} \longrightarrow \mathsf{M} \\ 0_M &: \;\longrightarrow \mathsf{M} \\ a_R &: \;\longrightarrow \mathsf{R} \qquad \text{for all } a \in R \\ +_R, \cdot_R &: \mathsf{R} \times \mathsf{R} \longrightarrow \mathsf{R} \\ -_R &: \mathsf{R} \longrightarrow \mathsf{R} \\ * &: \mathsf{R} \times \mathsf{M} \longrightarrow \mathsf{M} \end{aligned}$$

Apart from these symbols in F_{M} there are the free function symbols

$$f : \mathsf{M}^{\alpha(f)} \longrightarrow \mathsf{M}$$

in $F \setminus F_{\mathrm{M}}$. Terms of sort R denote elements of the ring R, and terms of sort M elements of the module. The following syntactical restrictions prevent equations between elements of R, which reflects that we have a hierarchical situation where R is fixed.

- Neither rewrite rules nor clauses may contain equations between R-terms.
- On the ground level only constants may occur as R-terms.
- On the nonground level, rules and literals may contain both constants and variables, but no nested terms of sort R.
- Constraints may contain arbitrary R-terms.

From now on we drop the indices that distinguish operations in M from operations in R, as the distinction will always be clear from the context. Let $\widehat{\mathrm{M}}$ be the following constrained term rewriting system modulo AC for a module over R.

$$x + 0 \Rightarrow x \tag{M.1}$$
$$0 * x \Rightarrow 0 \tag{M.2}$$
$$1 * x \Rightarrow x \tag{M.3}$$
$$v * 0 \Rightarrow 0 \tag{M.4}$$
$$v * (x + y) \Rightarrow v * x + v * y \tag{M.5}$$
$$v_1 * (v_2 * x) \Rightarrow v * x \;\; [v = v_1 \cdot v_2] \tag{M.6}$$
$$x + x \Rightarrow v * x \;\; [v = 1 + 1] \tag{M.7}$$
$$v_1 * x + x \Rightarrow v * x \;\; [v = v_1 + 1] \tag{M.8}$$
$$v_1 * x + v_2 * x \Rightarrow v * x \;\; [v = v_1 + v_2] \tag{M.9}$$
$$y + x + x \Rightarrow y + v * x \;\; [v = 1 + 1] \tag{M.7e}$$
$$y + v_1 * x + x \Rightarrow y + v * x \;\; [v = v_1 + 1] \tag{M.8e}$$
$$y + v_1 * x + v_2 * x \Rightarrow y + v * x \;\; [v = v_1 + v_2] \tag{M.9e}$$

We let $\mathrm{AC} = \mathrm{AC}(+_M)$ and $\mathrm{M} = \mathrm{AC}\backslash gnd(\widehat{\mathrm{M}})$. That is, M is a ground term rewriting system that does rewriting with AC-matching. It is obtained from $\widehat{\mathrm{M}}$ by instantiating variables of sort R in such a way by elements of R that the constraints are satisfied. In this way term rewriting with M incorporates computations in R.

Lemma 7.2 M *is locally ground confluent and locally ground coherent modulo* AC.

Proof: Local coherence modulo AC holds, since the corresponding cliff converges for (M.1), and since M contains the extended rules (M.7e)–(M.9e) for (M.7)–(M.9).

For local ground confluence modulo AC we have to show that any ground critical pair modulo AC can be joined by a ground rewrite proof of M modulo AC. For example, let us consider a critical pair

$$c * (s + a_3 * t) \overset{\mathrm{M.9e}}{\Leftarrow} c * (a_1 * t + s + a_2 * t) \overset{\mathrm{M.5}}{\Rightarrow} c * (s + a_1 * t) + c * (a_2 * t)$$

for some ground terms s and t and $a_1, a_2, a_3 \in R$ such that $a_3 = a_1 + a_2$. This is a ground instance of the constrained critical pair

$$v * (y + v_3 * x) \overset{\mathrm{M.9e}}{\Leftarrow} v * (v_1 * x + y + v_2 * x) \overset{\mathrm{M.5}}{\Rightarrow} v * (v_1 * x) + v * (y + v_2 * x) \;\; [v_3 = v_1 + v_2].$$

To show that all ground instances of this critical pair converge, we first normalize both terms of the critical pair while collecting the constraints of the rules we apply. We obtain the terms

$$\begin{array}{ll} v * y + v_4 * x \;\; [v_3 = v_1 + v_2 \, \wedge \, v_4 = v \cdot v_3] & \text{and} \\ v_7 * x + v * y \;\; [v_3 = v_1 + v_2 \, \wedge \, v_5 = v \cdot v_1 \, \wedge \, v_6 = v \cdot v_2 \, \wedge \, v_7 = v_5 + v_6] & \end{array}$$

as normal forms. We have to check that the normalization is possible for each ground instance of the critical pair. For M this is always the case, as the constraints do not restrict the applicability of rules. It remains to prove that the two irreducible ground terms obtained by reduction are AC-equivalent. On the nonground level this amounts to checking that the constraints on the normal forms together imply AC-equivalence of the two irreducible terms. The constraints can be simplified by substituting definitions and eliminating variables which not occur in the term:

$$\begin{array}{l} v * y + v_4 * x \;\; [v_4 = v \cdot (v_1 + v_2)] \\ v_7 * x + v * y \;\; [v_7 = v \cdot v_1 + v \cdot v_2] \end{array}$$

Since R is a commutative ring, we may use rewriting with $\widehat{\mathrm{CR}}$ to normalize the constraints:

$$\begin{array}{l} v * y + v_4 * x \;\; [v_4 = v \cdot v_1 + v \cdot v_2] \\ v_7 * x + v * y \;\; [v_7 = v \cdot v_1 + v \cdot v_2] \end{array}$$

Finally, we propagate the constraints into the terms and verify AC-equivalence with respect to the AC-axioms for both the ring and the module operations. Since associativity and commutativity of $+$ and $\cdot$ hold in any commutative ring, this implies equality of the constants from R and hence AC-equivalence with respect to the module operation $+$ for any two ground instances.

Technically, this procedure can be emulated by propagating the constraints into the rules of $\widehat{\mathrm{M}}$, resulting in a system $\widehat{\mathrm{M}}'$, and showing convergence of all critical peaks with respect to $\widehat{\mathrm{M}}' \cup \widehat{\mathrm{CR}}$ modulo AC. E.g., (M.9) becomes

$$v_1 * x + v_2 * x \Rightarrow (v_1 + v_2) * x.$$

Note that the left-hand sides of rules are the same in $\widehat{\mathrm{M}}$ and $\widehat{\mathrm{M}}'$, hence the critical pairs stay essentially the same. We have implemented this in Prolog and used it to show that all critical pairs between rules in $\widehat{\mathrm{M}}$ converge. □

For modules we use the notational conventions of abelian groups, and additionally the following: When we write $b * \alpha$ then this may also denote α for $b = 1$ and 0 for $b = 0$. Also note that an R-expression like for example $(a+b)$ in a ground term $(a+b)*t$ denotes the constant obtained by evaluating the expression, and not the expression itself, since a compound term of sort R is not allowed in this context. We will use $s-t$ as an abbreviation for $s + (-1) * t$ in M and $a - b$ as an abbreviation for $a + (-b)$ in R.

7.1 Termination

The termination ordering $\succeq_{\mathrm{M}}$ for modules follows the same schema as the termination ordering for commutative rings.

Again we use the lexicographic combination of a problem-specific TPO, an ordering by polynomial interpretation, and the AC-RPO. We let $F_I = F_{\mathrm{M}} = \{+, *, -, 0, 1\} \cup R$, assume a total precedence $\succeq_p$ on $F \setminus F_I$, and let $\succeq_p$ also denote its TPO-admissible extension to F. For the TPO-status we again use the method of Section 3.3. Let $\succeq_c$ be a quasi-ordering on F_C. We define an ordering $\succeq_t$ on terms over $F_I \cup F_C$ that extends $\succeq_c$ and satisfies the conditions of Lemma 3.9.

Let $\widehat{\mathrm{D}}$ be the term rewriting system consisting of the distributivity rule (M.5), and let $\mathrm{D} = \mathrm{AC}\backslash gnd(\widehat{\mathrm{D}})$. Then D is convergent modulo AC. We will assign a complexity κ to any ground term over $F_I \cup F_C$ in D-normal form. Let t be such a ground term. It has the form

$$t = a_{11} * \cdots * a_{1k_1} * c_1 + \cdots + a_{n1} * \cdots * a_{nk_n} * c_n$$

where $n \geq 1$, $k_i \geq 0$, $c_i \in F_C \cup \{0,1\}$, and $a_{ij} \in R$ for $i = 1, \ldots, n$ and $j = 1, \ldots, k_i$. In the ordering we need to identify constants c_i in the same $\sim_c$-equivalence class, hence we assume a function $\mathrm{rep} : F_C \rightarrow F_C$ such that $\mathrm{rep}(c_i) = \mathrm{rep}(c_j)$ if and only if $c_i \sim_c c_j$. We extend rep to $F_C \cup \{0,1\}$ by $\mathrm{rep}(0) = 0$ and $\mathrm{rep}(1) = 1$. The ordering $\succeq_c$ is extended to $F_C \cup \{0,1\}$ such that $c \succ_c 1 \succ_c 0$ for any constant c in F_C. We let

$$\begin{aligned}
\mathrm{occ}(t, c_i) &= \{j \mid c_j \sim_c c_i\} \\
\#(t, c_i) &= |\mathrm{occ}(t, c_i)| \\
\mathrm{cs}(t, c_i) &= \{\langle a_{11}, \ldots, a_{1k_1}\rangle, \ldots, \langle a_{n1}, \ldots, a_{nk_n}\rangle\}
\end{aligned}$$

That is, $\mathrm{occ}(t, c_i)$ is the set of indices of the occurrences of constants in the same $\sim_c$-equivalence class as c_i, $\#(t, c_i)$ is the number of these occurrences, and $\mathrm{cs}(t, c_i)$ is the multiset of the tuples of coefficients associated with these occurrences. To each equivalence class we associate the tuple $\langle \mathrm{rep}(c_i), \#(t, c_i), \mathrm{cs}(t, c_i)\rangle$. Finally, we let $\kappa(t)$ be the set of tuples for the constants from F_C that occur in t. We order these complexities according to the multiset extension of the lexicographic combination of $\succeq_c$, $>$ and the multiset extension of the length-lexicographic extension of $\succeq_R$. We denote the ordering on complexities by $\succeq_\kappa$. Then we define the ordering $\succeq_t$ on terms over $F_I \cup F_C$ by $s \succeq_t t$ if and only if $\kappa(\mathrm{D}(s)) \succeq_\kappa \kappa(\mathrm{D}(t))$ where s and t are terms over $F_I \cup F_C$. Finally we get the TPO-status $\succeq_1^{st}$ as the status derived from $\succeq_t$, and let

$$\succeq_1(\succeq_p) = \succeq_{tpo}(\succeq_p, \succeq_1^{st}).$$

Lemma 7.3 *Let $\succeq_p$ be a well-ordering on $F \setminus F_{\mathrm{M}}$. Then $\succeq_1(\succeq_p)$ is a total* $\mathrm{AC} \cup \mathrm{D}$*-compatible and* $\mathrm{AC} \cup \mathrm{D}$*-antisymmetric simplification quasi-ordering on ground terms that contains* $\mathrm{M} \setminus \mathrm{D}$.

Proof: (*Simplification quasi-ordering*) We begin by showing that $\succeq_1^{st}$ is a TPO-status by Lemma 3.9.

(*Strictly compatible with contexts*) Let $f_{u[]}$ be the function that maps any complexity $\kappa(\mathrm{D}(t))$ to the complexity $\kappa(\mathrm{D}(u[t]))$, and let

$$\mathrm{D}(t) = a_{11} * \cdots * a_{1k_1} * c_1 + \cdots + a_{n1} * \cdots * a_{nk_n} * c_n.$$

To see that $f_{u[]}$ is well-defined, observe that a term in normal form can be reconstructed up to $\mathrm{AC} \cup \sim_c$-equivalence from its complexity, and that κ maps terms in an $\mathrm{AC} \cup \sim_c$-equivalence class to the same complexity. We show that $f_{u[]}$ is strictly monotonic for any context $u[]$ by considering contexts of depth one.

(1) Consider $u = s + []$.
(1.1) Suppose $s = b_1 * \cdots * b_l * d$.
(1.1.1) Suppose no c_i is $\sim_c$-equivalent to d. Then

$$f_{u[]}(\kappa(t)) = \kappa(t) \cup \{\langle \mathrm{rep}(d), 1, \{\langle b_1, \ldots, b_l \rangle\}\rangle\}$$

and $f_{u[]}$ is strictly monotonic.

(1.1.2) Otherwise $d \sim_c c_i$ for some $i = 1, \ldots, k$. Suppose without loss of generality $d \sim_c c_1$. Then

$$\begin{aligned} f_{u[]}(\kappa(t)) = \{&\langle c_1, \#(t, c_1) + 1, \mathrm{cs}(t, c_1) \cup \{\langle b_1, \ldots, b_l \rangle\}\rangle\} \\ &\cup \{\langle c_i, \#(t, c_i), \mathrm{cs}(t, c_i)\rangle \mid c_i \not\sim_c c_1\}. \end{aligned}$$

Consider the function that maps a tuple $\langle c, n, M \rangle$ to $\langle c, n+1, M \cup \{\langle b_1, \ldots, b_l \rangle\}\rangle$ if $c \sim_c d$ and to $\langle c, n, M \rangle$ otherwise. This function is strictly monotonic according to Proposition 2.8(7). Hence its multiset extension $f_{u[]}$ is strictly monotonic.

(1.2) Otherwise s is a proper sum, and $f_{u[]}$ can be obtained as a finite composition of the strictly monotonic functions of case (1.1). Hence $f_{u[]}$ is strictly monotonic.

(2) Consider $u = a * []$. Then $f_{u[]}$ maps any tuple $\langle a_1, \ldots, a_k \rangle$ of coefficients in the multiset in the third component to $\langle a, a_1, \ldots, a_k \rangle$. This mapping is strictly monotonic. By multiset extension, lexicographic product with identity functions for the first two components and again multiset extension we obtain $f_{u[]}$, which we conclude to be strictly monotonic.

(*Subterm property*) It suffices to consider contexts of depth one, then the subterm property follows by structural induction. We have to show (1) $s + t \succeq_t t$ and (2) $a * t \succeq_t t$. For (1) we observe that $\kappa(\mathrm{D}(s + t))$ has at least one additional tuple or a tuple whose second component increases when compared to $\kappa(\mathrm{D}(t))$. For (2) we observe that in each tuple in $\kappa(\mathrm{D}(t))$ the lengths of the tuples in the multiset in the third component increase by one.

(*Decreases infinite derivations*) Suppose there is some infinite descending chain

$$\kappa(\mathrm{D}(t_1)) \succ_\kappa \kappa(\mathrm{D}(t_2)) \succ_\kappa \ldots .$$

Then there exists an infinite descending chain of tuples

$$\langle c_1, n_1, M_1 \rangle \succ \langle c_2, n_2, M_2 \rangle \succ \ldots .$$

Since $>$ on natural numbers and $\succ_R$ and its extensions are well-founded, there exists an infinite descending chain

$$c_1 \succ_c c_2 \succ_c \cdots$$

in the first component, where c_i occurs in some t_{j_i} for $1 \leq j_1 < j_2 < \ldots$. The constants $c_1, c_2, \ldots$ are from F_C, since there is no infinite descending chain starting in 0 or 1.

(*Constant dominance condition*) The constants are in the first component of each tuple, hence they dominate the ordering.

(AC$\cup$D-*compatible*) We have to show that $\succeq_t$ is AC$\cup$D-compatible. Suppose $s \Leftrightarrow_{\mathrm{AC}\cup\mathrm{D}} s' \succeq_t t' \Leftrightarrow_{\mathrm{AC}\cup\mathrm{D}} t$. Then

$$\kappa(\mathrm{D}(s)) = \kappa(\mathrm{D}(s')) \succeq_\kappa \kappa(\mathrm{D}(t)) = \kappa(\mathrm{D}(t'))$$

implies $s \succeq_t t$.

(AC∪D-*antisymmetric*) We have to show that the quasi-ordering $\succeq_t(\succeq_c)$ is AC∪D∪$\sim_c$-antisymmetric. Suppose $s \sim_t(\succeq_c)\, t$ and

$$\begin{aligned}\kappa(\mathrm{D}(s)) &= \{\langle c_1, m_1, M_1\rangle, \ldots, \langle c_k, m_k, M_k\rangle\} \text{ and}\\ \kappa(\mathrm{D}(t)) &= \{\langle d_1, n_1, N_1\rangle, \ldots, \langle d_l, n_l, N_l\rangle\}.\end{aligned}$$

Then $\kappa(\mathrm{D}(s)) \sim_\kappa \kappa(\mathrm{D}(t))$, $k = l$, and there exists a permutation π such that $c_i = d_{\pi(i)}$, $m_i = n_{\pi(i)}$ and $M_i = N_{\pi(i)}$ for $i = 1, \ldots, k$. We can now show that $\mathrm{D}(s) \overset{*}{\Leftrightarrow}_{\mathrm{AC}\cup\sim_c} \mathrm{D}(t)$. First we replace each constant c by its representative $\mathrm{rep}(c)$. Let s' and t' be the resulting terms. They can be written as $s' = s_1 + \cdots + s_k$ and $t' = t_1 + \cdots + t_l$ where s_i consists of the terms containing c_i and t_i of the terms containing d_i. From $M_i = N_{\pi(i)}$ we infer that c_i and $d_{\pi(i)}$ are associated with the same tuples of coefficients from R, and that hence $s_i =_{\mathrm{AC}} t_{\pi(i)}$ and in turn $s' =_{\mathrm{AC}} t'$. We conclude that $s =_{\mathrm{AC}\cup\mathrm{D}\cup\sim_c} t$.

(*Total*) Since $\succeq_c$, $\geq$ and $\succeq_R$ are total, and since multiset and lexicographic extension as well as lexicographic product preserve totality, $\succeq_t(\succeq_c)$ is total. Hence $\succeq_1$ is total by Lemma 3.10.

(*Orients rules from left to right*) Rules (M.1), (M.3) and (M.4) are oriented from left to right by the subterm property, rule (M.2) because 0 is the minimal term.

