An Ordered Theory Resolution Calculus

Peter Baumgartner
Universität Koblenz
Institut für Informatik
Rheinau 3-4
5400 Koblenz

Net: peter@infko.uni-koblenz.de

Abstract. *In this paper we present an ordered theory resolution calculus and prove its completeness. Theory reasoning means to relieve a calculus from explicitly drawing inferences in a given theory by special purpose inference rules (e.g. E-resolution for equality reasoning). We take advantage of orderings (e.g. simplification orderings) by disallowing to resolve upon clauses which violate certain maximality constraints; stated positively, a resolvent may only be built if all the selected literals are maximal in their clauses. By this technique the search space is drastically pruned. As an instantiation for theory reasoning we show that equality can be built in by rigid E-unification.*

Keywords: *Automated Theorem Proving, Theory Resolution.*

1. Introduction

The resolution principle ((Rob65)) is an important and well investigated calculus for automated reasoning. In this paper we will enrich the resolution calculus with a framework for ordered theory reasoning.

Theory reasoning ((Sti85)) means to relieve a calculus from explicit reasoning in some problem domain (e.g. equality, partial orders) by taking apart the domain knowledge and "building it into" the calculus by means of dedicated inference rules (e.g. paramodulation, E-resolution for equality). Theory reasoning is a very general scheme and thus has many applications, among them are the following: Reasoning with taxonomical knowledge (confer e.g. the Krypton system ((BFL83)), building in the theory of equality, building in theory-unification (e.g. AC-unification, universal unification for equational theories based on narrowing), combination of theorem provers with functional programming languages, building in arithmetic, reasoning with generalized clauses, where the literals in these clauses are conjunctions of ordinary literals, and building in the axioms of the "reachability" relation in the translation of modal logic to ordinary first order logic.

The advantages of theory reasoning compared to ordinary reasoning are the following: for the first, the theory inference system may be specially tailored for the theory to be reasoned with; thus higher efficiency can be achieved by a clever reasoner that takes advantage of the theories' properties. For the second, theory resolution steps are more “macroscopic”than ordinary resolution steps, in the sense that they may resolve upon more than two literals and so can hide a lot of computation that is not relevant for the overall proof plan. Thus proofs become shorter and are more compact, leading to better readability.

Ordering restrictions are a very effective technique to prune the huge search space coming up in the search for a proof. In our understanding of *ordered resolution* a partial ordering on the literals is used to disallow resolving two clauses which violate certain maximality constraints of the selected literals; stated positively, a resolvent may only be built if all the selected literals are maximal in their clauses.

By “ordered theory resolution” we mean the combination of both methods, i.e. imposing ordering restrictions on theory resolution. Thus we combine the advantage of theory reasoning with the advantage of ordering restrictions.

Our calculus can deal with arbitrary theories, provided that they are expressable as a universally quantified formulae, e.g. as a set of clauses. This restriction is motivated by our intended application of a Herbrand-theorem for the completeness proof; such a theorem can only be applied for universally quantified formulae. For this case we can prove the completeness of the calculus. This is our main result.

The rest of the paper is structured as follows: in section 2 we will informally describe the calculus by an example. In section 3 the calculus is formally defined, and in section 4 we prove that our calculus can immediately be instantiated to rigid E-unification. Section 5 contains the completeness proof. Section 6 deals with related work. Finally we draw some conclusions in section 7.

2. An Example

Let us compute an example to demonstrate the main ideas. Assume the taxonomical theory $\mathcal{T}$ be defined by the following clause set:

(T-1) $\forall X : boy(X) \rightarrow person(X)$	;; A boy is a person.
(T-2) $\forall X : girl(X) \rightarrow person(X)$	;; A girl is a person.
(T-3) $\forall X, Y : person(X) \wedge child(X, Y) \rightarrow person(Y)$	;; A child of a person ;; is a person.

Furthermore let $\mathcal{S}$ be the following clause set (in a logic programming style notation):

(1)	$boy(a)$	;; Some facts

(2)	$child(a, b)$	
(3)	$child(b, c)$	
(4)	$sex(a, male)$	
(5)	$sex(b, female)$	
(6)	$sex(c, male)$	
(7)	$child(X, Y) \rightarrow descendant(X, Y)$	;; Children are descendants
(8)	$child(X, Y) \wedge descendant(Y, Z) \rightarrow descendant(X, Z)$	;; Transitivity of *descendant*
(9)	$person(X) \wedge person(Y) \wedge$	
	$sex(X, Z) \wedge sex(Y, Z) \rightarrow samesex(X, Y)$	;; Same sex of persons
(10)	$descendant(X, Y) \wedge samesex(X, Y) \rightarrow$	;; "Query"

The clauses (1) – (6) in $\mathcal{S}$ are some facts about individuals a, b, c; clauses (7) and (8) define *descendant* -ship and (9) defines what it means for two persons to have the same sex. Finally, clause (10) asks whether there exists an X and an Y such that Y is an descendant of X, and both have the same sex.