(M.6) decreases the length of the tuples of coefficients. More formally, consider some ground instance (M.6)σ and let $t = P_{F_I}(x\sigma)$, $b_1 = v_1\sigma$, $b_2 = v_2\sigma$ and $b = v\sigma$. Then we have to show that

$$\kappa(\mathrm{D}(b_1 * (b_2 * t))) \succ_\kappa \kappa(\mathrm{D}(b * t)).$$

Now consider some triple $\langle c_i, n_i, \{\bar{a}_1, \ldots, \bar{a}_{k_i}\}\rangle$ in $\kappa(\mathrm{D}(t))$. This corresponds to the triple $\langle c_i, n_i, \{b_1 b_2 \bar{a}_1, \ldots, b_1 b_2 \bar{a}_{k_i}\}\rangle$ in $\kappa(\mathrm{D}(b_1 * (b_2 * t)))$, which is greater than the corresponding triple $\langle c_i, n_i, \{b\bar{a}_1, \ldots, b\bar{a}_{k_i}\}\rangle$ in $\kappa(\mathrm{D}(b * t))$, because tuples of coefficients are ordered by the length-lexicographic extension of $\succeq_R$.

(M.7)–(M.9e) decrease the number of occurrences of some constants in the sum. We consider some ground instance (M.7)σ and let $t = P_{F_I}(x\sigma)$ and $b = v\sigma$. Then we have to show that

$$\kappa(\mathrm{D}(t + t)) \succ_\kappa \kappa(\mathrm{D}(b * t)).$$

Suppose

$$\kappa(\mathrm{D}(t)) = \{\langle c_1, n_1, M_1\rangle, \ldots, \langle c_k, n_k, M_k\rangle\}.$$

Then $\kappa(\mathrm{D}(t + t))$ contains a triple $\langle c_i, 2n_i, M_i \cup M_i\rangle$ where $n_i > 0$, which is greater than $\langle c_i, n_i, M_i'\rangle$ in $\kappa(\mathrm{D}(b * t))$ for $i = 1, \ldots, k$. For (M.8)–(M.9e) the proof is essentially the same, differences affect only the third component of triples. □

We let $\succeq^p_{\mathrm{M}}$ be the quasi-ordering induced by the following polynomial interpretation. It is essentially the same as for commutative rings:

$$\begin{aligned} p_0^{\mathrm{M}} &= 2\\ p_1^{\mathrm{M}} &= 2\\ p_+^{\mathrm{M}}(x, y) &= x + y + 5\\ p_*^{\mathrm{M}}(v, x) &= v \cdot x\\ p_f^{\mathrm{M}}(x_1, \ldots, x_n) &= x_1 + \cdots + x_n + 2 \qquad \text{for } f \text{ free.}\end{aligned}$$

Lemma 7.4 *$\succeq_{\mathrm{M}}^{p}$ is a total* AC*-compatible simplification quasi-ordering on ground terms that orients* D *from left to right.*

We let $\succeq_{\mathrm{M}}$ be the lexicographic combination of $\succeq_1$, $\succeq_{\mathrm{M}}^{p}$ and $\succeq_{acrpo}$ over an arbitrary precedence.

Lemma 7.5 *Let $\succ_p$ be a total precedence on $F \setminus F_{\mathrm{M}}$. Then $\succeq_{\mathrm{M}}(\succeq_p)$ is a total* AC*-antisymmetric and* AC*-compatible simplification quasi-ordering on ground terms and contains* M.

Proof: Analogous to the case of commutative rings. □

We will use the next lemma to prove that rules in the symmetrization are oriented from left to right.

Lemma 7.6 *Let α be an* M*-atomic term in* M*-normal form, let r be a ground term such that $\alpha \succ_{\mathrm{M}} r$, and let b', b'' and m be elements of R such that $b' \succ_R b''$. Then $b' * \alpha \succ_{\mathrm{M}} b'' * \alpha + m * r$.*

Proof: By a similar argument as for commutative rings we can infer $\alpha \succ_1 r$ from α being in M-normal form and D ∪ AC-antisymmetry of $\succeq_1$. Suppose

$$\kappa(\mathrm{D}(r)) = \{\langle c_1, n_1, M_1\rangle, \ldots, \langle c_k, n_k, M_k\rangle\}.$$

Then

$$\kappa(\alpha) = \{\langle c_\alpha, 1, \{\langle\rangle\}\rangle\} \succ_\kappa \kappa(\mathrm{D}(r))$$

implies $c_\alpha \succ_c c_i$ for $i = 1, \ldots, k$, since the second and third component of the tuple in $\kappa(\alpha)$ are minimal. If we now compare

$$\kappa(\mathrm{D}(b' * \alpha)) = \kappa(b' * \alpha) = \{\langle c_\alpha, 1, \{\langle b'\rangle\}\rangle\}$$

and

$$\begin{aligned}\kappa(\mathrm{D}(b'' * \alpha + m * r)) &= \kappa(b'' * \alpha) \cup \kappa(\mathrm{D}(m * r)) \\ &= \{\langle c_\alpha, 1, \{\langle b''\rangle\}\rangle\} \cup \{\langle c_1, n_1, M_1'\rangle, \ldots, \langle c_k, n_k, M_k'\rangle\}\end{aligned}$$

where we see that all triples in $\kappa(\mathrm{D}(m * r))$ are smaller than $\langle c_\alpha, 1, \{\langle b'\rangle\}\rangle$ in the first component, and that $\langle c_\alpha, 1, \{\langle b''\rangle\}\rangle$ is smaller than $\langle c_\alpha, 1, \{\langle b'\rangle\}\rangle$ in the third component. □

Proposition 7.7 M *is ground convergent modulo* AC.

Proof: Termination is proved by inclusion of M in $\succeq_{\mathrm{M}}$, hence confluence and coherence follow from the local properties proven in Lemma 7.2. □

7.2 Symmetrization

We start by defining M-normal forms. Remember that for each equation there has to be an M-equivalent normal form. M-equivalence is preserved by multiplying both sides of an equation with a unit. For $R = \mathbb{Z}$ the units are 1 and -1, and for R a field any nonzero element is a unit. This allows to make the coefficient on the maximal summand positive in the case of integers, and 1 in the case of fields, which are the minimal choices in the given orderings. An equation $l \approx r$ is in M-*normal form* if either (i) $l = r = 0$, or (ii) $l \succ r$, l is irreducible with respect to M and $l \approx r$ has one of the forms (a) $\alpha \approx r$ where α is M-atomic, or (b) $b * \alpha \approx r$ where α is M-atomic, $R = \mathbb{Z}$ and $b > 1$. The set of equations in M-normal forms is denoted by $\mathrm{Norm}_{\mathrm{M}}$. We use the following symmetrization function, where we write $b' \Rightarrow_b^m b''$ for $b' = b'' + mb$ and $b' \succ_R b''$:

$$\mathcal{S}_{\mathrm{M}}(0 \approx 0) = \emptyset \tag{M.S1}$$

$$\mathcal{S}_{\mathrm{M}}(\alpha \approx r) = \{\alpha \Rightarrow r\} \tag{M.S2}$$

$$\mathcal{S}_{\mathrm{M}}(b * \alpha \approx r) = \{b' * \alpha \Rightarrow b'' * \alpha + m * r \mid b' \Rightarrow_b^m b''\} \tag{M.S3}$$

Lemma 7.8 *$\mathcal{S}_{\mathrm{M}}$ is a symmetrization function for* M.

Proof: $\mathcal{S}_{\mathrm{M}}(l \approx r)$ is terminating by Lemma 7.6. For left-minimality the only nontrivial case is (M.S3), where we observe that for $b > 0$ and any two distinct integers b_1 and b_2 with difference mb not both b_1 and b_2 can be in the interval $[0, b)$. Hence at least one of them is greater than or equal to b with respect to $\succ_R$.

Next we have to show that the rules in the symmetrization $\mathcal{S}_{\mathrm{M}}(l \approx r)$ are a consequence of $l \approx r$ and M. The only interesting case is again (M.S3). From the normal form $b * \alpha \approx r$ we can derive (M.S3) in two steps, using the rules (M.6) and (M.9). Consider the critical pair

$$m * r \Leftarrow m * (b * \alpha) \stackrel{\mathrm{M.6}}{\Rightarrow} (mb) * \alpha.$$

This covers all rules in $\mathcal{S}_{\mathrm{M}}(l \approx r)$ with $b'' = 0$. The remaining rules are then obtained from critical pairs of the form

$$b'' * \alpha + m * r \Leftarrow b'' * \alpha + (mb) * \alpha \stackrel{\mathrm{M.9}}{\Rightarrow} b' * \alpha,$$

where $b' = b'' + mb$.

It remains to show convergence modulo AC of $\mathrm{M} \cup \mathcal{S}_{\mathrm{M}}(l \approx r)$. Local coherence is not a problem, since no rule in $\mathcal{S}_{\mathrm{M}}(l \approx r)$ has the only AC-symbol $+$ at the root position. For local confluence the only cases to consider are again those of overlaps of rules in $\mathcal{S}_{\mathrm{M}}(l \approx r)$ into M strictly below the root position, and of overlaps of $\mathcal{S}_{\mathrm{M}}(l \approx r)$ with itself.

(1) Overlaps of rules in M into rules of $\mathcal{S}_{\mathrm{M}}(l \approx r)$ cannot occur, as the left-hand sides are normalized with respect to M. There is no critical overlap of (M.S2) into M, since (M.S2) has a free function symbol at the root. It remains to consider overlaps of (M.S3) into M. Since l is irreducible with respect to M, there can be no overlaps at the root position. There exist overlaps below the root position into (M.6), (M.8), (M.9), (M.8e) and (M.9e).

(1.1) We first consider the peak

$$a * (b'' * \alpha + m * r) \stackrel{\mathrm{M.S3}}{\Leftarrow} a * (b' * \alpha) \stackrel{\mathrm{M.6}}{\Rightarrow} (ab') * \alpha,$$

where $b' \Rightarrow_b^m b''$. By normalizing the left-hand side with M we obtain

$$(ab'') * \alpha + (am) * r.$$

(1.1.1) Suppose $ab' = ab''$. Since R is an integral domain, we may cancel a on both sides. We obtain $b' = b''$, which contradicts $b' \succ_R b''$.

(1.1.2) Suppose $ab' \succ_R ab''$. Then $\mathcal{S}_{\mathrm{M}}(b * \alpha \approx r)$ contains the rule

$$(ab') * \alpha \Rightarrow (ab'') * \alpha + (am) * r,$$

since $ab' \Rightarrow_b^{am} ab''$. By applying this rule to $(ab') * \alpha$ we obtain convergence.

(1.1.3) Otherwise $ab'' \succ_R ab'$. In this case the rule is oriented the other way:

$$(ab'') * \alpha \Rightarrow (ab') * \alpha + (-am) * r$$

Applying this rule to $(ab'') * \alpha + (am) * r$ and canceling $(am) * r$ against $(-am) * r$ yields $(ab') * \alpha$.

(1.2) The peaks with the remaining rules are all analogous to the following one:

$$b'' * \alpha + m * r + b_0 * \alpha \overset{\mathrm{M.S3}}{\Leftarrow} b' * \alpha + b_0 * \alpha \overset{\mathrm{M.9}}{\Rightarrow} (b' + b_0) * \alpha,$$

where $b' \Rightarrow_b^m b''$. Rewriting the left-hand side with (M.9) yields $(b'' + b_0) * \alpha + m * r$.

(1.2.1) Suppose $b' + b_0 = b'' + b_0$. This cannot be the case, since this would imply $b' = b''$ by cancellation, in contradiction to $b' \succ_R b''$.

(1.2.2) Suppose $b' + b_0 \succ_R b'' + b_0$. Then there exists a rule

$$(b' + b_0) * \alpha \Rightarrow (b'' + b_0) * \alpha + m * r$$

in $\mathcal{S}_{\mathrm{M}}(b * \alpha \approx r)$.

(1.2.3) Otherwise $b'' + b_0 \succ_R b' + b_0$. Then $\mathcal{S}_{\mathrm{M}}(b * \alpha \approx r)$ contains

$$(b'' + b_0) * \alpha \Rightarrow (b' + b_0) * \alpha + (-m) * r,$$

and canceling $m * r$ against $-m * r$ we again obtain convergence.

(2) It remains to show check overlaps of $\mathcal{S}_{\mathrm{M}}(l \approx r)$ into itself. Overlaps of (M.S2) into itself are trivial.

For (M.S3) the only overlap is at the top-level. It is

$$b_1'' * \alpha + m_1 * r \overset{\mathrm{M.S3}}{\Leftarrow} b' * \alpha \overset{\mathrm{M.S3}}{\Rightarrow} b_2'' * \alpha + m_2 * r,$$

where $b_1'' \Leftarrow_b^{m_1} b' \Rightarrow_b^{m_2} b_2''$.

(2.1) If $b_1'' = b_2''$ then $m_1 b = b' - b_1'' = b' - b_2'' = m_2 b$, by canceling b we get $m_1 = m_2$, and we conclude that the peak is trivial.

(2.2) If $b_1'' \succ_R b_2''$ then

$$b_1'' * \alpha + m_1 * r \overset{\mathrm{M.S3}}{\Rightarrow} b_2'' * \alpha + (m_2 - m_1) * r + m_1 * r \Rightarrow_{\mathrm{M}} b_2'' * \alpha + m_2 * r,$$

since $b_1'' = b' - m_1 b = b_2'' + m_2 b - m_1 b = b_2'' + (m_2 - m_1)b$ implies $b_1'' \Rightarrow_b^{m_2 - m_1} b_2''$.

(2.3) The remaining case of $b_2'' \succ_R b_1''$ is analogous. □

Note that it is crucial for this proof that R has no zero divisors. If this were not the case, say $b_1 b_2 = 0$ for $b_1, b_2 \in R \setminus \{0\}$, then for a rule $b_2 * \alpha \Rightarrow r$ the critical pair

$$b_1 * r \overset{\text{M,S3}}{\Leftarrow} b_1 * (b_2 * \alpha) \overset{\text{M,6}}{\Rightarrow} 0 * \alpha$$

would not converge. To make it convergent the symmetrization would need a rule that does not contain α, which would violate the condition of left-minimality. Hence such rings are beyond the scope of this approach.

The symmetrization function defined above allows any rewriting step that decreases the coefficient in the ordering, hence in general there exist many rules with the same left-hand side. This simplifies the confluence proof above, because any peak can be closed with a single $\mathcal{S}_{\mathrm{M}}$-step. But it would also lead to superfluous Superposition inferences if used without the restriction that only Superposition inferences with minimal right-hand side are needed.

7.3 Critical extension peaks and transitivity

Theorem 7.9 *There are no extension peaks with respect to* M.

Proof: For fields there are no extended rules in the symmetrization, so $R = \mathbb{Z}$ is the only interesting case. Now let $b_i * \alpha_i \Rightarrow r_i$ be a rewrite rule in $\mathrm{Norm}_{\mathrm{M}}$ and $S_i = \mathcal{S}_{\mathrm{M}}(b_i * \alpha_i \Rightarrow r_i)$ for $i = 1, 2$. Clearly S_1 and S_2 can only overlap if $\alpha =_{\mathrm{AC}} \alpha_1 =_{\mathrm{AC}} \alpha_2$. Assume without loss of generality that $b_1 \geq b_2$. Then $b_1 * \alpha_1$ can be reduced by the rule $b_1 * \alpha_2 \Rightarrow b_1 - b_2 * \alpha_2 + r_2$ in S_2. Hence there is no extension peak. □

There are no critical terms with respect to M, hence $\mathrm{cc}_{\mathrm{M}}(t) = \mathrm{cc}_{\mathrm{M}}(C) = \top$ for any ground term t and ground clause C.

Corollary 7.10 *Transitivity holds in I_N for all sets of ground clauses N.*

7.4 Simplification

For modules we will use the following Isolation rule. Its conclusion is almost in M-normal form, the only conditions missing are that s' is atomic and irreducible with respect to M. Note that the conclusion is uniquely determined up to the sign of the unit u and the order of p and q in the difference. Which variant occurs depends on the orientation of the premise. We do not break the symmetry between the two sides of the premise by an ordering restriction, because this simplifies the proof that Isolation is a simplification. Here our goal is to prove that Isolation is a simplification even without these restrictions, later we will add them in order to decrease the number of inferences.

M-*Isolation* $$\frac{[\neg](b' * s' + p \approx b'' * s'' + q)}{[\neg](b * s' \approx u * (q - p))}$$

if (i) $s' =_{\mathrm{AC}} s''$, (ii) $b = u(b' - b'')$ where (a) $R = \mathbb{Z}$, $b \geq 1$ and $u = \mathrm{sign}(b' - b'')$ for $b' \neq b''$, or (b) R is a field, $b = 1$ and $u = (b' - b'')^{-1}$ for $b' \neq b''$, or (c) $b = 0$ and $u = 1$ for $b' = b''$, (iii) $b' * s' + p \succ b * s'$ or $b'' * s'' + q \succ b * s'$, and (iv) $s' \succ p$ and $s'' \succ q$.

Remember that by our notational convention a coefficient of 1 is tacitly assumed wherever a coefficient is missing.

Lemma 7.11 M-*Isolation is a simplification rule.*

Proof: Since Isolation preserves M-equivalence and transitivity holds universally, and since the conclusion is smaller than the premise, Isolation is a simplification. □

Like in the preceding chapters the simplification function will consists of a subset of these simplifications which obeys additional ordering restrictions. Analogously to the case of commutative rings we restrict $\mathrm{Simp}_{\mathrm{M}}$-Rewriting on a literal L to redexes either inside or above a maximal summand. We restrict M-Isolation to irreducible atomic terms. Finally, as always the premise of the simplifications has to be oriented such that its left-hand side is not smaller than the right hand side.

$\mathrm{Simp}_{\mathrm{M}}$-*Sum Contraction* $$\frac{[\neg](b' * s' + b'' * s'' + p \approx q)}{[\neg](b * s' + p \approx q)}$$

if (i) $s' =_{\mathrm{AC}} s''$, (ii) $b = b' + b''$, (iii) $b' * s'$ is a maximal summand in $b' * s' + b'' * s'' + p$, and (iv) $b' * s' + b'' * s'' + p \succeq q$.

$\mathrm{Simp}_{\mathrm{M}}$-*Summand Rewriting* $$\frac{[\neg](u[l] + p \approx q)}{[\neg](u[r] + p \approx q)}$$

if (i) $l \Rightarrow r$ is a rule in M, (ii) $u[l]$ is a maximal summand in $u[l]+p$, and (iii) $u[l]+p \succeq q$.

$\mathrm{Simp}_{\mathrm{M}}$-*Isolation* $$\frac{[\neg](b' * \alpha' + p \approx b'' * \alpha'' + q)}{[\neg](b * \alpha' \approx u * (q - p))}$$

if (i) $\alpha' =_{\mathrm{AC}} \alpha''$, (ii) $b = u(b' - b'')$ where (a) $R = \mathbb{Z}$, $b \geq 1$ and $u = \mathrm{sign}(b' - b'')$ for $b' \neq b''$, or (b) R is a field, $b = 1$ and $u = (b' - b'')^{-1}$ for $b' \neq b''$, or (c) $b = 0$ and $u = 1$ for $b' = b''$, (iii) either $b' \succ_R b$ or $p \neq 0$, (iv) α' is M-atomic and irreducible with respect to M, (v) $\alpha' \succ p$ and $\alpha'' \succ q$. and (vi) $b' * \alpha' + p \succeq b'' * \alpha'' + q$.

Note that the conclusion of a $\mathrm{Simp}_{\mathrm{M}}$-Isolation is in M-normal form. We let $\mathrm{Simp}_{\mathrm{M}}(L)$ consist of all literals L' such that there exists a simplification by $\mathrm{Simp}_{\mathrm{M}}$-Rewriting or $\mathrm{Simp}_{\mathrm{M}}$-Isolation with premise L and conclusion L'.

Lemma 7.12 $\mathrm{Simp}_{\mathrm{M}}$ *is an admissible simplification function for* $\mathrm{Norm}_{\mathrm{M}}$.

Proof: $\mathrm{Simp}_{\mathrm{M}}$-Rewriting and $\mathrm{Simp}_{\mathrm{M}}$-Isolation are restrictions of the rules proven to be simplifications in Lemma 7.11, hence they are simplification rules.

It remains to show that $\mathrm{Simp}_{\mathrm{M}}$ is admissible for $\mathrm{Norm}_{\mathrm{M}}$. That is, we have to show that any literal is either in $\mathrm{Norm}_{\mathrm{M}}$ or be simplified by $\mathrm{Simp}_{\mathrm{M}}$. Let $L = [\neg](p \approx q)$ be some ground literal. We may assume without loss of generality that $p \succeq q$.

(1) If $p = q = 0$ then L is in M-normal form.

(2) Otherwise p contains a maximal summand $b' * s'$. Then $p = b' * s' + p'$.

(2.1) Suppose p' contains another summand $b'' * s''$ with $s'' =_{\mathrm{AC}} s'$. That is, $p = b' * s' + b'' * s'' + p''$. Then $\mathrm{Simp}_{\mathrm{M}}$-Sum Contraction applies.

(2.2) Suppose $b' * s'$ is reducible by M. Then $\mathrm{Simp}_{\mathrm{M}}$-Summand Rewriting applies.

(2.3) Otherwise s' is atomic, and $b' * s' = b' * \alpha'$ is the only summand in p containing α'. Let $q = b'' * \alpha'' + q'$ where $\alpha' =_{\mathrm{AC}} \alpha''$ and $b'' = 0$ denotes absence of α''. There can be at most one summand containing α', otherwise q would be greater than p.

(2.3.1) If $b' > 0$, $b'' = 0$ and $p = 0$ then L is in M-normal form.

(2.3.2) Otherwise condition (ii) determines b such that b is minimal and either $b' \succ_R b$ or $p \neq 0$, and Isolation applies. □

7.5 The inference system

We obtain the ground inference system below. Note that we have omitted the ordering restrictions that select the maximal literal; these are the same as in the general case.

M-*Sum Contraction*

$$\frac{[\neg](b * s + b' * s' + p \approx q) \vee C}{[\neg]((b + b') * s + p \approx q) \vee C}$$

if (i) $s =_{\mathrm{AC}} s'$, (ii) $b * s$ is a maximal summand in $b * s + b' * s' + p$, and (iii) $b * s + b' * s' + p \succeq q$.

M-*Summand Rewriting*

$$\frac{[\neg](u[l] + p \approx q) \vee C}{[\neg](u[r] + p \approx q) \vee C}$$

if (i) $l \Rightarrow r$ is a rule in M, (ii) $u[l]$ is a maximal summand in $u[l]+p$, and (iii) $u[l]+p \succeq q$.

M-*Isolation*

$$\frac{[\neg](b' * \alpha' + p \approx b'' * \alpha'' + q) \vee C}{[\neg](b * \alpha' \approx u * (q - p)) \vee C}$$

if (i) $\alpha' =_{\mathrm{AC}} \alpha''$, (ii) $b = u(b' - b'')$ where (a) $R = \mathbb{Z}$, $b \geq 1$ and $u = \mathrm{sign}(b' - b'')$ for $b' \neq b''$, or (b) R is a field, $b = 1$ and $u = (b' - b'')^{-1}$ for $b' \neq b''$, or (c) $b = 0$ and $u = 1$ for $b' = b''$, (iii) either $b' \succ_R b$ or $p \neq 0$, (iv) α' is M-atomic and irreducible with respect to M, (v) $\alpha' \succ p$ and $\alpha'' \succ q$. and (vi) $b' * \alpha' + p \succeq b'' * \alpha' + q$.

M-*Superposition*

$$\frac{b * \alpha \approx r \vee D \qquad [\neg](p[b' * \alpha'] \approx q) \vee C}{[\neg](p[b'' * \alpha + m * r] \approx q) \vee C \vee D}$$

if (i) $\alpha =_{\mathrm{AC}} \alpha'$, (ii) $b' \Rightarrow_b^m b''$, (iii) b'' is the minimal ring element in $b'+Rb$, (iv) $b*\alpha \approx r$ is in M-normal form, and (v) $[\neg](p[b' * \alpha'] \approx q)$ is in M-normal form.

Condition (iii) becomes integer division in the case of $R = \mathbb{Z}$, where b is the smallest positive remainder that can be obtained. For fields b'' is always zero.

Reflexivity Resolution

$$\frac{p \not\approx q \vee C}{C}$$

if (i) $p =_{\mathrm{AC}} q$, (ii) $p \not\approx q$ is in M-normal form.

M-*Equality Factoring*

$$\frac{s \approx r \vee t \approx r' \vee C}{r \not\approx r' \vee t \approx r' \vee C}$$

if (i) $s =_{\mathrm{AC}} t$, (ii) $s \approx r$ and $t \approx r'$ are in M-normal form, and (iii) $r \succeq r'$.

We let $\mathsf{Sup}_{\mathrm{M}}$ be the set of these inferences.

Theorem 7.13 $\mathsf{Sup}_{\mathrm{M}}$ *is refutationally complete for* M_1.

Proof: Strictly analogous to the proof for abelian groups, this follows from Theorem 4.20 in combination with Propositions 7.7 and 7.5, and Lemmas 7.8 and 7.12. □

7.6 Improving superpositions at the root position

We will now exploit that the integers are an Euclidean ring with respect to $\succ_R$. That is, the remainder of an integer division is smaller then the divisor in $\succ_R$.

Example 7.14 *Suppose we have two equations* $10 * c \approx r_1$ *and* $6 * c \approx r_2$ *where* $c \succ r_1$ *and* $c \succ r_2$. *We get the following sequence of superpositions, where the first column gives the results of the superpositions and the second the equations in* M*-normal form (except for the last):*

$$\begin{array}{rr} & 10 * c \approx r_1 \\ & 6 * c \approx r_2 \\ 4 * c + r_2 \approx r_1 & 4 * c \approx r_1 + (-1) * r_2 \\ 2 * c + r_1 + (-1) * r_2 \approx r_2 & 2 * c \approx (-1) * r_1 + 2 * r_2 \\ (-2) * r_1 + 4 * r_2 \approx r_1 + (-1) * r_2 & 0 \approx 3 * r_1 + (-5) * r_2 \end{array}$$

We notice that this sequence computes the greatest common divisor for the coefficients on c, using Euclid's algorithm.

More generally, consider two positive ground literals $b_1 * \alpha \approx r_1$ and $b_2 * \alpha \approx r_2$ where $b_1 > b_2 \geq 1$. By M-Superposition and contraction of summands containing r_1 and r_2 we get the following general sequence:

$$\begin{aligned} b_1 * \alpha &\approx r_1 \\ b_2 * \alpha &\approx r_2 \\ b_3 * \alpha &\approx m'_3 * r_1 + m''_3 * r_2 \\ &\vdots \\ b_n * \alpha &\approx m'_n * r_1 + m''_n * r_2 \\ 0 &\approx m'_{n+1} * r_1 + m''_{n+1} * r_2 \end{aligned}$$

Equation number i is obtained by superposing with equation $i-1$ into $i-2$ for $3 \leq i \leq n+1$. Then m_i and b_i are the quotient and remainder, respectively, of the integer division of b_{i-2} by b_{i-1}. That is, $b_{i-2} = m_i b_{i-1} + b_i$ and $b_{i-1} > b_i$. We let $m'_1 = 1$, $m''_1 = 0$, $m'_2 = 0$, $m''_2 = 1$, and $m'_i = m'_{i-2} - m_i m'_{i-1}$ and $m''_i = m''_{i-2} - m_i m''_{i-1}$ for $3 \leq i \leq n+1$, which preserves the invariant $b_i = b_1 m'_i + b_2 m''_i$. Finally, b_n is the greatest common divisor of b_1 and b_2, and $b_n = m'_n b_1 + m''_n b_2$. In the presence of the last two equations the other equations become redundant. Their left-hand sides can be reduced by equation n such that it no longer contains the maximal term α, and the resulting equation is a consequence of equation $n+1$. Note that those two equations are smaller than the equations to be shown redundant.

The computation on the coefficients is the extended Euclidean algorithm, which not only computes the greatest common divisor $b = b_n$, but also the *Bézout coefficients* $m_1 = m'_n$ and $m_2 = m''_n$ such that $b = m_1 b_1 + m_2 b_2$.

Lemma 7.15 *(von zur Gathen and Gerhard 1999, Lemma 3.12)*

$$|m'_i| \leq \frac{b_2}{b_{i-1}} \text{ and } |m''_i| \leq \frac{b_1}{b_{i-1}} \text{ for } 2 \leq i \leq n+1.$$

For $i = n$ this implies $|m_1| \leq \frac{b_2}{2b}$ and $|m_2| \leq \frac{b_1}{2b}$, since b_{n-1} is at least $2b$. Moreover, with this restriction m_1 and m_2 are determined uniquely. For $i = n+1$ we get $|m'_{n+1}| \leq \frac{b_2}{b}$ $|m''_{n+1}| \leq \frac{b_1}{b}$. On the other hand $0 = b_1 m'_{n+1} + b_2 m''_{n+1}$ implies that $|b_1 m'_{n+1}| = |b_2 m''_{n+1}|$ is a common multiple of b_1 and b_2. The least common multiple is $b_1 b_2/b$, hence $|b_1 m'_{n+1}| \geq b_1 b_2/b$ and $|b_2 m''_{n+1}| \geq b_1 b_2/b$, which implies $|m'_{n+1}| \geq b_2/b$ and $|m''_{n+1}| \geq b_1/b$. We conclude $|m'_{n+1}| = b_2/b$ and $|m''_{n+1}| = b_1/b$. Hence we can write equation $n+1$ equivalently as

$$(b_2/b) * r_1 \approx (b_1/b) * r_2.$$

We can now give two inferences which replace the sequence of superpositions. In the inference system they replace the corresponding M-Superposition inferences in the top-level sum. Note that the standard inference is kept for $b_1 = b_2$.

M-*GCD Superposition A* $$\frac{b_1 * \alpha_1 \approx r_1 \vee C \qquad b_2 * \alpha_2 \approx r_2 \vee D}{b * \alpha_1 \approx m_1 * r_1 + m_2 * r_2 \vee C \vee D}$$

if (i) $\alpha_1 =_{\mathrm{AC}} \alpha_2$, (ii) $b = m_1 b_1 + m_2 b_2$, (iii) $b = \gcd(b_1, b_2)$, (iv) $-\frac{b_2}{2b} \leq m_1 \leq \frac{b_2}{2b}$, (v) $-\frac{b_1}{2b} \leq m_2 \leq \frac{b_1}{2b}$, (vi) $b_1 > b_2 \geq 1$, and (vii) $b_i * \alpha_i \approx r_i$ is in M-normal form for $i = 1, 2$.

M-*GCD Superposition B* $$\frac{b_1 * \alpha_1 \approx r_1 \vee C \qquad b_2 * \alpha_2 \approx r_2 \vee D}{(b_2/b) * r_1 \approx (b_1/b) * r_2 \vee C \vee D}$$

if (i) $\alpha_1 =_{\mathrm{AC}} \alpha_2$, (ii) $b = m_1 b_1 + m_2 b_2$, (iii) $b = \gcd(b_1, b_2)$, (iv) $-\frac{b_2}{2b} \leq m_1 \leq \frac{b_2}{2b}$, (v) $-\frac{b_1}{2b} \leq m_2 \leq \frac{b_1}{2b}$, (vi) $b_1 > b_2 \geq 1$, and (vii) $b_i * \alpha_i \approx r_i$ is in M-normal form for $i = 1, 2$.

Here the expensive computation of the GCD on the term level is replaced by a computation of the extended GCD in the integers. Analogous inferences were used by Kandri-Rody and Kapur (Kandri-Rody and Kapur 1988) for the computation of Gröbner bases over a Euclidean domain and by Wang (Wang 1993) for integer module reasoning. Note that for fields this is not useful, as the remainder is already zero after a single Superposition step.

8

Commutative Algebras

Let $\langle R, +_R, \cdot_R, -_R, 0_R, 1_R \rangle$ be a commutative ring. A *(commutative) R-algebra* is a commutative ring

$$\langle \mathrm{CA}, +_{\mathrm{CA}}, \cdot_{\mathrm{CA}}, -_{\mathrm{CA}}, 0_{\mathrm{CA}}, 1_{\mathrm{CA}} \rangle$$

together with a ring homomorphism $\langle \_ \rangle : R \to \mathrm{CA}$. In this chapter we will develop a superposition calculus for theorem proving with respect to all R-algebras over a fixed R. For R we allow the same rings as in the case of modules, namely the ring of integers and fields. Like for modules we drop the subscripts indicating whether the ring operation belongs to R or CA, as this will be clear from the context. We assume that the CA-operations $-$, 0 and 1 are eliminated in a preprocessing step using the term rewriting system

$$\begin{aligned} -x &\Rightarrow \langle -1 \rangle x \\ 0 &\Rightarrow \langle 0 \rangle \qquad \text{and} \\ 1 &\Rightarrow \langle 1 \rangle. \end{aligned}$$

This terminates, since in each step one of the symbols is removed, and it is confluent, since there are no critical pairs. This leads to the signature

$$\begin{aligned} +,\ \cdot &: \mathsf{CA} \times \mathsf{CA} \longrightarrow \mathsf{CA} \\ a &: \ \longrightarrow \mathsf{R} \qquad \text{for all } a \in R \\ +,\ \cdot &: \mathsf{R} \times \mathsf{R} \longrightarrow \mathsf{R} \\ - &: \mathsf{R} \longrightarrow \mathsf{R} \\ \langle \_ \rangle &: \mathsf{R} \longrightarrow \mathsf{CA}. \end{aligned}$$

Free function symbols operate again on CA:

$$f : \mathsf{CA}^{\alpha(f)} \longrightarrow \mathsf{CA}.$$

Terms of sort R are interpreted by elements of R, while terms of sort CA are interpreted by elements of the algebra CA. We impose the same restrictions on the use of R-terms as for modules. We use the same notational conventions as for commutative rings, except for the meta notation $n\phi$. Instead we often omit $\cdot$ and use juxtaposition to indicate multiplication. We will use $s - \langle b \rangle t$ as an abbreviation for $s + \langle -b \rangle t$, and we will identify t and $\langle 1 \rangle t$ where appropriate. Let $\widehat{\mathrm{CA}}$ be the following constrained term rewriting system modulo $\mathrm{AC} = \mathrm{AC}(+) \cup \mathrm{AC}(\cdot)$. This is essentially the term rewriting system of Bachmair and Ganzinger (1994b).

$$x + \langle 0 \rangle \Rightarrow x \tag{CA.1}$$
$$\langle 0 \rangle x \Rightarrow \langle 0 \rangle \tag{CA.2}$$
$$\langle 1 \rangle x \Rightarrow x \tag{CA.3}$$
$$x(y+z) \Rightarrow xy + xz \tag{CA.4}$$
$$\langle v_1 \rangle + \langle v_2 \rangle \Rightarrow \langle v \rangle \;\; [v = v_1 + v_2] \tag{CA.5}$$
$$\langle v_1 \rangle \langle v_2 \rangle \Rightarrow \langle v \rangle \;\; [v = v_1 \cdot v_2] \tag{CA.6}$$
$$x + x \Rightarrow \langle v \rangle x \;\; [v = 1 + 1] \tag{CA.7}$$
$$\langle v_1 \rangle x + x \Rightarrow \langle v \rangle x \;\; [v = v_1 + 1] \tag{CA.8}$$
$$\langle v_1 \rangle x + \langle v_2 \rangle x \Rightarrow \langle v \rangle x \;\; [v = v_1 + v_2] \tag{CA.9}$$
$$y + \langle v_1 \rangle + \langle v_2 \rangle \Rightarrow y + \langle v \rangle \;\; [v = v_1 + v_2] \tag{CA.5e}$$
$$y\langle v_1 \rangle \langle v_2 \rangle \Rightarrow y\langle v \rangle \;\; [v = v_1 \cdot v_2] \tag{CA.6e}$$
$$y + x + x \Rightarrow y + \langle v \rangle x \;\; [v = 1 + 1] \tag{CA.7e}$$
$$y + \langle v_1 \rangle x + x \Rightarrow y + \langle v \rangle x \;\; [v = v_1 + 1] \tag{CA.8e}$$
$$y + \langle v_1 \rangle x + \langle v_2 \rangle x \Rightarrow y + \langle v \rangle x \;\; [v = v_1 + v_2] \tag{CA.9e}$$

We let $\mathrm{CA} = \mathrm{AC}\backslash gnd(\widehat{\mathrm{CA}})$, which again consists of those ground instances of $\mathrm{AC}\backslash\widehat{\mathrm{CA}}$ that satisfy the constraints.