$\mathcal{S}$ is unsatisfiable in the theory $\mathcal{T}$, as can be seen by instantiating X with a and Y with c in clause (10). We will now develop a formal proof by ordered theory resolution. In order to do so an ordering on the literals is needed. For this example the following ordering works fine:

- For terms we order $a \succ b \succ c$.
- For predicate symbols we order: $samesex \succ sex \succ person \succ child \succ descendant$. The ordering for predicate symbols not listed is immaterial.
- For literals with different predicate symbols the previous ordering on their predicate symbols is used (thus e.g. $child(X, Y) \succ descendant(X, Y)$).
- For literals with the same predicate symbol but different sign, the positive is greater than the negative (thus e.g. $descendant(X, Y) \succ \neg descendant(X, Y)$).
- For literals with same predicate symbol and same sign the above term ordering is used (thus e.g. $person(a) \succ person(b)$). If variables appear, comparison is undefined (thus e.g. $person(X)$ and $person(Y)$ are incomparable).

As mentioned above, only maximal literals in clauses (these are "potentially biggest" literals, i.e. literals that can be made biggest in a clause by instantiation) may be selected for ordered theory inferences. Thus by the choosen ordering the following controlled use of clauses is achieved: since $child \succ descendant$ only the $child$-literals can be selected in (7) and (8). So (infinite) resolution among clauses (7) and (8) and the *descendant* literal in (10) is avoided; in other words, (7) and (8) are "blocked" unless the *child* -literal is resolved away. But even if the *child* -literals are resolved away in (7) and (8), say for example with clause (2), no infinite resolution among the resulting clauses

$(7')$ $descendant(a, b)$
$(8')$ $descendant(b, Z) \rightarrow descendant(a, Z)$

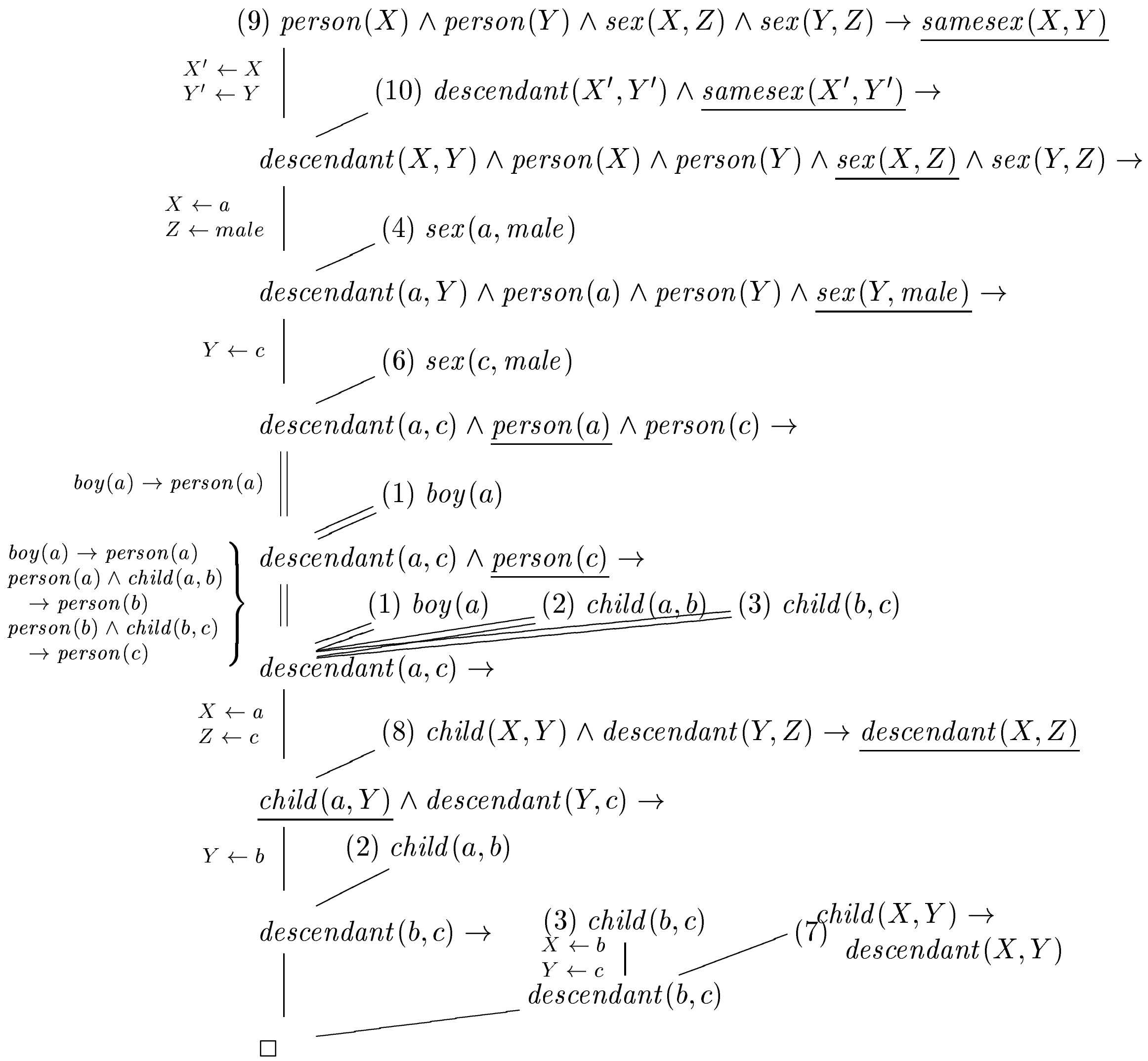
will occur, because in $(8')$ only the positive *descendant* -literal may be selected for resolution. Thus there exists no ordered theory resolvent among $(7')$ and $(8')$. This example demonstrates one of the main applications of ordering restrictions: they help to avoid infinite and redundant derivations in proofs. Note that below we will show that completeness is not affected by this restrictions.

An example for a legal theory resolution step is the following:

Ex1:
$$\frac{\begin{array}{l}\underline{boy(a)} \vee boy(c)\\ \underline{child(a,b)}\\ \underline{child(b,c)}\\ \underline{\neg person(Y)} \vee descendant(Y,Z)\end{array}}{boy(c) \vee descendant(c,Z)}$$

The selected and thus maximal literals are underlined. By resolution on the meta-level it can be seen that the conjunction of the selected literals is $\mathcal{T}$ -unsatisfiable. Thus an ordered theory-resolution step can be applied. As usual, a most general substitution (here: $\{Y \leftarrow c\}$) has to be computed, and the resolvent can then be built by an application of the substitution to the disjunction of the nonselected literals.

A complete proof of the example is as follows (the selected literals in the clauses are underlined, the double-lined inferences indicate true theory resolution, whereas the single-lined inferences indicate ordinary resolution; the labels show the used substitutions in case of ordinary resolution, or the instances of theory clauses that justify the true theory resolution steps):



3. The Calculus

3.1. Preliminaries

A *clause* is a set of literals $\{L_1, \ldots, L_n\}$, often written as $L_1 \vee \ldots \vee L_n$. Instead of $\{L\} \cup R$ we will also write $L \vee R$. A *unit clause* contains exactly one element.

As motivated in the introduction we take apart the knowledge of the domain (i.e. the *theory*) from the program clauses. More technically, the *axioms of the theory* (or simply the *theory*) $\mathcal{T}$ is a satisfiable set of clauses.

Concerning model theory it is sufficient to consider Herbrand-interpretations only, which assign a fixed meaning to all language elements short of atoms; thus we define a *(Herbrand-) interpretation* to be any total function from the set of ground atoms to $\{true, false\}$. Let $\mathcal{T}$ be a theory. A *(Herbrand-) $\mathcal{T}$-interpretation* is an interpretation satisfying the theory $\mathcal{T}$. Since we are dealing with $\mathcal{T}$-interpretations only, the prefix $\mathcal{T}$- can unambigiously be omitted in the sequel.

A clause set M is *satisfiable* iff there exists an interpretation that simultaneously assigns *true* to all ground instances of its members, or else it is *unsatisfiable*.

3.2. Orderings

Next we will introduce orderings, which our inference rules below will take advantage of.

DEFINITION 3.1
(ORDERING) Let $\succeq$ be a partial ordering on terms and let $\succ$ denote the strict subset of $\succeq$. Let $\succ$ satisfy the following conditions, where $(X, Y) \in$ Term $\times$ Term or $(X, Y) \in$ Literal $\times$ Literal:

1. $\succ$ is stable, i.e. for all substitutions σ: if $X \succ Y$ then $X\sigma \succ Y\sigma$.
2. $\succ$ is total on ground terms and $\succ$ is total on ground literals.