Lemma 8.1 CA *is locally ground confluent and locally ground coherent modulo* AC.

Proof: We have shown this by the same approach as for modules. □

8.1 Termination

We slightly modify the ordering for modules to obtain an ordering $\succeq_{\mathrm{CA}}$ for commutative algebras. For the first component let $F_I = F_{\mathrm{CA}} = \{+, \cdot, \langle \_ \rangle\} \cup R$. We assume a well-ordering $\succeq_p$ on $F \setminus F_I$, and let $\succeq_p$ also denote its TPO-admissible extension to F. To apply Section 3.3 we assume as given a quasi-ordering $\succeq_c$ on F_C and extend it to a quasi-ordering $\succeq_t$ on terms over $F_I \cup F_C$. Let $\widehat{\mathrm{D}}$ be the term rewriting system consisting of the distributivity rule (CA.4), and let $\mathrm{D} = \mathrm{AC}\backslash gnd(\widehat{\mathrm{D}})$. Then D is convergent modulo AC. Let t be any ground term t over $F_I \cup F_C$ that is in D-normal form. Then

$$t = \langle a_{11} \rangle \cdots \langle a_{1k_1} \rangle \phi_1 + \cdots + \langle a_{n1} \rangle \cdots \langle a_{nk_n} \rangle \phi_n$$

where $n \geq 1$, $k_i \geq 0$, $\phi_i = c_{i1} \cdots c_{il_i}$ and $l_i \geq 0$ for $i = 1, \ldots, n$. We use a similar complexity measure as in the case of modules, with two modifications. The constants in the first component are replaced by products, and the tuples in the multiset in the third components become multisets. Here we identify products in the same $\mathrm{AC} \cup \sim_c$-equivalence class, or equivalently the multisets M_ϕ of coefficients in ϕ in the same $\sim_{c,mul}$-equivalence class. Hence we extend the representation function rep from F_C to products over F_C so that it maps products in an $\mathrm{AC} \cup \sim_c$-equivalence class to a multiset over F_C representing the product. That is, if $\phi =_{\mathrm{AC}\cup\sim_c} \psi$ then $\mathrm{rep}(\phi) = \mathrm{rep}(\psi)$. We let

$$\begin{aligned} \mathrm{occ}(t, \phi_i) &= \{ j \mid \phi_j =_{\mathrm{AC}\cup\sim_c} \phi_i \} \\ \#(t, \phi_i) &= |\mathrm{occ}(t, \phi_i)| \\ \mathrm{cs}(t, \phi_i) &= \{\{a_{11}, \ldots, a_{1k_1}\}, \ldots, \{a_{n1}, \ldots, a_{nk_n}\}\} \end{aligned}$$

That is, $\mathrm{occ}(t, \phi_i)$ is the set of occurrences of products in the same $\mathrm{AC} \cup \sim_c$-equivalence class as ϕ_i, $\#(t, \phi_i)$ is the number of these occurrences, and $\mathrm{cs}(t, \phi_i)$ is the multiset of the multisets of coefficients associated with these occurrences. To each $\mathrm{AC} \cup \sim_c$-equivalence class of products we associate a tuple $\langle \mathrm{rep}(\phi_i), \#(t, \phi_i), \mathrm{cs}(t, \phi_i)\rangle$. Finally, we let $\kappa(t)$ be the set containing the tuples for the products occurring in t. We let the ordering $\succeq_\kappa$ be the multiset extension of the lexicographic combination of $\succeq_c$, $>$ and the multiset extension of the size-multiset extension of $\succeq_R$. Then we define the ordering $\succeq_t$ on terms over $F_I \cup F_C$ by $s \succeq_t t$ if and only if $\kappa(\mathrm{D}(s)) \succeq_\kappa \kappa(\mathrm{D}(t))$ where s and t are terms over $F_I \cup F_C$. We get the TPO-status $\succeq_1^{st}$ as the status derived from $\succeq_t$, and let $\succeq_1(\succeq_p) = \succeq_{tpo}(\succeq_p, \succeq_1^{st})$.

Lemma 8.2 *Let $\succeq_p$ be a well-ordering on $F \setminus F_{\mathrm{CA}}$. Then $\succeq_1(\succeq_p)$ is a total* $\mathrm{AC} \cup \mathrm{D}$-*compatible and* $\mathrm{AC} \cup \mathrm{D}$-*antisymmetric simplification quasi-ordering on ground terms that contains* $\mathrm{CA} \setminus \mathrm{D}$.

Proof: (*Simplification quasi-ordering*) We begin by showing that $\succeq_1^{st}$ is a TPO-status by Lemma 3.9.

(*Strictly compatible with contexts*) Let $f_{u[]}$ be the function that maps any complexity $\kappa(\mathrm{D}(t))$ to the complexity $\kappa(\mathrm{D}(u[t]))$, and let

$$\mathrm{D}(t) = \langle a_{11}\rangle \cdots \langle a_{1k_1}\rangle \phi_1 + \cdots + \langle a_{n1}\rangle \cdots \langle a_{nk_n}\rangle \phi_n$$

where $n \geq 1$, $k_i \geq 0$ $\phi_i = c_{i1} \cdots c_{il_i}$ and $l_i \geq 0$ for $i = 1, \ldots, n$. By the same argument as for modules $f_{u[]}$ is well-defined. We show that $f_{u[]}$ is strictly monotonic for any context $u[]$ by considering contexts of depth one.

(1) Consider $u = s + []$.

(1.1) Suppose $s = \langle b_1\rangle \cdots \langle b_l\rangle \psi$.

(1.1.1) Suppose no ϕ_i is $\mathrm{AC} \cup \sim_c$-equivalent to ψ. Then

$$f_{u[]}(\kappa(t)) = \kappa(t) \cup \{\langle \mathrm{rep}(\psi), 1, \{\{b_1, \ldots, b_l\}\}\rangle\}$$

and $f_{u[]}$ is strictly monotonic.

(1.1.2) Otherwise $\psi =_{\mathrm{AC}\cup\sim_c} c_i$ for some $i = 1, \ldots, k$. Suppose without loss of generality $\psi =_{\mathrm{AC}\cup\sim_c} \phi_1$. Then

$$\begin{aligned} f_{u[]}(\kappa(t)) = {} & \{\langle \mathrm{rep}(\phi_1), \#(t, \phi_1) + 1, \mathrm{cs}(t, \phi_1) \cup \{\{b_1, \ldots, b_l\}\}\rangle\} \\ & \cup \{\langle \mathrm{rep}(\phi_i), \#(t, \phi_i), \mathrm{cs}(t, \phi_i)\rangle \mid \phi_i \neq_{\mathrm{AC}\cup\sim_c} \phi_1\}. \end{aligned}$$

Consider the function that maps a tuple $\langle \mathrm{rep}(\phi), n, M\rangle$ to

$$\langle \mathrm{rep}(\phi), n + 1, M \cup \{\{b_1, \ldots, b_l\}\}\rangle$$

if $\phi =_{\mathrm{AC}\cup\sim_c} \psi$ and to $\langle \mathrm{rep}(\phi), n, M\rangle$ otherwise. This function is strictly monotonic according to Proposition 2.8(7). Hence its multiset extension $f_{u[]}$ is strictly monotonic.

(1.2) Otherwise s is a proper sum, and $f_{u[]}$ can again be obtained as a finite composition of the strictly monotonic functions of case (1.1). Hence $f_{u[]}$ is strictly monotonic.

(2) Consider $u = s \cdot []$.

(2.1) Suppose $s = \langle b_1\rangle \cdots \langle b_l\rangle \psi$. Then $f_{u[]}$ maps any tuple $\langle \mathrm{rep}(\phi), n, \{M_1, \ldots, M_k\}\rangle$ in $\kappa(t)$ to

$$\langle \mathrm{rep}(\phi) \cup \mathrm{rep}(\psi), n, \{M_1 \cup N, \ldots, M_k \cup N\}\rangle$$

where $N = \{b_1, \ldots, b_l\}$, which clearly is strictly monotonic.

(2.2) Otherwise s is a proper sum, and $f_{u[]}$ is a finite union of the functions of case (2.1), which again is strictly monotonic.

(*Subterm property*) It suffices to consider contexts of depth one, then the subterm property follows by structural induction. We have to show (1) $s + t \succeq_t t$ and (2) $s \cdot t \succeq_t t$.

In case (1) $\kappa(\mathrm{D}(s+t))$ contains at least one additional tuple or a tuple whose second component increases when compared to $\kappa(\mathrm{D}(t))$. In case (2) the sizes of the multisets in the third component strictly increase for each tuple in $\kappa(\mathrm{D}(t))$.

(*Decreases infinite derivations*) Suppose there is some infinite descending chain

$$\kappa(\mathrm{D}(t_1)) \succ_\kappa \kappa(\mathrm{D}(t_2)) \succ_\kappa \ldots .$$

Then there exists an infinite descending chain of tuples

$$\langle \mathrm{rep}(\phi_1), n_1, M_1 \rangle \succ \langle \mathrm{rep}(\phi_2), n_2, M_2 \rangle \succ \ldots .$$

Since $>$ on natural numbers and $\succ_R$ and its extensions are well-founded, there exists an infinite descending chain

$$\mathrm{rep}(\phi_{j_1}) \succ_{mul}(\succeq_c) \mathrm{rep}(\phi_{j_2}) \succ_{mul}(\succeq_c) \ldots$$

in the first component. By Proposition 2.6(2) there exists an infinite descending chain

$$c_1 \succ_c c_2 \succ_c \ldots$$

where c_i occurs in some ϕ_{j_i} and hence in t_{j_i} for $1 \le j_1 < j_2 < \ldots$.

(*Constant dominance condition*) The constants dominate the first component of each tuple, hence they dominate the ordering.

(AC∪D-*compatible*) We have to show that $\succeq_t$ is AC∪D-compatible. Suppose $s \Leftrightarrow_{\mathrm{AC}\cup\mathrm{D}} s' \succeq_t t' \Leftrightarrow_{\mathrm{AC}\cup\mathrm{D}} t$. Then

$$\kappa(\mathrm{D}(s)) = \kappa(\mathrm{D}(s')) \succeq_\kappa \kappa(\mathrm{D}(t)) = \kappa(\mathrm{D}(t'))$$

implies $s \succeq_t t$.

(AC∪D-*antisymmetric*) We have to show that the quasi-ordering $\succeq_t(\succeq_c)$ is AC∪D∪$\sim_c$-antisymmetric. Suppose $s \sim_t(\succeq_c)\, t$ and

$$\kappa(\mathrm{D}(s)) = \{\langle \mathrm{rep}(\phi_1), m_1, M_1\rangle, \ldots, \langle \mathrm{rep}(\phi_k), m_k, M_k\rangle\} \text{ and}$$
$$\kappa(\mathrm{D}(t)) = \{\langle \mathrm{rep}(\psi_1), n_1, N_1\rangle, \ldots, \langle \mathrm{rep}(\psi_l), n_l, N_l\rangle\}.$$

Then $\kappa(\mathrm{D}(s)) \sim_\kappa \kappa(\mathrm{D}(t))$, $k = l$, and there exists a permutation π such that $\mathrm{rep}(\phi_i) = \mathrm{rep}(\psi_{\pi(i)})$, $m_i = n_{\pi(i)}$ and $M_i = N_{\pi(i)}$ for $i = 1, \ldots, k$. We can now show that $\mathrm{D}(s) \overset{*}{\Leftrightarrow}_{\mathrm{AC}\cup\sim_c} \mathrm{D}(t)$. We associate to each product ϕ the product $\phi' = c_1 \cdots c_n$ derived from the representative $\mathrm{rep}(\phi) = \{c_1, \ldots, c_n\}$. Let s' and t' be the terms resulting from replacing any product ϕ in $\mathrm{D}(s)$ and $\mathrm{D}(t)$ by the corresponding product ϕ'. They can be written as $s' = s_1 + \cdots + s_k$ and $t' = t_1 + \cdots + t_l$ where s_i consists of the terms containing ϕ'_i and t_i of the terms containing ψ'_i. From $M_i = N_{\pi(i)}$ we infer that ϕ'_i and $\psi'_{\pi(i)}$ are associated with the same multisets of coefficients from R, and that hence $s_i =_{\mathrm{AC}} t_{\pi(i)}$ and in turn $s' =_{\mathrm{AC}} t'$. We conclude that $s =_{\mathrm{AC}\cup\mathrm{D}\cup\sim_c} t$.

(*Total*) Since $\succeq_c$, $\geq$ and $\succeq_R$ are total, and since multiset and lexicographic extension as well as lexicographic product preserve totality, $\succeq_t(\succeq_c)$ is total. Hence $\succeq_1$ is total by Lemma 3.10.

(*Orients rules from left to right*) Rules (CA.1)–(CA.3) are oriented from left to right by the subterm property. The distributivity rule (CA.4) is explicitly excluded here. For (CA.5) and (CA.6) we have

$$\begin{aligned}\kappa(\mathrm{D}(\langle a_1\rangle + \langle a_2\rangle)) &= \{\langle\emptyset, 2, \{\{a_1\}, \{a_2\}\}\rangle\}\\ \kappa(\mathrm{D}(\langle a_1\rangle\langle a_2\rangle)) &= \{\langle\emptyset, 1, \{\{a_1, a_2\}\}\rangle\}\\ \kappa(\mathrm{D}(\langle a\rangle)) &= \{\langle\emptyset, 1, \{\{a\}\}\rangle\},\end{aligned}$$

hence $\langle a_1\rangle + \langle a_2\rangle \succ_1 \langle a\rangle$ and $\langle a_1\rangle\langle a_2\rangle \succ_1 \langle a\rangle$ for any a_1, a_2, a in R.

From (CA.7)–(CA.9) we consider only (CA.7), the others are analogous. We consider some ground instance (CA.7)σ, let $t = P_{F_I}(x\sigma)$ be the term resulting from replacing atomic subterms by constants, and let $b = v\sigma$. Then we have to show that

$$\kappa(\mathrm{D}(t+t)) \succ_\kappa \kappa(\mathrm{D}(\langle b\rangle t)).$$

Suppose

$$\kappa(\mathrm{D}(t)) = \{\langle\mathrm{rep}(\phi_1), n_1, M_1\rangle, \ldots, \langle\mathrm{rep}(\phi_k), n_k, M_k\rangle\}.$$

Then

$$\kappa(\mathrm{D}(t+t)) = \{\langle\mathrm{rep}(\phi_1), 2n_1, 2M_1\rangle, \ldots, \langle\mathrm{rep}(\phi_k), 2n_k, 2M_k\rangle\}$$

which is greater than

$$\kappa(\mathrm{D}(\langle b\rangle t)) = \{\langle\mathrm{rep}(\phi_1), n_1, M_1'\rangle, \ldots, \langle\mathrm{rep}(\phi_k), n_k, M_k'\rangle\}.$$

(CA.5e)–(CA.9e) follow from (CA.5)–(CA.9) by compatibility with contexts. □

We let $\succeq^p_{\mathrm{CA}}$ be the quasi-ordering induced by the following polynomial interpretation, which is again derived from the one for commutative rings:

$$\begin{aligned}p_a^{\mathrm{CA}} &= 2 \qquad \text{for } a \in R\\ p_{\langle\_\rangle}^{\mathrm{CA}}(v) &= v+1\\ p_+^{\mathrm{CA}}(x,y) &= x+y+5\\ p^{\mathrm{CA}}(x,y) &= xy\\ p_f^{\mathrm{CA}}(x_1,\ldots,x_n) &= x_1+\cdots+x_n+2 \qquad \text{for } f \text{ free.}\end{aligned}$$

Lemma 8.3 *$\succeq^p_{\mathrm{CA}}$ is a total* AC*-compatible simplification quasi-ordering on ground terms that orients* D *from left to right.*

We let $\succeq_{\mathrm{CA}}$ be the lexicographic combination of $\succeq_1$, $\succeq^p_{\mathrm{CA}}$ and $\succeq_{acrpo}$ over an arbitrary precedence.

Proposition 8.4 *Let $\succ_p$ be a total precedence on $F \setminus F_{\mathrm{CA}}$. Then $\succeq_{\mathrm{CA}}(\succeq_p)$ is a total* AC*-antisymmetric and* AC*-compatible simplification quasi-ordering on ground terms and contains* CA.

Proof: Analogous to the case of commutative rings. □

Lemma 8.5 *Let ϕ be a product in* CA*-normal form, let r be a ground term such that $\phi \succ_{\mathrm{CA}} r$, and let b', b'' and m be elements of R such that $b' \succ_R b''$. Then $\langle b'\rangle\phi \succ_{\mathrm{CA}} \langle b''\rangle\phi + \langle m\rangle r$.*

Proof: Again we can use D ∪ AC-antisymmetry of $\succeq_1$ to infer $\phi \succ_1 r$ from ϕ being in CA-normal form. Suppose

$$\kappa(\mathrm{D}(r)) = \{\langle \mathrm{rep}(\phi_1), n_1, M_1\rangle, \ldots, \langle \mathrm{rep}(\phi_k), n_k, M_k\rangle\}.$$

Then

$$\kappa(\phi) = \{\langle \mathrm{rep}(\phi), 1, \{\emptyset\}\rangle\} \succ_\kappa \kappa(\mathrm{D}(r))$$

implies $\mathrm{rep}(\phi) \succ_{mul}(\succeq_c)\ \mathrm{rep}(\phi_i)$ for $i = 1, \ldots, k$. If we now compare

$$\kappa(\mathrm{D}(\langle b'\rangle\phi)) = \kappa(\langle b'\rangle\phi) = \{\langle \mathrm{rep}(\phi), 1, \{\{b'\}\}\rangle\}$$

and

$$\begin{aligned}\kappa(\mathrm{D}(\langle b''\rangle\phi + \langle m\rangle r)) &= \kappa(\langle b''\rangle\phi) \cup \kappa(\mathrm{D}(\langle m\rangle r)\\ &= \{\langle \mathrm{rep}(\phi), 1, \{\{b''\}\}\rangle\} \cup \{\langle \mathrm{rep}(\phi_1), n_1, M_1'\rangle, \ldots, \langle \mathrm{rep}(\phi_k), n_k, M_k'\rangle\}\end{aligned}$$

we see that all triples in $\kappa(\mathrm{D}(\langle m\rangle r))$ are smaller than $\langle \mathrm{rep}(\phi), 1, \{\{b'\}\}\rangle$ in the first component, and that $\langle \mathrm{rep}(\phi), 1, \{\{b''\}\}\rangle$ is smaller than $\langle \mathrm{rep}(\phi), 1, \{\{b'\}\}\rangle$ in the third component. □

Proposition 8.6 CA *is convergent modulo* AC.