As usual we define $X \preceq Y$ iff $Y \succeq X$ and $X \prec Y$ iff $Y \succ X$. Let M be a literal set. A literal $L \in M$ is the *biggest literal in* M iff for all $L' \in M, L' \neq L$ it holds that $L' \prec L$. $L \in M$ is *maximal in* M iff for all $L' \in M$ it holds that $L \not\prec L'$ (or, equivalently, iff there does not exist a $L' \in M$ s.t. $L \prec L'$). $max(M)$ denotes the set of all maximal literals of M. □

An example for such an ordering are the extensions from terms to literals of the well-known simplification orderings (e.g. recursive path orderings, lexicographic path ordering) which are mainly used in the term rewriting paradigm. But note that we do not require the ordering to be noetherian. This is not required since in contrast to term rewriting calculi in our calculus no chains are built at all. Note also that in the ordering of literals we make no assumption about the treatment of the negation sign. Hence, if A is an atom we may allow to compare $A \prec \neg A$ or $\neg A \prec A$, whatever seems more appropriate for the application.
See (Der87) for an overview about orderings.

Examples: 1. The literal $person(father(x))$ is both, the biggest and a maximal element in
$\{person(father(x)),\ person(x)\}$, while $person(x)$ is neither the biggest nor maximal.
2. $\{child(x, y),\ child(y, x)\}$ has no biggest element, since with $\sigma_1 = \{y \leftarrow father(x)\}$ we have $child(x, y)\sigma_1 \prec child(y, x)\sigma_1$ and with $\sigma_2 = \{x \leftarrow father(y)\}$ we have

$child(y, x)\sigma_2 \prec child(x, y)\sigma_2$ (if the arguments of *child* are lexicographically ordered). However, both elements are maximal.

3.3. Substitutions

As with non-theory calculi the refutations should be computed at a most general level; this is usually achieved by most general unifiers. In the presence of theories however, unifiers need not be unique, and they are replaced by the more general and dual concept of *theory refuting substitutions*.

DEFINITION 3.2
(THEORY REFUTING SUBSTITUTION) Let $\mathcal{L}$ be a literal set. $\mathcal{L}$ is *$\mathcal{T}$-complementary*[1] iff for all ground substitutions γ the set $\mathcal{L}\gamma$ is $\mathcal{T}$-unsatisfiable[2]. $\mathcal{L}$ is *minimal $\mathcal{T}$-complementary* iff $\mathcal{L}$ is $\mathcal{T}$-complementary and all subsets $\mathcal{L}' \subset \mathcal{L}$ are not $\mathcal{T}$-complementary.

We say that $\mathcal{L}$ is (minimal) $\mathcal{T}$-refutable by σ iff $\mathcal{L}\sigma$ is (minimal) $\mathcal{T}$-complementary.

A set of substitutions is a *complete and most general set of $\mathcal{T}$-refuting substitutions for $\mathcal{L}$* (or short: $CSR_{\mathcal{T}}(\mathcal{L})$) iff

1. for all $\sigma \in CSR_{\mathcal{T}}(\mathcal{L})$: $\mathcal{L}$ is $\mathcal{T}$-refutable by σ *(Correctness)*
2. for all substitutions θ such that $\mathcal{L}$ is $\mathcal{T}$-refutable by θ:
   there exists a $\sigma \in CSR_{\mathcal{T}}(\mathcal{L})$ and a substitution σ' such that $\theta = \sigma\sigma'|var(\theta)$
   (Completeness)

The members of $CSR(\mathcal{L})$ are also called *most general $\mathcal{T}-$refuters (MGR) for $\mathcal{L}$*. The prefix $\mathcal{T}$- is often omitted in the sequel. □

Example: Assume the theory consists solely of (T-1) from the example in the introduction: $\mathcal{T} = \{\forall x : boy(x) \rightarrow person(x)\}$. Consider the set
$S = \{boy(x),\ \neg person(father(y))\}$. S is $\mathcal{T}$-refutable by MGR $\sigma = \{x \leftarrow father(y)\}$ because every ground instance of $S\sigma = \{boy(father(y)),\ \neg person(father(y))\}$ is unsatisfiable in the theory (T-1). S is even minimally refutable by σ, as any true subset of $S\sigma$ can be ground instantiated to an (T-1)-satisfiable set. The substitution $\gamma = \{x \leftarrow father(z)\}$ is not a refuting substitution for S, because $S\gamma = \{boy(father(z)),\ \neg person(father(y))\}$ is not complementary. This can be seen by applying, say, $\theta = \{z \leftarrow a, y \leftarrow b\}$ to $S\gamma$ and finding a model.

[1] this definition is intended as a generalization of standard "syntactically complementary" which means that two literals are syntactic complementary iff one of them is the negation of the other.
[2] $\mathcal{L}\sigma$ is the set that results from applying σ to the elements of $\mathcal{L}$

3.4. The inference rules

Next we will apply the previous concepts of orderings and theory refuting substitutions in the inference rules of our calculus.

DEFINITION 3.3
(OTR-CALCULUS) Let $\mathcal{T}$ be a theory. The inference rules of the ordered theory resolution calculus (OTR-Resolution) are defined as follows:

Ordered Factoring:

$$\frac{C}{C\sigma} \quad \begin{cases} \text{if (1) } \sigma \text{ is a most general (syntactical) unifier for some } \{L_1, \ldots, L_n\} \subseteq C, \\ \text{and (2) } L_1\sigma \text{ is maximal in } C\sigma \end{cases}$$

Ordered theory resolution:

$$\frac{C_1 \quad \ldots \quad C_n}{(C_1\sigma - \{L_1\sigma\}) \cup \ldots \cup (C_n\sigma - \{L_n\sigma\})} \quad \begin{cases} \text{if (1) } \sigma \in CSR_{\mathcal{T}}(\{L_1, \ldots, L_n\}) \text{ for some } L_1 \in C_1, \ldots, L_n \in C_n, \\ \text{and (2) } L_i\sigma \text{ is maximal in } C_i\sigma \text{ (for } i = 1 \ldots n) \end{cases}$$

The inference rules of the ordered theory resolution calculus.

In these inference rules, the L_i are called the *selected literals*. Let M be a clause set. An $\mathrm{OTR}(\mathcal{T})$-derivation of C_n from M is a sequence $C_1, \ldots, C_n$ where each $C_i \in M$ or is obtained by an application of the above inference rules to k variable disjoint copies of clauses $C_{j_1} \ldots C_{j_k}$ where $j_1 < i, \ldots, j_k < i$. A *ground derivation* is a derivation where every clause is ground. A *refutation of* M is a derivation of the empty clause $\Box$ from M.
$\Box$

If two literals are syntactically complementary then they are $\mathcal{T}$-complementary in any theory $\mathcal{T}$. Hence the ordered resolution inference rule subsumes the well-known standard resolution rule (modulo ordering).

As an example for an ordered theory resolution inference see the example Ex1 in section 2 again; in the same section also a complete refutation can be found.

A standard problem in resolution calculi is the question whether tautologies, i.e. clauses of the form $A \vee \neg A \vee R$ are neccessary for refutational completeness. In standard non-theory resolution, tautologies may savely be deleted. The following example shows that this is *not* the case for theory resolution: let the clause set be

$$M_1 = \{\underline{A} \vee B,\ \underline{\neg A} \vee \neg B\}$$

and let the theory state that "A is logically equivalent to B". Assume an ordering such that the underlined literals are maximal in their clauses. M_1 is theory-unsatisfiable, and although there exists a refutation, the only ordered theory resolvent of the clauses in M_1 is the clause $B \vee \neg B$, which is a tautology.

4. Treating Equality by Rigid E-unification

In (GNPS90) a first order calculus with equality is defined. The base calculus is Andrews method of matings ((And81)), and equality is treated by a device called *rigid E-unification*. The base calculus is not of crucial importance here, but the treatment of equality is, since the results obtained by these authors are immediately applicable to our calculus when instantiating the theory to equality.

In the mentioned calculus, inferences are carried out *modulo an equational theory*. More precisely, instead of computing the well-known most general unifier, the key concept of *rigid E-unifier* is used ((GNPS90), Problem 2):

> Given a finite set $E = \{u_1 = v_1, \ldots, u_n = v_n\}$ of equations and a pair $\langle u, v\rangle$ of terms, is there a substitution σ such that, treating $E\sigma$ as a set of ground equations, $u\sigma \overset{*}{\cong}_{E\sigma} v\sigma$, that is, $u\sigma$ and $v\sigma$ are congruent modulo $E\sigma$ (by congruence closure)?
>
> The substitution σ is called a *rigid E-unifier of* u *and* v.

Most exciting, the authors show that rigid E-unification is *decidable*. This result is applicable in our calculus if we can show that rigid E-unifiers coincide with our $\mathcal{T}$-refuting substitutions (again, if the theory is equality), because then we can compute complete set of refuting substitutions with a rigid E-unification algorithm.

In order to compare concepts, the following observation is helpful: σ is a rigid E-unifier of $\langle u, v\rangle$ wrt. E iff $E\sigma \cup \{\neg u\sigma = v\sigma\}$ is E-unsatisfiable, when all variables are treated as constants.[3] This reformulation will be the starting point for the comparison to rigid E-unification. More formally we arrive at the following proposition.