Proof: Termination follows again from inclusion of CA in $\succeq_{\mathrm{CA}}$, hence confluence and coherence follow from the local properties of Lemma 8.1. □

8.2 Symmetrization

An equation $l \approx r$ is in CA-*normal form* if either (i) $l = r = 0$, or (ii) $l \succ_{\mathrm{CA}} r$, l is irreducible with respect to CA, and $l \approx r$ has one of the forms (a) $\langle 1\rangle \approx \langle 0\rangle$, (b) $\alpha \approx r$, (c) $\phi \approx r$ where ϕ is a proper product of atomic terms, or (d) $\langle b\rangle\phi \approx r$ where $R = \mathbb{Z}$ and $b > 1$, and ϕ is a proper product of atomic terms. The symmetrization function $\mathcal{S}_{\mathrm{CA}}$ maps each case in the definition of an CA-normal form to a set of rewrite rules as follows:

$$\begin{aligned}
\mathcal{S}_{\mathrm{CA}}(0 \approx 0) &= \emptyset && \text{(CA.S1)}\\
\mathcal{S}_{\mathrm{CA}}(\langle 1\rangle \approx \langle 0\rangle) &= \{t \Rightarrow \langle 0\rangle \mid t \neq \langle 0\rangle\} && \text{(CA.S2)}\\
\mathcal{S}_{\mathrm{CA}}(\alpha \approx r) &= \{\alpha \Rightarrow r\} && \text{(CA.S3)}\\
\mathcal{S}_{\mathrm{CA}}(\phi \approx r) &= \{\phi \Rightarrow r\} && \text{(CA.S4a)}\\
&\quad \cup\, gnd(\{x\phi \Rightarrow xr\}) && \text{(CA.S4b)}\\
\mathcal{S}_{\mathrm{CA}}(\langle b\rangle\phi \approx r) &= \{\langle b'\rangle\phi \Rightarrow \langle b''\rangle\phi + \langle m\rangle r \mid b' \Rightarrow_b^m b''\} && \text{(CA.S5a)}\\
&\quad \cup\, gnd(\{\langle b'\rangle x\phi \Rightarrow \langle b''\rangle x\phi + \langle m\rangle xr \mid b' \Rightarrow_b^m b''\}) && \text{(CA.S5b)}
\end{aligned}$$

Lemma 8.7 $\mathcal{S}_{\mathrm{CA}}$ *is a symmetrization function for* CA.

Proof: For termination the only nontrivial cases are (CA.S5a) and (CA.S5b), which are handled by Lemma 8.5.

Left-minimality follows in the same way as in the case of modules.

Next we consider soundness. We can infer (CA.S2) from $\langle 1\rangle \approx \langle 0\rangle$ by multiplying both sides with t and using the rules (CA.2) and (CA.3). Note that $t \neq 0$ is required by the ordering. There's nothing to prove for (CA.S3) and (CA.S4a). (CA.S4b) is obtained by compatibility with contexts. Soundness of (CA.S5a) can be shown analogously to the case of modules, using critical pairs with the rules (CA.6e) and (CA.9). This also works for (CA.S5b), one only has to add the context $x[]$ in the beginning.

It remains to show convergence modulo AC of $\mathrm{CA} \cup \mathcal{S}_{\mathrm{CA}}(l \approx r)$. Local coherence is obtained by the included AC-extensions. For local confluence the only cases to consider are again those of overlaps of rules in $\mathcal{S}_{\mathrm{CA}}(l \approx r)$ into CA strictly below the root position, and of overlaps of $\mathcal{S}_{\mathrm{CA}}(l \approx r)$ with itself.

(1) We have checked convergence of the overlaps of rules in the symmetrization into CA. We present only the two most interesting cases. We consider an equation $\langle b\rangle\phi \approx r$ in CA-normal form.

(1.1) The first is the overlap of (CA.S5a) into (CA.9):

$$\langle b''\rangle\phi + \langle m\rangle r + \langle b_0\rangle\phi \overset{\mathrm{CA.S5a}}{\Leftarrow} \langle b'\rangle\phi + \langle b_0\rangle\phi \overset{\mathrm{CA.9}}{\Rightarrow} \langle b' + b_0\rangle\phi,$$

where $b' \Rightarrow_b^m b''$. By the cancellation law $b' \neq b''$ implies $b'+b_0 \neq b''+b_0$. If $b'+b_0 \succ_R b''+b_0$ then

$$\langle b' + b_0\rangle\phi \overset{\mathrm{CA.S5a}}{\Rightarrow} \langle b'' + b_0\rangle\phi + \langle m\rangle r,$$

since $b' + b_0 \Rightarrow_b^m b'' + b_0$. Otherwise $b'' + b_0 \succ_R b' + b_0$ and

$$\langle b'' + b_0\rangle\phi + \langle m\rangle r \overset{\mathrm{CA.S5a}}{\Rightarrow} \langle b' + b_0\rangle\phi + \langle -m\rangle r + \langle m\rangle r \overset{*}{\Rightarrow}_{\mathrm{CA}} \langle b' + b_0\rangle\phi,$$

since $b'' + b_0 \Rightarrow_b^{-m} b' + b_0$.

(1.2) The other overlap which we consider is of (CA.S5b) into (CA.6):

$$\langle m_0\rangle(\langle b''\rangle\phi + \langle m\rangle r) \overset{\mathrm{CA.S5b}}{\Leftarrow} \langle m_0\rangle\langle b'\rangle\phi \overset{\mathrm{CA.6}}{\Rightarrow} \langle m_0 b\rangle\phi,$$

where again $b' \Rightarrow_b^m b''$. We can reduce the left-hand side to $\langle m_0 b''\rangle\phi + \langle m_0 m\rangle r$. This time we use the cancellation law for multiplication, which follows from R being an integral domain, to infer $m_0 b' \neq m_0 b''$ from $b' \neq b''$. If $m_0 b' \succ_R m_0 b''$ then

$$\langle m_0 b'\rangle\phi \overset{\mathrm{CA.S5a}}{\Rightarrow} \langle m_0 b''\rangle\phi + \langle m m_0\rangle r,$$

since $m_0 b' \Rightarrow_b^{m m_0} m_0 b''$. Otherwise $m_0 b'' \succ_R m_0 b'$ and

$$\langle m_0 b'\rangle\phi + \langle m m_0\rangle r \overset{\mathrm{CA.S5a}}{\Rightarrow} \langle m_0 b\rangle\phi + \langle -m m_0\rangle r + \langle m m_0\rangle r \overset{*}{\Rightarrow}_{\mathrm{CA}} \langle m_0 b\rangle\phi,$$

since $m_0 b'' \Rightarrow_b^{-m m_0} m_0 b'$.

(2) It remains to consider overlaps of $\mathcal{S}_{\mathrm{CA}}(\langle b\rangle\phi \approx r)$ into itself.

(2.1) An overlap at the top-level is

$$\langle b_1''\rangle\phi + \langle m_1\rangle r \overset{\mathrm{CA.S5a}}{\Leftarrow} \langle b'\rangle\phi \overset{\mathrm{CA.S5b}}{\Rightarrow} \langle b_2''\rangle\phi + \langle m_2\rangle r,$$

where $b_1'' \Leftarrow_b^{m_1} b' \Rightarrow_b^{m_2} b_2''$.

(2.1.1) If $b_1'' = b_2''$ then $m_1 b = b' - b_1'' = b' - b_2'' = m_2 b$, by canceling b we get $m_1 = m_2$, and we conclude that the peak is trivial.

(2.1.2) If $b_1'' \succ_R b_2''$ then

$$\langle b_1''\rangle\phi + \langle m_1\rangle r \stackrel{\text{CA.S5a}}{\Rightarrow} \langle b_2''\rangle\phi + \langle m_2 - m_1\rangle r + \langle m_1\rangle r \Rightarrow_{\text{CA}} \langle b_2''\rangle\phi + \langle m_2\rangle r,$$

since $b_1'' = b' - m_1 b = b_2'' + m_2 b - m_1 b = b_2'' + (m_2 - m_1)b$ implies $b_1'' \Rightarrow_b^{m_2-m_1} b_2''$.

(2.1.3) The remaining case of $b_2'' \succ_R b_1''$ is analogous.

(2.2) Between extensions with respect to multiplication we get the overlap

$$\langle b_2'\rangle(\langle b_1''\rangle\phi + \langle m_1\rangle r) \stackrel{\text{CA.S5b}}{\Leftarrow} \langle b_1'\rangle\langle b_2'\rangle\phi \stackrel{\text{CA.S5b}}{\Rightarrow} \langle b_1'\rangle(\langle b_2''\rangle\phi + \langle m_2\rangle r),$$

where $b_1' \Rightarrow_b^{m_1} b_1''$ and $b_2' \Rightarrow_b^{m_2} b_2''$. We have the reductions

$$\begin{aligned}
\langle b_2'\rangle(\langle b_1''\rangle\phi + \langle m_1\rangle r) &\Rightarrow_{\text{CA}} \langle b_1''\rangle\langle b_2'\rangle\phi + \langle b_2'\rangle\langle m_1\rangle r \\
&\stackrel{\text{CA.S5b}}{\Rightarrow} \langle b_1''\rangle(\langle b_2''\rangle\phi + \langle m_2\rangle r) + \langle b_2'\rangle\langle m_1\rangle r \\
&\stackrel{*}{\Rightarrow}_{\text{CA}} \langle b_1'' b_2''\rangle\phi + \langle m_2 b_1'' + m_1 b_2'\rangle r
\end{aligned}$$

and

$$\begin{aligned}
\langle b_1'\rangle(\langle b_2''\rangle\phi + \langle m_2\rangle r) &\Rightarrow_{\text{CA}} \langle b_2''\rangle\langle b_1'\rangle\phi + \langle b_1'\rangle\langle m_2\rangle r \\
&\stackrel{\text{CA.S5b}}{\Rightarrow} \langle b_2''\rangle(\langle b_1''\rangle\phi + \langle m_1\rangle r) + \langle b_1'\rangle\langle m_2\rangle r \\
&\stackrel{*}{\Rightarrow}_{\text{CA}} \langle b_1'' b_2''\rangle\phi + \langle m_1 b_2'' + m_2 b_1'\rangle r.
\end{aligned}$$

The last terms of these reductions are equal, since

$$\begin{aligned}
m_2 b_1'' + m_1 b_2' &= m_2 b_1'' + m_1(b_2'' + m_2 b) = m_2 b_1'' + m_1 b_2'' + m_1 m_2 b \\
m_1 b_2'' + m_2 b_1' &= m_1 b_2'' + m_2(b_1'' + m_1 b) = m_1 b_2'' + m_2 b_1'' + m_2 m_1 b
\end{aligned}$$

are equal. □

8.3 Critical extension peaks and transitivity

The critical extension peaks for algebras turn out to be analogous to those of commutative rings.

Theorem 8.8 *Let $\langle b_i\rangle\phi_i \Rightarrow r_i$ be a rewrite rule in* Norm_{CA} *for $i = 1, 2$ and assume without loss of generality $b_1 \succeq_R b_2$.*

These two rules have at most the single critical extension peak

$$r_1\psi_1 \Leftarrow \langle b_1\rangle\phi \Rightarrow \langle b_1 - m_2 b_2\rangle\phi + \langle m_2\rangle r_2\psi_2$$

where $\phi =_{\text{AC}} \text{lcm}(\phi_1, \phi_2) =_{\text{AC}} \phi_1\psi_1 =_{\text{AC}} \phi_2\psi_2$ and $b_1 - m_2 b_2 \prec_R b_2$, if (i) $\psi_1 \neq 1$, (ii) $b_1 \succ_R b_2$ or $\psi_2 \neq 1$, and (iii) either (a) $b_1 \succ_R 1$ and $b_2 \succ_R 1$, (b) $b_1 \succ_R 1$, $b_2 = 1$ and ϕ_2 is a proper product with $\gcd(\phi_1, \phi_2) \neq 1$, or (c) $b_1 = b_2 = 1$ and ϕ_1 and ϕ_2 are proper products with $\gcd(\phi_1, \phi_2) \neq 1$.

Otherwise there is no critical extension peak between these two rules.

$$\begin{tikzcd} & \langle b_1\rangle\phi_1\phi_2 \arrow[dl, "S_1"'] \arrow[dr, "S_2"] & \\ r_1\phi_2 \arrow[dr, dashed, "S_2"'] & & \langle b_1\rangle\phi_1 r_2 \arrow[dl, dashed, "S_1"] \\ & r_1 r_2 & \end{tikzcd}$$

Figure 8.1: Convergence of an extended peak

Proof: The peak is clearly the minimal peak of the two rules. There are two things to show, namely that nonminimal peaks are redundant, and that minimal peaks violating the condition are either not extension peaks or redundant. We start with the latter. Let $S_i = \mathcal{S}_{\mathrm{CA}}(\langle b_i\rangle\phi_i \Rightarrow r_i)$ for $i = 1, 2$.

(1) Suppose one of the conditions (i)–(iii) is violated. First note that condition (i) is equivalent to $\langle b_1\rangle\phi_1$ being irreducible with respect to S_2, and that condition (ii) is equivalent to $\langle b_2\rangle\phi_2$ being irreducible with respect to S_1, and that this also holds for a symmetrization of the form (CA.S2). Hence if (i) or (ii) is violated then the peak is not an extension peak.

Now suppose (iii) is violated. Then $b_2 = 1$, as otherwise (a) would hold. If $b_i = 1$ and ϕ_i is not a proper product then S_i is of the form (CA.S3) and contains no extended rules. If $b_2 = 1$ and $\gcd(\phi_1, \phi_2) = 1$ then the peak is redundant, as Figure 8.1 shows.

(2) It remains to show that any peak $t_1 \Leftarrow s \Rightarrow t_2$ between $\langle b_1\rangle\phi_1 \Rightarrow r_1$ and $\langle b_2\rangle\phi_2 \Rightarrow r_2$ that is not of the form above is redundant. That is, the corresponding Extension Superposition inference

$$\frac{\langle b_1\rangle\phi_1 \approx r_1 \qquad \langle b_2\rangle\phi_2 \approx r_2}{t_1 \approx t_2}$$

with main premise $C = t_1 \not\approx s \vee s \not\approx t_2 \vee t_1 \approx t_2$ has to be redundant. If s is reducible by CA then C is redundant by Lemma 4.6, which implies redundancy of the peak by Lemma 4.25. So let us assume from now on that s is irreducible with respect to CA. This is the case if and only if s has the form $\langle b\rangle\phi$. Note that for $b = 1$ we write $\langle b\rangle\phi$ to represent ϕ. This poses no problems, since it leads to essentially the same rewrite steps. To show redundancy of the peak, we have to show

$$\mathrm{Trans}_C \cup \{\langle b_i\rangle\phi_i \Rightarrow r_i \mid i = 1, 2\} \models_I t_1 \approx t_2.$$

Consider some interpretation I_N that satisfies $\mathrm{Trans}_C \cup \{\langle b_i\rangle\phi_i \Rightarrow r_i \mid i = 1, 2\}$. Then R_N contains these two rules, $S_N = \mathcal{S}_{\mathrm{CA}}(R_N)$, and $\mathrm{CA} \cup S_N$ is Church-Rosser below s. Let $S_i = \mathcal{S}_{\mathrm{CA}}(\langle b_i\rangle\phi_i \approx r_i)$ for $i = 1, 2$. The extended rules are of the forms (CA.S4b), (CA.S5a) and (CA.S5b). We can ignore extended rules of the form (CA.S2), since these reduce any nonzero term and there can thus be no extension peak. To simplify matters we consider only (CA.S5b), which modulo ACU is a generalization of (CA.S4b) and (CA.S5a). Thus we assume that the overlapping rules have the form

$$\langle b\rangle\phi \Rightarrow \langle b_i'\rangle\phi + \langle m_i'\rangle\psi_i r_i$$

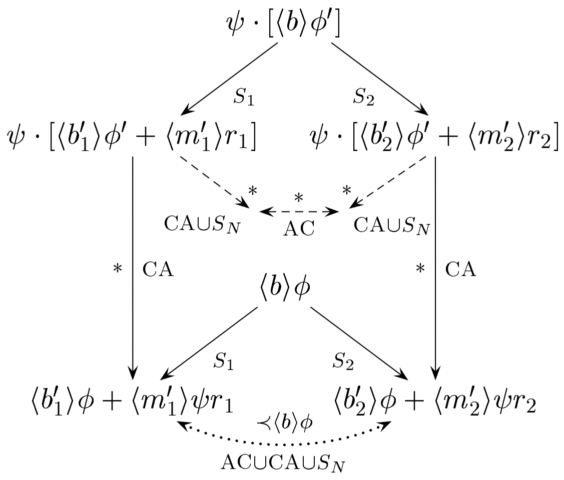


Figure 8.2: Redundancy of an extended peak

where $b_i' = b - m_i'b_i$ and $\phi =_{\mathrm{AC}} \phi_i\psi_i$ for $i = 1, 2$. We assume without loss of generality that ψ_1 and ψ_2 are chosen such that the rules overlap at the root position. ψ_1 and ψ_2 are products, since as subterms of s they are irreducible with respect to CA. Then $s = \langle b\rangle\phi$ and $t_i = \langle b_i'\rangle\phi + \langle m_i'\rangle\psi_i r_i$ for $i = 1, 2$. We have to find a proof $t_1 \Downarrow_{\mathrm{CA}\cup S_N} t_2$ for each case where the peak is not of the form above. To this end it suffices to exhibit a proof $t_1 \overset{*}{\Leftrightarrow}{}^{\prec s} t_2$, since $\mathrm{CA} \cup S_N$ is Church-Rosser below s.