[3] In the following argumentation variables are treated as constants: for the only-if direction note that $u\sigma \overset{*}{\cong}_{E\sigma} v\sigma$ implies by reflexivity of equality that $\neg u\sigma = v\sigma$ is false in all E-interpretations; for the if-direction recognize that equality can be axiomatized in definite logic, and thus a single "query" $\neg u\sigma = v\sigma$ suffices for E-unsatisfiability; thus $u\sigma = v\sigma$ is a logical consequence of a set of positive equations $E\sigma$. Then by Birckhoff's completeness theorem $u\sigma \overset{*}{\cong}_{E\sigma} v\sigma$

PROPOSITION 4.1
Let M be a literal set and σ be a substitution. Then M is E-refutable by σ iff $M\sigma$ is E-unsatisfiable, where the variables of $M\sigma$ are treated as constants.

PROOF. Let X be the set of variables of $M\sigma$.
(Only if) We prove the contraposition. Thus let $M\sigma$ be not E-unsatisfiable; hence $M\sigma$ is E-satisfiable. Let I^X be a model for $M\sigma$ where the variables of $M\sigma$ are treated as constants. Define $\theta := \{x \leftarrow \mathrm{sk}^x | x \in X\}$ where sk^x is drawn from a X-indexed set of new constant symbols. As a consequence of this definition $M\sigma\theta$ is ground. Let I be the partial interpretation that is equal to I^x but is undefined for (the constants) X. Define I^{sk} as the interpretation that extends I with the assignments $I^{\mathrm{sk}}(\mathrm{sk}^x) = I^X(x)$ for all $x \in X$. By structural induction we see that $I^{\mathrm{sk}}(L\sigma\theta) = I^X(L\sigma)$ for all $L \in M$. With I^X being a model for $M\sigma$ we have thus found a ground substitution θ s.t. $M\sigma\theta$ is E-satisfiable. So $M\sigma$ is not E-refutable, and the contraposition is proved.

(If) We prove the contraposition. Thus assume that by some grounding substitution θ $M\sigma\theta$ is E-satisfiable. Let I be a (Herbrand-) model. Let I^X be the interpretation that extends I by $I^X(x) = x\theta \quad (\text{for all } x \in X)$ where x is treated as a constant in I^X. By structural induction we see that $I(L\sigma\theta) = I^X(L\sigma)$ for all $L \in M$. With I being a model for $M\sigma\theta$ we have thus found a model I^X for $M\sigma$ where the variables are treated as constants. Hence the contraposition is proved. Q.E.D.

For our purpose the main application of the equivalence results in this section is to build in rigid E-unification into the resolution calculus. To the best of our knowledge this is an original result. We conclude this discussion with the note that rigid E-unification is NP-complete.

5. Soundness and completeness

Soundness can be stated as follows:

THEOREM 5.1
(SOUNDNESS) *Let $\mathcal{T}$ be a theory and M be a clause set. If there exists an $\mathrm{OTR}(\mathcal{T})$ refutation of M then M is $\mathcal{T}$-unsatisfiable.*

It is much more difficult to prove the *completeness*, which shall be done next.

There is a canonical way for completenes proofs of first order calculi: first show the desired result for the ground case, and then apply a lifting lemma to show that the ground refutation can also be carried out with variables. We will also follow this strategy.

The proof technique for the ground case is interesting of its own; it is a generalization of the technique based on the "excess literal parameter" ((AB70)). Informally, the excess literal parameter is a measure for the complexity of clause sets, and sets consisting of unit clauses only have the lowest complexity. Now, to show completeness of a calculus one has to split an unsatisfiable clause set into unsatisfiable sets of lower complexity, and "assemble" the existsing refutations of these split sets into a refutation of the original set. However in this process more care need to be taken in our case than in the original unordered case: in the unordered case the splitting may be carried out on *any* non-unit clause and on *any* literal in that clause; this does not work in the ordered case. We have to select for splitting that clause that contains the *smallest* literal wrt. all literals occuring in non-unit-clauses.

In the following let $min(X)$ denote the smallest literal wrt. $\prec$ occuring in a clause or clause set X. By ground totality of $\prec$ (definition 3.1) such a literal always exist.

LEMMA 5.2
Let M be a ground clause set and $L \in M$ be a ground unit clause. Suppose that L does not occur in a non-unit clause in M. Then there does not exist a derivation of a non-unit clause $L \vee R$.

PROOF. Assume, to the contrary that that there exists a derivation of some non-unit-clause $L \vee R$. Let

$$D = C_1, \ldots, C_k, L \vee R$$

be a shortest derivation (i.e. with minimal index k). Since L does not occur in a non-unit clause in M, $L \vee R$ cannot be in M. Hence $L \vee R$ is derived in an (un-)ordered theory resolution step. The last inference in D is of the form

$$C_{i_1}, \ldots, C_{i_n} \vdash L \vee R$$

L must occur in one of the clauses $C_{i_1}, \ldots, C_{i_n}$. Let B be such a clause. B must contain at least *two* literals, the one that is resolved upon in the inference, and L. So B is a non-unit clause and containing L. However B is obtained in a shorter derivation than $L \vee R$ in D. Contradiction to the assumption that D is the shortest derivation. Q.E.D.

LEMMA 5.3
Let M be a ground clause set and let $R \in M$. Let L be a literal such that (1) $L \prec min(R)$ and such that (2) L is $\preceq$ than all literals occuring in non-unit clauses in M. Let $M' = M - \{R\} \cup \{L \vee R\}$. Suppose that there exists an ordered derivation of some literal C_k from M.
Then there exists an ordered derivation of ${C_k}'$ from M' where ${C_k}' = C_k$ or ${C_k}' = C_k \vee L$. Furthermore, L is the smallest literal in every clause that contains it in that derivation (3).

PROOF. Let

$$D = C_1, \ldots, C_k$$

be the given derivation. By induction on k we will construct the desired derivation D'. If $k = 0$ then the lemma holds immediately by setting $D' = D$, which is the empty derivation. Otherwise, if $k > 0$ we distinguish 3 disjoint cases.

1. $C_k = R$. Let $D' = L \vee R$. By (1) above L is the smallest literal in the single clause in that refutation and thus (3) holds.

2. $C_k \neq R$ but $C_k \in M$. In this case $C_k \in M'$. So $D' = C_k$ is the desired derivation; If C_k does not contain L then (3) holds immediately, or else (3) follows from (2).

3. $C_k \notin M$. C_k must be obtained by an ordered theory resolution step. Suppose that C_k is inferred from the clauses $C_{i_1}, \ldots, C_{i_k}$. By definition of derivation $i_j < k$ (for all $j = 1 \ldots k$) and hence by the induction hypotheses there exist k ordered derivations $D_1{}', \ldots, D_k{}'$ of clauses $C_{i_1}{}', \ldots, C_{i_k}{}'$ of M', where

   $$C_{i_j}{}' = C_{i_j} \quad \text{or} \quad C_{i_j}{}' = L \vee C_{i_j}$$

   Since (3) holds for these derivations the maximal literal in $C_{i_j}{}'$ is the same as the maximal literal in C_{i_j}. Hence $C_k{}'$ can be obtained in an ordered theory resolution from the clauses $C_{i_1}{}', \ldots, C_{i_k}{}'$, where $C_k{}' = C_k \quad$ or $\quad C_k{}' = L \vee C_k$. Thus by concatenating $D_1{}', \ldots, D_1{}'$ and $C_k{}'$ we obtain an ordered derivation D' of $C_k{}'$

   It remains to show that (3) holds for $C_k{}'$. If no $C_{i_j}{}'$ contains L then $C_k{}'$ also does not contain L and (3) holds immediately. Otherwise, suppose to the contrary that

   $$P \prec L \text{ for some literal } P \text{ in } C_k{}' \qquad (*)$$

   Since P occurs in $C_k{}'$ it must occur also in one of the clauses $C_{i_j}{}'$. Let B be such a clause. B must contain at least *two* literals, the one that is resolved upon in the inference, and P. So B is a non-unit clause $(**)$.

   By (2) P cannot occur in a non-unit clause in M. Also P cannot occur in a non-unit clause in M' $(***)$. Proof of $(***)$: assume to the contrary that P occurs in M'. Since M' differs from M only in the clause $L \vee R$, which is in M' but not in M, P must occur in $L \vee R$ and thus in R. However, by (1) $L \prec min(R)$ which contradicts the assumption $P \prec L$ in $(*)$. So $(***)$ is proved.

   From $(***)$ it follows that P occurs as a unit-clause in M'. But then by lemma (5.2) there does not exist a derivation of B as deduced in $(**)$. From this contradiction it follows that the assumption $(*)$ is wrong. Thus (3) is proved.

Q.E.D.

LEMMA 5.4
(GROUND COMPLETENESS) *Let $\mathcal{T}$ be a theory and M_g be an unsatisfiable ground clause set. Furthermore suppose a complete inference system for $\mathcal{T}$ as given. Then there exists a ground OTR-refutation of M_g.*

PROOF. Since we deal with ground clauses here, the notions of "complementary" and "unsatisfiable" are equivalent and will be used interchangeable in the proof.

Let M be a literal set. Then $k(M)$ denotes the number of occurences of literals in M minus the number of clauses in M ($k(M)$ is called the *excess literal parameter* in ((AB70))). Now we prove the claim by induction on $k(M)$.

1. $\boxed{k(M) = 0}$: M must be a set of unit clauses

   $$M = \{L_1, \ldots, L_n\}$$

   Since M is unsatisfiable a (ground) *ordered theory resolution* step can be applied to $L_1, \ldots, L_n$. This step results in the empty clause. Hence we have found a refutation

   $$L_1, \ldots, L_n, \Box$$

   for M.