(2.1) Suppose ϕ is not the least common multiple of ϕ_1 and ϕ_2. Then ψ_1 and ψ_2 have a common factor $\psi \neq 1$, and we let ψ_1' and ψ_2' such that $\psi_i =_{\mathrm{AC}} \psi\psi_i'$. This case is illustrated in Figure 8.2. We can also factor out ψ from ϕ, i.e. $\phi =_{\mathrm{AC}} \phi_i\psi_i'\psi =_{\mathrm{AC}} \phi'\psi$ where we let $\phi' =_{\mathrm{AC}} \phi_i\psi_i'$. Furthermore, we let $s' = \langle b\rangle\phi'$ and $t_i' = \langle b_i'\rangle\phi' + \langle m_i'\rangle\psi_i' \cdot r_i$, S_i contains a rule $s' \Rightarrow t_i'$ and there is a peak $t_1' \Leftarrow_{S_1} s' \Rightarrow_{S_2} t_2'$. Since s' is smaller than s, the corresponding transitivity instance

$$t_1' \not\approx s' \;\vee\; s' \not\approx t_2' \;\vee\; t_1' \approx t_2'$$

is smaller than C, and $t_1' \approx t_2'$ is true in I_N by induction hypothesis. By putting the context $\psi \cdot [\,]$ around every term in the proof $t_1' \Downarrow t_2'$ (dashed) and normalizing with respect to CA, we obtain a proof $t_1 \overset{*}{\Leftrightarrow} t_2$ (dotted). Since any term t in $t_1' \Downarrow t_2'$ is smaller than s', the CA-normal form of $\psi \cdot t$ is smaller than $s =_{\mathrm{AC}} \psi \cdot s'$. These normal forms bound the terms in $t_1 \overset{*}{\Leftrightarrow} t_2$, hence all terms in $t_1 \overset{*}{\Leftrightarrow} t_2$ are smaller than s. We conclude $t_1 \Downarrow_{\mathrm{CA}\cup S_N} t_2$.

(2.2) Suppose this peak is not minimal due to the choice of coefficients.

(2.2.1) Suppose $b \succ_R b_1$. By definition of the symmetrization function this can only be the case for $R = \mathbb{Z}$. We may assume that the smaller peak

$$\psi_1 r_1 \Leftarrow \langle b_1\rangle\phi \Rightarrow \langle b_1 - m_2 b_2\rangle\phi + \langle m_2\rangle\psi_2 r_2$$

converges. We let $m = 1$ for $b > 0$ and $m = -1$ for $b < 0$, and put the context

$$u[\,] = \langle b\rangle\phi + \langle m\rangle([\,] - \langle b_1\rangle\phi)$$

around every term in the peak and the dashed valley proof (see Figure 8.3). We normalize

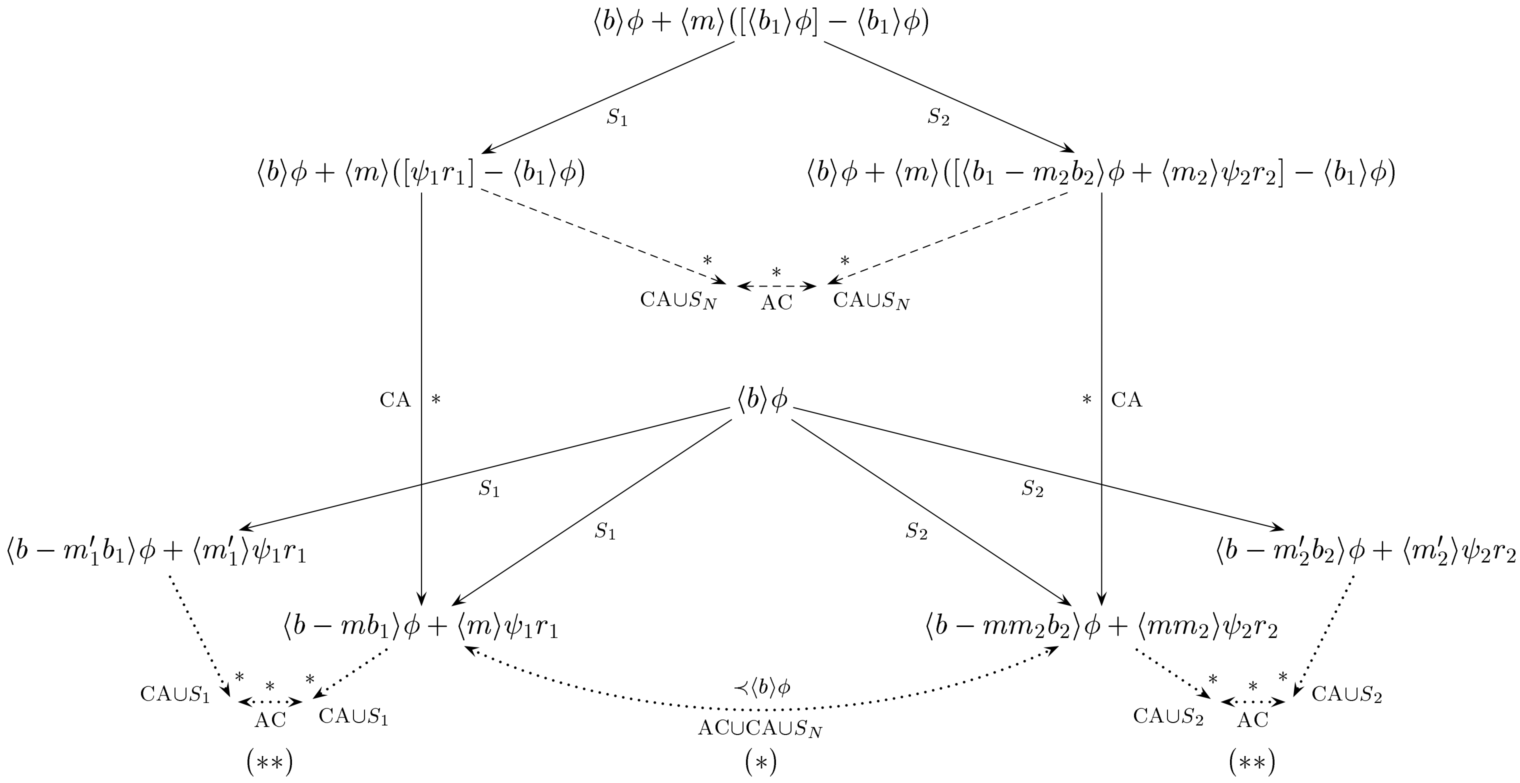


Figure 8.3: Redundancy of an extended peak due to a nonminimal coefficient at the top

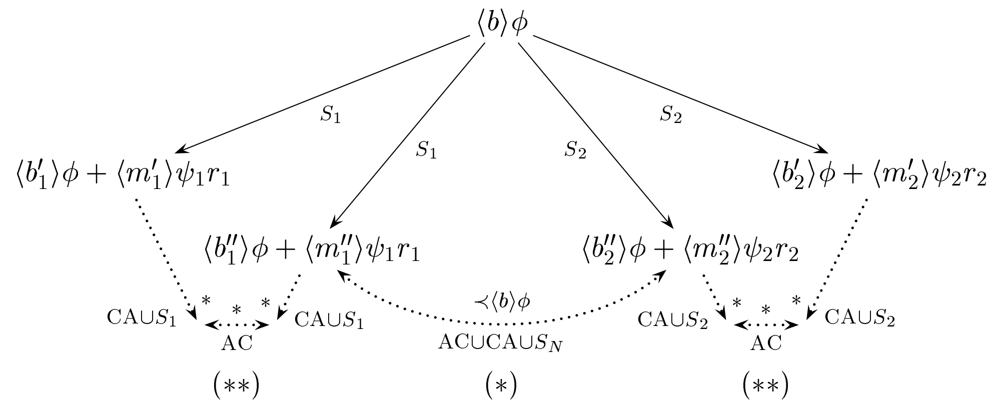


Figure 8.4: Redundancy of an extended peak due to nonminimal coefficients at the sides

with respect to CA and obtain a proof $(*)$ that connects the side terms of the inner peak

$$\langle b - mb_1\rangle\phi + \langle m\rangle\psi_1 r_1 \Leftarrow \langle b\rangle\phi \Rightarrow \langle b - mm_2 b_2\rangle\phi + \langle mm_2\rangle\psi_2 r_2.$$

Each rewrite step of the original proof becomes a valley proof in $(*)$ by Lemma 4.30. Since valley proofs are bounded by the terms at their ends, it suffices to show that the normalized terms of the original proof stay below $\langle b\rangle\phi$. We consider some term $\langle b'\rangle\phi$ with $b' \prec_R b_1$:

$$u[\langle b'\rangle\phi + r] = \langle b\rangle\phi + \langle m\rangle(\langle b'\rangle\phi + r - \langle b_1\rangle\phi)$$

where $\phi \succ_{\mathrm{CA}} r$. It is normalized to

$$\langle b + m(b' - b_1)\rangle\phi + \langle m\rangle r.$$

m was chosen such that $b_1 \succ_R b'$ implies $b \succ_R b + m(b' - b_1)$: the interval $[0, b_1)$ is mapped to $[b - b_1, b)$ for $b > 0$ and $m = 1$, and to $(b, b + b_1]$ for $b < 0$ and $m = -1$, and in both cases b is greater with respect to $\succ_R$ than any element of the interval. Hence $\langle b\rangle\phi \succ_R \langle b + m(b' - b_1)\rangle\phi + \langle m\rangle r$. The existence of the rewrite proofs $(**)$ on the left and the right follows from S_1 and S_2 being strongly symmetrized with respect to CA.

(2.2.2) Suppose $b = b_1$ but the side terms of the peak are not minimal. Then there exists a peak $b_1'' \Leftarrow b \Rightarrow b_2''$ such that $\{b_1'', b_2''\} \prec \{b_1', b_2'\}$. We then have the situation of Figure 8.4. The dotted proof stays below $\langle b\rangle\phi$. The proof $(*)$ exists since the peak in the middle is smaller than the original peak at the outside; hence it converges by assumption. The valley proofs $(**)$ on the sides exist by strong symmetrization. □

Corollary 8.9 *1. For $R = \mathbb{Z}$ any critical ground term t is of the form $\langle b\rangle\phi$ where $b > 0$.*

2. For a field R any critical ground term t is of the form ϕ.

We will now use this to give a sufficient criterion for a term t being in the critical closure of a term s.

Lemma 8.10 *Let s and t be* CA*-irreducible ground terms, and let $\langle b\rangle\phi$ be the maximal monomial of t. If $\langle b\rangle\phi \preceq s$ then $t \in \mathrm{cc}_{\mathrm{CA}}(s)$.*

Proof: Analogous to the case of commutative rings. □

We will need this to show that CA-Isolation is a simplification.

8.4 Simplification

The rules for simplification are strictly analogous to those for modules over R. However, the proof that these transformations are indeed simplifications is more complicated in this case, since transitivity is not universally valid.

CA-*Isolation* $$\frac{[\neg](\langle b'\rangle s' + p \approx \langle b''\rangle s'' + q)}{[\neg](\langle b\rangle s' \approx \langle u\rangle (q-p))}\rightarrow$$

if (i) $s' =_{\mathrm{AC}} s''$, (ii) $b = u(b'-b'')$ where (a) $R=\mathbb{Z}$, $b \geq 1$ and $u = \mathrm{sign}(b'-b'')$ for $b' \neq b''$, or (b) R is a field, $b=1$ and $u=(b'-b'')^{-1}$ for $b'\neq b''$, or (c) $b=0$ and $u=1$ for $b'=b''$, (iii) $\langle b'\rangle s' + p \succ \langle b\rangle s'$ or $\langle b''\rangle s'' + q \succ \langle b\rangle s'$, and (iv) $s' \succ p$ and $s'' \succ q$.

Lemma 8.11 CA-*Isolation is a simplification rule.*

Proof: In the proof we write s for $s' =_{\mathrm{AC}} s''$. Let

$$L = [\neg](\langle b'\rangle s + p \approx \langle b''\rangle s + q)$$
$$L' = [\neg](\langle b\rangle s \approx \langle u\rangle (q-p)).$$

Using monotonicity with respect to the contexts

$$\langle u\rangle ([\,] - \langle b''\rangle s - p)$$

for positive and

$$\langle u^{-1}\rangle [\,] + \langle b''\rangle s + p$$

for negative literals and normalizing with CA we can transform $\langle b'\rangle s + p \approx \langle b''\rangle s + q$ into $\langle b\rangle s \approx \langle u\rangle (q-p)$ and vice-versa. Thus $\{L\} \cup \mathrm{CA}_1 \models L'$. It remains to show

$$\{L'\} \cup \mathrm{Trans}_L \models_I L.$$

Suppose that L is the smallest literal where this does not hold. We will show that this leads to a contradiction.

(1) Suppose s is reducible by CA, say to t. Then by induction hypothesis the smaller instance

$$\frac{[\neg](\langle b'\rangle t + p \approx \langle b''\rangle t + q)}{[\neg](\langle b\rangle t \approx \langle u\rangle (q-p))}\rightarrow$$

of CA-Isolation is a simplification, which by Lemma 4.11 implies that the original instance is also a simplification.

(2) Otherwise s is irreducible with respect to CA. Then either $s = \phi$ or $s = \langle d\rangle \phi$ where ϕ is irreducible. For the case $s=\phi$ we let $d=1$ in the following. We also assume without loss of generality $\langle b'\rangle s' \succeq \langle b''\rangle s''$.

(2.1) Suppose L_1 and L_2 are positive. Then we have to show

$$\{\langle b\rangle s \approx \langle u\rangle (q-p)\} \cup \mathrm{Trans}_L \models_I \langle b'\rangle s + p \approx \langle b''\rangle s + q.$$

Let I_N be an interpretation such that $\{\langle b\rangle s \approx \langle u\rangle (q-p)\} \cup \mathrm{Trans}_L$ is true in I_N. Then there exists a valley proof $\langle b\rangle s \Downarrow_{\mathrm{CA}\cup S_N} \langle u\rangle (q-p)$ (dashed in Figure 8.5). We put the

$$
\begin{tikzcd}
\langle u^{-1}\rangle[\langle b\rangle s] + \langle b''\rangle s + p \arrow[dr, dashed, "*", "\mathrm{CA}\cup S_N"'] \arrow[dd, "*", "\mathrm{CA}"'] & & & \langle u^{-1}\rangle[\langle u\rangle(q-p)] + \langle b''\rangle s + p \arrow[dl, dashed, "*"', "\mathrm{CA}\cup S_N"] \arrow[dd, "*", "\mathrm{CA}"'] \\
& {} \arrow[r, dashed, leftrightarrow, "*", "\mathrm{AC}"'] & {} & \\
\langle db'\rangle\phi + p \arrow[rrr, dotted, leftrightarrow, bend right, "\preceq\langle b'\rangle s", "\mathrm{AC}\cup\mathrm{CA}\cup S_N"'] & & & \langle db''\rangle\phi + q
\end{tikzcd}
$$

Figure 8.5: Positive isolation is a simplification

context $\langle u^{-1}\rangle[\,] + \langle b''\rangle s + p$ around every term in this proof and normalize with respect to CA. We obtain a proof $\langle db'\rangle\phi + p \overset{*}{\Leftrightarrow} \langle db''\rangle\phi + q$ (dotted). It remains to check that any monomial in this proof is bounded by $\langle b'\rangle s$.

If $s = \langle d\rangle\phi$ for $d \neq 1$ then any of the normalized terms contains only one coefficient. Hence it is smaller than $\langle b'\rangle s = \langle b'\rangle\langle d\rangle\phi$, which contains two.

Otherwise $s = \phi$ and $d = 1$. For the case where R is a field the coefficients do not matter, because only ϕ alone is a critical term. For the case $R = \mathbb{Z}$ the coefficients in the original proof are in the interval $[0, b]$. If $b' \geq b'' \geq 0$ then $u = 1$ and the coefficients are mapped to $[b'', b']$. Otherwise $b' < b''$ and $b' < 0$, $u = -1$ and the coefficients are mapped to $[b', b'']$. In both cases the interval is bounded by b' with respect to $\succ_R$.

Hence transitivity holds for the terms in the dotted proof and there exists a rewrite proof $\langle db'\rangle\phi + p \Downarrow_{\mathrm{CA}\cup S_N} \langle db''\rangle\phi + q$. By composing this proof with $\langle b'\rangle s \overset{*}{\Rightarrow}_{\mathrm{CA}} \langle db'\rangle\phi$ and $\langle db''\rangle\phi \overset{*}{\Leftarrow}_{\mathrm{CA}} \langle b''\rangle s$ at the ends we obtain a rewrite proof showing that $\langle b'\rangle s + p \approx \langle b''\rangle s + q$ is true in I_N.

(2.2) Otherwise L and L' are negative, and we have to show

$$\{\langle b'\rangle s + p \approx \langle b''\rangle s + q\} \cup \mathrm{Trans}_L \models_I \langle b\rangle s \approx \langle u\rangle(q - p).$$

Let I_N be an interpretation such that $\{\langle b'\rangle s + p \approx \langle b''\rangle s + q\} \cup \mathrm{Trans}_L$ is true in I_N. Then there exists a valley proof $\langle b'\rangle s + p \Downarrow_{\mathrm{CA}\cup S_N} \langle b''\rangle s + q$ (dashed in Figure 8.6). We put the context $\langle u\rangle([\,] - \langle b''\rangle s - p)$ around every term in this proof and normalize with respect to CA. We obtain a proof $\langle db\rangle\phi \overset{*}{\Leftrightarrow} \langle u\rangle(q - p)$ (dotted). Again we have to verify that the monomials in this proof are bounded by $\langle b'\rangle s$.

If $s = \langle d\rangle\phi$ for $d \neq 1$ then again any of the normalized terms is smaller than $\langle b'\rangle s$ due to a smaller number of coefficients.

Otherwise $s = \phi$ and $d = 1$. For fields it suffices to note that the greatest product in the proof is ϕ. For integers $b' \succeq_R b''$ implies that either $b' \geq b'' \geq 0$ or $b' < 0$ and $b' < b''$. In the first case the coefficients of the untransformed proof are in $[b'', b']$, $u = 1$ and the coefficients are mapped by $x \mapsto u(x - b'')$ to $[0, b]$. In the second case $u = -1$ and the coefficients are mapped from $[b', b'']$ to $[0, b]$.

Again all the terms in the dotted proof stay below the bound, transitivity holds, and there exists a rewrite proof $\langle db\rangle\phi \Downarrow_{\mathrm{CA}\cup S_N} \langle u\rangle(q - p)$. By composing this with $\langle b\rangle s \overset{*}{\Rightarrow}_{\mathrm{CA}} \langle db\rangle\phi$ we obtain a rewrite proof for $\langle b\rangle s \approx \langle u\rangle(q - p)$. $\square$

$$
\begin{tikzcd}
\langle u\rangle([\langle b'\rangle s+p]-\langle b''\rangle s-p) \arrow[dr, dashed, "*", "\mathrm{CA}\cup S_N"'] \arrow[dd, "\mathrm{CA}"', "*"] & & & \langle u\rangle([\langle b''\rangle s+q]-\langle b''\rangle s-p) \arrow[dl, dashed, "*"', "\mathrm{CA}\cup S_N"] \arrow[dd, "\mathrm{CA}"', "*"] \\
& {} \arrow[r, dashed, leftrightarrow, "*", "\mathrm{AC}"'] & {} & \\
\langle db\rangle\phi \arrow[rrr, dotted, leftrightarrow, bend right, "\preceq_{\langle b'\rangle s}", "\mathrm{AC}\cup\mathrm{CA}\cup S_N"'] & & & \langle u\rangle(q-p)
\end{tikzcd}
$$

Figure 8.6: Negative isolation is a simplification

Like in the preceeding chapters the simplification function will consists of a subset of these simplifications with additional ordering restrictions.

Simp$_{\mathrm{CA}}$-*Sum Contraction* $\qquad \dfrac{[\neg](\langle b\rangle s+\langle b'\rangle s'+p\approx q)}{[\neg](\langle b+b'\rangle s+p\approx q)}$

if (i) $s=_{\mathrm{AC}}s'$, (ii) $\langle b\rangle s$ is a maximal summand in $\langle b\rangle s+\langle b'\rangle s'+p$, and (iii) $\langle b\rangle s+\langle b'\rangle s'+p\succeq q$.

Simp$_{\mathrm{CA}}$-*Summand Rewriting* $\qquad \dfrac{[\neg](u[l]+p\approx q)}{[\neg](u[r]+p\approx q)}$

if (i) $l\Rightarrow r$ is a rule in CA, (ii) $u[l]$ is a maximal summand in $u[l]+p$, and (iii) $u[l]+p\succeq q$.

Simp$_{\mathrm{CA}}$-*Isolation* $\qquad \dfrac{[\neg](\langle b'\rangle\phi'+p\approx\langle b''\rangle\phi''+q)}{[\neg](\langle b\rangle\phi'\approx\langle u\rangle(q-p))}$

if (i) $\phi'=_{\mathrm{AC}}\phi''$, (ii) $b=u(b'-b'')$ where (a) $R=\mathbb{Z}$, $b\geq 1$ and $u=\mathrm{sign}(b'-b'')$ for $b'\neq b''$, or (b) R is a field, $b=1$ and $u=(b'-b'')^{-1}$ for $b'\neq b''$, or (c) $b=0$ and $u=1$ for $b'=b''$, (iii) $b'\succ_R b$ or $p\neq 0$, (iv) $\phi'=\alpha_1\ldots\alpha_k$, $k\geq 0$, α_i is CA-atomic for $i=1,\ldots,k$, (v) ϕ' is irreducible with respect to CA, (vi) $\phi'\succ p$ and $\phi''\succ q$, and (vii) $\langle b'\rangle\phi'+p\succeq\langle b''\rangle\phi''+q$.

We let Simp$_{\mathrm{CA}}(L)$ consist of all literals L' such that there exists a simplification by Simp$_{\mathrm{CA}}$-Rewriting or Simp$_{\mathrm{CA}}$-Isolation with premise L and conclusion L'.

Lemma 8.12 Simp$_{\mathrm{CA}}$ *is an admissible simplification function for* Norm$_{\mathrm{CA}}$.

Proof: Simp$_{\mathrm{CA}}$-Rewriting and Simp$_{\mathrm{CA}}$-Isolation are restrictions of the rules proven to be simplifications in Lemma 8.11, hence they are simplification rules.

It remains to show that Simp$_{\mathrm{CA}}$ can simplify any literal that is not in CA-normal form. This proof is strictly analogous to that for modules. □

8.5 The inference system

This leads to the following ground inference system.

CA-*Sum Contraction*
$$\frac{[\neg](\langle b\rangle s + \langle b'\rangle s' + p \approx q) \ \vee\ C}{[\neg](\langle b + b'\rangle s + p \approx q) \ \vee\ C}$$

if (i) $s =_{\mathrm{AC}} s'$, (ii) $\langle b\rangle s$ is a maximal summand in $\langle b\rangle s + \langle b'\rangle s' + p$, and (iii) $\langle b\rangle s + \langle b'\rangle s' + p \succeq q$.

CA-*Summand Rewriting*
$$\frac{[\neg](u[l] + p \approx q) \ \vee\ C}{[\neg](u[r] + p \approx q) \ \vee\ C}$$

if (i) $l \Rightarrow r$ is a rule in CA, (ii) $u[l]$ is a maximal summand in $u[l] + p$, and (iii) $u[l] + p \succeq q$.

CA-*Isolation*
$$\frac{[\neg](\langle b'\rangle \phi' + p \approx \langle b''\rangle \phi'' + q) \ \vee\ C}{[\neg](\langle b\rangle \phi' \approx \langle u\rangle (q - p)) \ \vee\ C}$$

if (i) $\phi' =_{\mathrm{AC}} \phi''$, (ii) $b = u(b' - b'')$ where (a) $R = \mathbb{Z}$, $b \geq 1$ and $u = \mathrm{sign}(b' - b'')$ for $b' \neq b''$, or (b) R is a field, $b = 1$ and $u = (b' - b'')^{-1}$ for $b' \neq b''$, or (c) $b = 0$ and $u = 1$ for $b' = b''$, (iii) $b' \succ_R b$ or $p \neq 0$, (iv) $\phi' = \alpha_1 \ldots \alpha_k$, $k \geq 0$, α_i is CA-atomic for $i = 1, \ldots, k$, (v) ϕ' is irreducible with respect to CA, (vi) $\phi' \succ p$ and $\phi'' \succ q$, and (vii) $\langle b'\rangle \phi' + p \succeq \langle b''\rangle \phi' + q$.

CA-*Superposition A*
$$\frac{\phi \approx r \ \vee\ D \qquad [\neg](s[\phi'] \approx t) \ \vee\ C}{[\neg](s[r] \approx t) \ \vee\ C \ \vee\ D}$$

if (i) $\phi =_{\mathrm{AC}} \phi'$, (ii) $[\neg](s[\phi'] \approx t)$ is in CA-normal form, and (iii) $\phi \approx r$ is in CA-normal form.

The preceding rule combines (CA.S3) and (CA.S4a).

CA-*Superposition B*
$$\frac{\phi \approx r \ \vee\ D \qquad [\neg](s[\phi'] \approx t) \ \vee\ C}{[\neg](s[\psi r] \approx t) \ \vee\ C \ \vee\ D}$$

if (i) $\psi\phi =_{\mathrm{AC}} \phi'$, (ii) ϕ is a proper product, (iii) $[\neg](s[\phi'] \approx t)$ is in CA-normal form, and (iv) $\phi \approx r$ is in CA-normal form.

This is for superposition with rules of the form (CA.S4b).

CA-*Superposition C*
$$\frac{\langle b\rangle \phi \approx r \ \vee\ D \qquad [\neg](s[l] \approx t) \ \vee\ C}{[\neg](s[\langle b''\rangle \phi + \langle m\rangle r] \approx t) \ \vee\ C \ \vee\ D}$$

if (i) $l =_{\mathrm{AC}} \langle b'\rangle \phi$, (ii) $b' \Rightarrow_b^m b''$, (iii) $R = \mathbb{Z}$ and $b > 1$, (iv) ϕ is a proper product of atomic terms, (v) $0 \leq b'' < b$, (vi) $[\neg](s[l] \approx t)$ is in CA-normal form, and (vii) $\langle b\rangle \phi \approx r$ is in CA-normal form.

This is for case (CA.S5a).

CA-*Superposition D* $\dfrac{\langle b\rangle\phi \approx r \;\vee\; D \qquad [\neg](s[l] \approx t) \;\vee\; C}{[\neg](s[\langle b''\rangle\psi\phi + \langle m\rangle\psi r] \approx t) \;\vee\; C \;\vee\; D}$

if (i) $l =_{\mathrm{AC}} \langle b'\rangle\psi\phi$, (ii) $b' \Rightarrow_b^m b''$, (iii) $R = \mathbb{Z}$ and $b > 1$, (iv) ϕ is a proper product of atomic terms, (v) $0 \leq b'' < b$, (vi) $[\neg](s[l] \approx t)$ is in CA-normal form, and (vii) $\langle b\rangle\phi \approx r$ is in CA-normal form.

This is for rules of the form (CA.S5b) in the symmetrization.

CA-*Extension Superposition* $\dfrac{\langle b_1\rangle\phi_1 \approx r_1 \;\vee\; C_1 \qquad \langle b_2\rangle\phi_2 \approx r_2 \;\vee\; C_2}{r_1\psi_1 \approx \langle b_1 - m_2b_2\rangle\phi + \langle m_2\rangle r_2\psi_2 \;\vee\; C_1 \;\vee\; C_2}$

if (i) $\phi =_{\mathrm{AC}} \mathrm{lcm}(\phi_1, \phi_2) =_{\mathrm{AC}} \phi_1\psi_1 =_{\mathrm{AC}} \phi_2\psi_2$, (ii) $0 \leq b_1 - m_2b_2 < b_2$, (iii) $\psi_1 \neq 1$, (iv) $b_1 > b_2$ or $\psi_2 \neq 1$, (v) either (a) $b_1 > 1$ and $b_2 > 1$, (b) $b_1 > 1$, $b_2 = 1$ and ϕ_2 is a proper product with $\gcd(\phi_1, \phi_2) \neq 1$, or (c) $b_1 = b_2 = 1$ and ϕ_1 and ϕ_2 are proper products with $\gcd(\phi_1, \phi_2) \neq 1$, (vi) $\langle b_i\rangle\phi_i \Rightarrow r_i$ is in CA-normal form for $i = 1, 2$.

The main premise of this inference is the transitivity instance

$$r_1\psi_1 \not\approx \langle b_1\rangle\phi \;\vee\; \langle b_1\rangle\phi \not\approx \langle b_1 - m_2b_2\rangle\phi + \langle m_2\rangle r_2\psi_2 \;\vee\; r_1\psi_1 \approx \langle b_1 - m_2b_2\rangle\phi + \langle m_2\rangle r_2\psi_2.$$

Reflexivity Resolution $\dfrac{p \not\approx q \;\vee\; C}{C}$

if (i) $p =_{\mathrm{AC}} q$, (ii) $p \not\approx q$ is in CA-normal form.

CA-*Equality Factoring* $\dfrac{s \approx r \;\vee\; t \approx r' \;\vee\; C}{r \not\approx r' \;\vee\; t \approx r' \;\vee\; C}$

if (i) $s =_{\mathrm{AC}} t$, (ii) $s \approx r$ and $t \approx r'$ are in CA-normal form, and (iii) $r \succeq r'$.

We call the set of these rules $\mathsf{Sup}_{\mathrm{CA}}$.

Theorem 8.13 $\mathsf{Sup}_{\mathrm{CA}}$ *is refutationally complete for* CA_1*.*

Proof: Strictly analogous to the proof for commutative rings, this follows from Theorem 4.20 in combination with Propositions 8.6 and 8.4, and Lemmas 8.7 and 8.12. By using Theorem 8.8 we again restrict CA-Extension Superposition to critical extension peaks. $\square$

Like in the case of modules it is possible to replace sequences of superpositions in the root context by GCD Superposition inferences, as described in Section 7.6. After the first superposition inference has unified the products ϕ_1 and ϕ_2 from the two clauses, resulting in their least common multiple ϕ, the sequence of superpositions proceeds with maximal product ϕ strictly analogously to the case of modules.

9

Lifting

In this chapter we discuss informally how our calculi can be lifted, the specific problems that arise and how they can be resolved. We do not do this formally, as this is a technically involved and tedious exercise that does not in itself provide new insights. Nevertheless it will be needed for the implementation of these calculi.

9.1 Free variable lifting

Free variable lifting represents a set of ground terms (or formulas) as the set of all ground instances of some nonground term. This limits the expressibility, and leads to problems when the set of ground instances is not uniform. Take for instance the inference rule AG-Isolation. If we apply this rule to a nonground equation of the form $x + p \approx q$ then we can in general instantiate x by a sum of maximal and nonmaximal terms, and these are treated differently in the ground inference. This difference must be reflected in the lifted inference.

AG-*Isolation*
$$\frac{[\neg](p \approx q) \,\vee\, C}{([\neg]((n_1 - n_2)s \approx y_2 - y_1) \,\vee\, C)\sigma}$$

if (i) $\sigma \in \mathit{CSS}(p \approx_{\mathrm{AC}} n_1 s + y_1 \,\wedge\, q \approx_{\mathrm{AC}} n_2 s + y_2)$ (ii) $s\sigma$ is AG-atomic or a variable, (iii) $s\sigma$ is irreducible with respect to AG, (iv) $n_1 \geq n_2$, (v) $n_2 \neq 0$ or $y_1\sigma \neq 0$, and (vi) $s\sigma \not\preceq y_1\sigma$ and $s\sigma \not\preceq y_2\sigma$.

Here $\mathit{CSS}(p \approx_{\mathrm{AC}} n_1 x + y_1 \,\wedge\, q \approx_{\mathrm{AC}} n_2 x + y_2)$ denotes a *complete set of solutions* of the unification problem, analogously to the complete set of AC-unifiers for a single equation. The main problem with this inference is that it can introduce additional term structure below free variables, and that it does this in a rather prolific way. Since n_1 and n_2 are not bounded, there is an infinite number of conclusions if p or q contain free variables. To preserve refutational completeness of the calculus the enumeration of conclusions has to be interleaved with other inferences.

Since inferences are restricted to maximal terms the AC-unification problems will generally be smaller than with more general calculi. However, the double exponential case can still occur for abelian groups and commutative rings, as superpositions with an equations of the form $x + \cdots + x \approx r$ into a term $y_1 + \cdots + y_k$ may occur (Domenjoud 1992).

Condition (iii) of the inference shows that we can apply the technique of Boudet, Contejean and Marché (1996) to prune the set of AC-unifiers.

Without experiments it is not clear how these problems affect the performance of theorem proving, and in particular how they compare to the problems more general calculi have with the axioms. However, these problems are caused by free variables in sums, which can be eliminated in some cases (see Section 9.3).

9.2 Lifting by constraints

There are many advantages of using constraints to lift our calculi. Constraints avoid the double exponential number of AC-unifiers, it suffices to check AC-unifiability, which is NP-complete (Kapur and Narendran 1992). The technique for proving refutational completeness of constrained calculi is via reduced ground instances of clauses, which implies the basic restriction (Bachmair, Ganzinger, Lynch and Snyder 1995, Nieuwenhuis and Rubio 1995). That is, no superpositions into the substitution part are needed, and the information about the substitution is inherited via the constraint. Furthermore, it is possible to keep the information which term has been considered maximal in an inference in the constraint of the conclusion, in order to restrict further inferences to those which are consistent with this assumption. This is important especially in the case of free variables in sums, where otherwise the splitting of a free variable into a maximal and a nonmaximal part can be repeated without limit. The drawback of constrained calculi is that the need to keep track of reducedness complicates the completeness proof further. Therefore we have chosen not to do it here. Previously we have carried this out for the special case of modules (Stuber 1996, Stuber 1998a). From this experience we are confident that it also works for the calculi here, in a straightforward way.

For the case of abelian groups the constraints contain only AC-unification problems and ordering restriction, which can be solved by the method of Comon, Nieuwenhuis and Rubio (1995). For modules and algebras it is not clear how to handle the constraints for the computation in the base rings, as these contain both addition and multiplication and do not fall into the scope of decision procedures for Presburger arithmetic. It is however at least possible to give a semi-decision procedure that interleaves constraint-solving with the computation of inferences.

9.3 Elimination of free variables

As we have seen variables occurring in certain contexts give rise to a particularly huge number of inferences. The most problematic case is that of variables in top positions, like x in $x + p \approx q$, where x can contain the maximal term. This happens only if the variable is not *shielded*, that is it does not occur below a free function symbol somewhere else in C. In this case inferences below x are necessary.

We use modules to demonstrate some techniques to remove these free variables. For modules over the integers also variables immediately below $*$ are problematic, say x in some subterm $v * x$. Any productive equation $b * \alpha \approx r$ where $b \geq 2$ gives rise to a superposition inference with such a subterm. In this case there are also many inferences with M, in particular with distributivity, which replaces $v*x$ by $v*y+v*z$, and with (M.6), which replaces $v * x$ by $v' * x$ and adds a constraint $v' = vv''$.

We now investigate situations where these problems can be avoided or at least alleviated somewhat. Let us first consider the general case for unshielded variables at the top. We try to eliminate these variables by simplification. As an example consider the clause

$$4 * x \not\approx c_1 \vee 6 * x \not\approx c_2 \vee 3 * x \approx c_3.$$

Under the assumption that x is the maximal term, it can be simplified to

$$2 * x \not\approx (-1) * c_1 + c_2 \ \vee\ 3 * c_1 \not\approx 2 * c_2 \ \vee\ x \approx c_1 + (-1) * c_2 + c_3.$$

In general there remains at most one negative literal where the coefficient b on x is the greatest common divisor of the coefficients of the negative literals in the original clause (cf. Section 7.6). It can be used to reduce all coefficients on x in positive literals, which thus become smaller than b. If the GCD is 1, x can be eliminated completely.[1] Since x need not be maximal, one has to do a case split with respect to x being maximal or not, which can be represented by suitable constraints. Note that we cannot simplify clauses where x occurs only in positive literals in this way; take for instance $2 * x \approx c_1 \vee 3 * x \approx c_2$.

One can carry this further for modules over fields, since there each equation $b * \alpha \approx r$ can be simplified to a normal form $\alpha \approx r'$. In this case any negative literal $b * x \not\approx r$ can be used to eliminate x from a clause completely.

If additionally we know that all models are infinite, we can eliminate the positive part as well. Suppose we are given the clause $C = x \approx r_1 \vee \ldots \vee x \approx r_n \vee C'$, where x occurs neither in C' nor in any r_i, which is true in an infinite model I. Then any assignment of values in I to variables in C satisfies C. Given any assignment, since the model is infinite there exists some value in the model which is distinct from all the r_i under that assignment. If we assign this to x, leaving other variables unchanged, C' must be true under that assignment. Since x does not occur in C', all assignments satisfy C' and we may simplify C to C'.

Also, if all left-hand sides of rules in R_C have the form α instead of $b * \alpha$, no overlaps with subterms of the form $b * x$ need to be considered.

[1] This technique can be obtained by applying ground completion to the negated clause.

10

Conclusion

We have presented refutationally complete ground superposition calculi for abelian groups, commutative rings, modules and algebras over a fixed ring. Compared to previously known calculi they promise to improve theory reasoning by a more directed use of the theory, through the use of macro inferences, stronger ordering restrictions, and theory-specific redundancy criteria. Whether these calculi live up to their promise in practice remains to be investigated. As they have not yet been implemented, we have not had the opportunity to carry out experiments.

On the theoretical side we have gained a better understanding of techniques to build equational theories into superposition calculi. Our formalism works for the well-behaved commutative theories that we consider, and it is easy to see that it also works for the empty theory, AC, and ACU. Other extensions of AC such as ACI should also be easy. Boolean rings are very close to the algebra over the two-element field. Modules and algebras over slightly more general base rings are another possible extension. In particular, reduction rings (Buchberger 1984) are suitable for everything except to prove that isolation is a simplification. It seems feasible to find a slight strengthening of the reduction ring axioms that still covers the common examples of reduction rings. It is not clear how well theories without commutativity such as the theory of groups can be handled by our formalism. While a symmetrization exists, the lack of a strong symmetrization makes the manipulation of equational proofs more difficult, which could cause problems with simplification.

For other theories the underlying framework needs to be extended. For commutative algebras over base rings with zero divisors we cannot achieve convergence of all critical peaks required by symmetrization. A possible solution would be to handle these critical peaks outside of a symmetrization, so that they must converge before a symmetrization is added to a candidate model. Syntactically this would lead to additional inferences.

Further away are extensions that require a more general notion of term rewriting such as nonsymmetric rewriting for transitive relations (Levy and Agustí 1993, Bachmair and Ganzinger 1998b). The simplest case would be ordered abelian groups, as this would require only the unconditional case. More difficult would be ordered rings, because monotonicity of the ordering with respect to multiplication has as a side condition that the multiplier is positive. This needs a notion of symmetrization that includes conditions. The same problem also occurs for fields, where the inverse exists only for nonzero elements. This requires negative conditions, which leads even outside of the horn clause fragment.

Bibliography

Baader, F. (1997). Combination of compatible reduction orderings that are total on ground terms. In *12th Ann. IEEE Symp. on Logic in Computer Science*, Warsaw, pp. 2–13. IEEE Computer Society Press.

Baader, F. and Nipkow, T. (1998). *Term rewriting and all that.* Cambridge University Press, Cambridge, UK.

Bachmair, L. and Ganzinger, H. (1991). Perfect model semantics for logic programs with equality. In *Proc. 8th Int. Conf. on Logic Programming.* MIT Press.

Bachmair, L. and Ganzinger, H. (1994a). Associative-commutative superposition. In N. Dershowitz and N. Lindenstrauss (eds), *Proc. 4th Int. Workshop on Conditional and Typed Rewriting*, Jerusalem, LNCS 968, pp. 1–14. Springer.

Bachmair, L. and Ganzinger, H. (1994b). Buchberger's algorithm: A constraint-based completion procedure. In *Proc. 1st Int. Conf. on Constraints in Computational Logics*, Munich, Germany, LNCS 845, pp. 285–301. Springer.

Bachmair, L. and Ganzinger, H. (1994c). Rewrite-based equational theorem proving with selection and simplification. *Journal of Logic and Computation* **4**(3): 217–247.

Bachmair, L. and Ganzinger, H. (1998a). Equational reasoning in saturation-based theorem proving. In *Automated Deduction - A Basis for Applications. Volume I*, chapter 11, pp. 353–397. Kluwer, Dordrecht, The Netherlands.

Bachmair, L. and Ganzinger, H. (1998b). Ordered chaining calculi for first-order theories of transitive relations. *Journal of the ACM* **45**(6): 1007–1049.

Bachmair, L. and Plaisted, D. (1985). Termination orderings for associative-commutative rewriting systems. *Journal of Symbolic Computation* **1**: 329–349.

Bachmair, L. and Tiwari, A. (1997). D-bases for polynomial ideals over commutative noetherian rings. In H. Comon (ed.), *8th Intl. Conf. on Rewriting Techniques and Applications*, Sitges, Spain, LNCS 1103, pp. 113–127. Springer.

Bachmair, L., Ganzinger, H. and Stuber, J. (1995). Combining algebra and universal algebra in first-order theorem proving: The case of commutative rings. In *Proc. 10th Workshop on Specification of Abstract Data Types*, Santa Margherita, Italy, LNCS 906, pp. 1–29. Springer.

BACHMAIR, L., GANZINGER, H. AND WALDMANN, U. (1994). Refutational theorem proving for hierarchic first-order theories. *Applicable Algebra in Engineering, Communication and Computing* **5**(3/4): 193–212.

BACHMAIR, L., GANZINGER, H., LYNCH, C. AND SNYDER, W. (1995). Basic paramodulation. *Information and Computation* **121**(2): 172–192.

BECKER, T. AND WEISPFENNING, V. (1993). *Gröbner Bases: A Computational Approach to Commutative Algebra*. Springer, Berlin.

BEN CHERIFA, A. AND LESCANNE, P. (1987). Termination of rewriting systems by polynomial interpretation and its implementation. Rapports de Recherche 677, INRIA, Rocquencourt.

BOUDET, A., CONTEJEAN, E. AND MARCHÉ, C. (1996). AC-complete unification and its application to theorem proving. In *Proc. 7th Int. Conf. on Rewriting Techniques and Applications*, New Brunswick, NJ, USA, LNCS 1103, pp. 18–32. Springer.

BOYER, R. S. AND MOORE, J. S. (1988). Integrating decision procedures into heuristic theorem provers: A case study of linear arithmetic. In J. E. Hayes, D. Michie and J. Richards (eds), *Machine Intelligence 11*, chapter 5, pp. 83–124. Clarendon Press, Oxford.

BUCHBERGER, B. (1970). Ein algorithmisches Kriterium für die Lösbarkeit eines algebraischen Gleichungssystems. *Aequationes Mathematicae* **4**: 374–383.

BUCHBERGER, B. (1979). A criterion for detecting unnecessary reductions in the construction of Gröbner-bases. In *Proc. Int. Symp. on Symbolic and Algebraic Manipulation (EUROSAM '79)*, Marseille, LNCS 72, pp. 3–21. Springer.

BUCHBERGER, B. (1984). A critical-pair/completion algorithm for finitely generated ideals in rings. In E. Börger (ed.), *Logic and Machines: Decision Problems and Complexity*, LNCS 171, pp. 137–161. Springer.

BUCHBERGER, B. (1987). History and basic features of the critical pair/completion procedure. *Journal of Symbolic Computation* **3**: 3–38.

BUCHBERGER, B. AND LOOS, R. (1983). Algebraic simplification. In *Computer Algebra: Symbolic and Algebraic Computation*, 2nd edn, pp. 11–43. Springer.

BÜNDGEN, R. (1991). *Term Completion Versus Algebraic Completion*, PhD thesis, Fakultät für Informatik, Universität Tübingen.

BÜNDGEN, R. (1996). Buchberger's algorithm: The term rewriter's point of view. *Theoretical Computer Science* **159**: 143–190.

COMON, H., NIEUWENHUIS, R. AND RUBIO, A. (1995). Orderings, AC-theories and symbolic constraint solving. In *10th Annual IEEE Symp. on Logic in Computer Science, LICS-95*, San Diego, CA, pp. 375–385. IEEE Comp. Soc. Press.

CONTEJEAN, E. AND MARCHÉ, C. (1996). CiME: Completion modulo E. In *Proc. 7th Int. Conf. on Rewriting Techniques and Applications*, New Brunswick, NJ, USA, LNCS 1103, pp. 18–32. Springer. CiME is available at `http://www.lri.fr/~demons/cime.html`.

DEHN, M. (1911). Über unendliche diskontinuierliche Gruppen. *Mathematische Annalen* **71**: 116–144.

DELOR, C. AND PUEL, L. (1993). Extension of the associative path ordering to a chain of associative commutative symbols. In *Proc. 5th Int. Conf. on Rewriting Techniques and Applications*, LNCS 690, pp. 389–404. Springer.

DERSHOWITZ, N. (1987). Termination of rewriting. *Journal of Symbolic Computation* **3**: 69–116.

DERSHOWITZ, N. AND JOUANNAUD, J.-P. (1990). Rewrite systems. In J. van Leeuwen (ed.), *Handbook of Theoretical Computer Science: Formal Models and Semantics*, Vol. B, chapter 6, pp. 243–320. Elsevier/MIT Press.

DERSHOWITZ, N. AND MANNA, Z. (1979). Proving termination with multiset orderings. *Communications of the ACM* **22**(8): 465–476.

DOMENJOUD, E. (1992). A technical note on AC-unification. The number of minimal unifiers of the equation $\alpha x_1 + \ldots + \alpha x_p \doteq_{\mathrm{AC}} \beta y_1 + \ldots + \beta y_q$. *Journal of Automated Reasoning* **8**: 39–44.

FITTING, M. (1996). *First-Order Logic and Automated Theorem Proving*, 2nd edn. Springer.

GANZINGER, H. AND WALDMANN, U. (1996). Theorem proving in cancellative abelian monoids (extended abstract). In *13th Int. Conf. on Automated Deduction*, New Brunswick, NJ, USA, LNAI 1104, pp. 388–402. Springer.

GESER, A. (1996). An improved general path order. *Applicable Algebra in Engineering, Communication and Computing* **7**: 469–511.

GREENDLINGER, M. (1960a). Dehn's algorithm for the word problem. *Communications on Pure and Applied Mathematics* **13**: 67–83.

GREENDLINGER, M. (1960b). On Dehn's algorithm for the word and conjugacy problems with applications. *Communications on Pure and Applied Mathematics* **13**: 641–677.

JOUANNAUD, J.-P. AND KIRCHNER, H. (1986). Completion of a set of rules modulo a set of equations. *SIAM Journal on Computing* **15**(4): 1155–1194.

JOUANNAUD, J.-P. AND LESCANNE, P. (1982). On multiset orderings. *Information Processing Letters* **15**(2): 57–63.

KANDRI-RODY, A. AND KAPUR, D. (1988). Computing a Gröbner basis of a polynomial ideal over a Euclidean domain. *Journal of Symbolic Computation* **6**: 19–36.

KAPUR, D. AND NARENDRAN, P. (1992). Complexity of unification problems with associative-commutative operators. *Journal of Automated Reasoning* **9**: 261–288.

KAPUR, D. AND SIVAKUMAR, G. (1997). A total ground path ordering for proving termination of AC-rewrite systems. In *Proc. 8th Int. Conf. on Rewriting Techniques and Applications*, Sitges, Spain, LNCS 1103, pp. 142–156. Springer.

KAPUR, D., MUSSER, D. R. AND NARENDRAN, P. (1988). Only prime superpositions need be considered in the Knuth-Bendix completion procedure. *Journal of Symbolic Computation* **6**: 19–36.

KNUTH, D. E. AND BENDIX, P. B. (1970). Simple word problems in universal algebras. In J. Leech (ed.), *Computational Problems in Abstract Algebra*, pp. 263–297. Pergamon Press, Oxford.

LANG, S. (1993). *Algebra*, 3rd edn. Addison-Wesley, Reading, Mass.

LE CHENADEC, P. (1986). *Canonical Forms in Finitely Presented Algebras.* Pitman, London.

LEVY, J. AND AGUSTÍ, J. (1993). Bi-rewriting, a term rewriting technique for monotonic order relations. In *5th Int. Conf. on Rewriting Techniques and Applications*, Montreal, LNCS 690, pp. 17–31. Springer.

MARCHÉ, C. (1994). Normalised rewriting and normalised completion. In *Proc. 9th Ann. IEEE Symp. on Logic in Computer Science*, Paris, pp. 394–403. IEEE Computer Society Press.

MARCHÉ, C. (1996). Normalised rewriting: an alternative to rewriting modulo a set of equations. *Journal of Symbolic Computation* **21**: 253–288.

MOSER, G. (1997). Some remarks on transfinite E-semantic trees and superposition. In *Int. Workshop on First-Order Theorem Proving*, Castle Hagenberg, Austria. Appeared in the RISC-Linz Report Series, No. 97-50, Johannes Kepler Universität Linz (Austria), 1997.

NIEUWENHUIS, R. AND RUBIO, A. (1994). AC-superposition with constraints: no AC-unifiers needed. In *Proc. 12th Int. Conf. on Automated Deduction*, Nancy, France, LNCS 814, pp. 545–559. Springer.

NIEUWENHUIS, R. AND RUBIO, A. (1995). Theorem proving with ordering and equality constrained clauses. *Journal of Symbolic Computation* **19**: 321–351.

NIEUWENHUIS, R. AND RUBIO, A. (1997). Paramodulation with built-in AC-theories and symbolic constraints. *Journal of Symbolic Computation* **23**: 1–21.

PETERSON, G. E. AND STICKEL, M. E. (1981). Complete sets of reductions for some equational theories. *Journal of the ACM* **28**(2): 233–264.

ROBINSON, G. AND WOS, L. (1969). Paramodulation and theorem-proving in first-order theories with equality. In B. Meltzer and D. Michie (eds), *Machine Intelligence 4*, chapter 8, pp. 135–150. Edinburgh University Press, Edinburgh.

ROBINSON, J. A. (1965). A machine-oriented logic based on the resolution principle. *Journal of the ACM* **12**(1): 23–41.

RUBIO, A. (1994). *Automated Deduction with Constrained Clauses*, PhD thesis, Departament de Llenguatges i Sistemes Informàtics de la Universitat Politècnica de Catalunya, Barcelona.

RUBIO, A. (1999). A fully syntactic AC-RPO. In *Proc. 10th Int. Conf. on Rewriting Techniques and Applications (RTA-99)*, Trento, Italy, LNCS 1631, pp. 148–162. Springer.

RUBIO, A. AND NIEUWENHUIS, R. (1995). A total AC-compatible ordering based on RPO. *Theoretical Computer Science* **142**: 209–227.

SCHEJA, G. AND STORCH, U. (1994). *Lehrbuch der Algebra*, 2nd edn. B. G. Teubner, Stuttgart.

STICKEL, M. E. (1985). Automated deduction by theory resolution. *Journal of Automated Reasoning* **1**(4): 333–355.

STIFTER, S. (1987). A generalization of reduction rings. *Journal of Symbolic Computation* **4**: 351–364.

STIFTER, S. (1991). The reduction ring property is hereditary. *Journal of Algebra* **2**: 399–414.

STIFTER, S. (1993). Gröbner bases of modules over reduction rings. *Journal of Algebra* **1**: 54–63.

STUBER, J. (1996). Superposition theorem proving for abelian groups represented as integer modules. In H. Ganzinger (ed.), *Proc. 7th Int. Conf. on Rewriting Techniques and Applications*, New Brunswick, NJ, USA, LNCS 1103, pp. 33–47. Springer.

STUBER, J. (1997). Strong symmetrization, semi-compatibility of normalized rewriting and first-order theorem proving. In M. P. Bonacina and U. Furbach (eds), *Proc. Int. Workshop on First-Order Theorem Proving. RISC-Linz Report 97-50*, Castle Hagenberg, Linz, Austria, pp. 125–129. Research Institute for Symbolic Computation, Linz, Austria.

STUBER, J. (1998a). Superposition theorem proving for abelian groups represented as integer modules. *Theoretical Computer Science* **208**(1–2): 149–177.

STUBER, J. (1998b). Superposition theorem proving for commutative rings. In W. Bibel and P. H. Schmitt (eds), *Automated Deduction - A Basis for Applications. Volume III. Applications*, chapter 2, pp. 31–55. Kluwer, Dordrecht, The Netherlands.

STUBER, J. (1999). Theory path orderings. In *Proc. 10th Int. Conf. on Rewriting Techniques and Applications (RTA-99)*, Trento, Italy, LNCS 1631, pp. 148–162. Springer.

VIGNERON, L. (1994). Associative-commutative deduction with constraints. In *Proc. 12th Int. Conf. on Automated Deduction*, Nancy, France, LNCS 814, pp. 530–544. Springer.

VON ZUR GATHEN, J. AND GERHARD, J. (1999). *Modern Computer Algebra.* Cambridge University Press, Cambridge, UK.

WALDMANN, U. (1997). *Cancellative Abelian Monoids in Refutational Theorem Proving*, Dissertation, Universität des Saarlandes, Saarbrücken.

WALDMANN, U. (1998). Superposition for divisible torsion-free abelian groups. In *Proc. 15th Int. Conf. on Automated Deduction (CADE-15)*, Townsville, Australia, LNAI 1421, pp. 144–159. Springer.

WANG, T.-C. (1993). Z-module reasoning: An equality-oriented proving method with built-in ring axioms. *Journal of the ACM* pp. 558–606.

WERTZ, U. (1992). First-order theorem proving modulo equations. Technical Report MPI-I-92-216, Max-Planck-Institut für Informatik, Saarbrücken.

WINKLER, F. AND BUCHBERGER, B. (1983). A criterion for eliminating unnecessary reductions in the Knuth-Bendix algorithm. In *Proc. Coll. on Algebra, Combinatorics and Logic in Computer Science*, Győr, Hungary.

ZHANG, H. (1993). A case study of completion modulo distributivity and abelian groups. In *Proc. 5th Int. Conf. on Rewriting Techniques and Applications*, Montreal, LNCS 690, pp. 32–46. Springer.

List of Symbols

Index