2. $\boxed{k(M) > 0}$: Suppose that the result holds for unsatisfiable ground clause sets M' such that $k(M') < k(M)$. Since $k(M) > 0$, M contains at least one non-unit clause. Let L_{min} be the smallest literal wrt. $\prec$ of all literals occuring in non-unit clauses in M. Hence there exists a clause $C = L_{min} \vee R$ where L is a literal and R is a non-empty clause. Now consider

   $$\begin{aligned} &M_L = M - \{C\} \cup \{L_{min}\}, \text{ and} \\ &M_R = M - \{C\} \cup \{R\} \end{aligned}$$

   Both M_L and M_R are unsatisfiable, since otherwise a model for one of them were a model for M, which contradicts the assumption that M is unsatisfiable. Since $k(M_L) < k(M)$ we can apply the induction hypothesis and obtain an ordered refutation of M_L.

   Now Consider M_R. Since $k(M_R) < k(M)$ we can apply the induction hypothesis again and obtain an ordered refutation of M_R. Since $L_{min} \prec min(R)$ we can apply lemma (5.3) and obtain either an ordered refutation of $M_R - \{R\} \cup \{L_{min} \vee R\} = M$ or an ordered derivation of L_{min}. In the first case we have immediately found the desired refutation; in the second case we append to that derivation the above refutation of M_L and thus obtain an ordered refutation of M.

Q.E.D.

Next we turn to lifting. As a preliminary we need the following lemma that states that the notion of maximality can be lifted from instances to more general terms.

LEMMA 5.5
Let S be a literal set, L be a literal in S, σ and δ be substitutions such that $\sigma \leq \delta$. If $L\delta \in max(S\delta)$ then $L\sigma \in max(S\sigma)$.

PROOF.

$$\begin{aligned}
& & & L\delta \in max(S\delta) \\
&(\text{By def. of } max\ ) &\Longleftrightarrow\ & \forall L' \in S\delta : L\delta \not\prec L' \\
&(\delta = \sigma\sigma') &\Longleftrightarrow\ & \forall L' \in (S\sigma)\sigma' : (L\sigma)\sigma' \not\prec L' \\
&(\text{Contrapos. of stability of } \succ) &\Longrightarrow\ & \forall L'' \in S\sigma : L\sigma \not\prec L'' \\
&(\text{By def. of } max\ ) &\Longleftrightarrow\ & L\sigma \in max(S\sigma).
\end{aligned}$$

Q.E.D.

LEMMA 5.6

(LIFTING LEMMA) *Suppose θ is a substitution and $C\theta$ is an ordered theory resolvent of some clauses $C_1\theta \ldots C_n\theta$. Then there exists a derivation of a clause C' from $C_1 \ldots C_n$, obtained from zero or one application of an ordered factoring step, followed by a single application of an ordered theory resolution step such that $C\theta$ is an instance of C'.*

PROOF. In the given ordered theory resolution step, every $C_i\theta$ ($i = 1 \ldots n$) takes the form:

$$C_i\theta = L_i\theta \vee R_i\theta$$

where the $L_i\theta$ are the selected literals. For every clause C_i, θ is a unifier for k_i ($i \geq 1$) literals

$$F_i = \{L_{i,1}, \ldots, L_{i,k_i}\}$$

Now let γ_i be a most general unifier for F_i, i.e. $L_{i,1}\gamma_i \equiv L_{i,k_i}\gamma_i$. We may assume that γ_i introduces no new variables to F_i. $L_i\theta$ is the selected literal in $C_i\theta$ and thus is maximal. By lemma (5.5) L_i is maximal in C_i. Thus $C_i\gamma$ can be derived from C_i by an application of an ordered factoring step.

Let

$$\gamma = \gamma_1 \ldots \gamma_n$$

Since all C_i are variable disjoint, and γ_i introduces no new variables it follows that

$$C_i\gamma_i = C_i\gamma = L_{i,1}\gamma \vee R_i\gamma$$

Since γ_i is most general, $L_{i,1}\theta$ is an instance of $L_{i,1}\gamma$, say by δ_i:

$$L_{i,1}\theta \equiv L_{i,1}\gamma\delta_i$$

Clearly we may assume that δ_i acts only on variables in $L_{i,1}\gamma$. Furthermore, since all $C_i\gamma$ are variable disjoint we may build

$$\delta = \delta_1 \ldots \delta_n$$

and obtain

$$C_i\theta = (C_i\gamma)\delta_i = (C_i\gamma)\delta$$

In the given resolution step the selected literals may be written as

$$\{(L_{1,1}\gamma)\delta, \ldots (L_{n,1}\gamma)\delta\}$$

By definition of ordered theory resolution this set is $\mathcal{T}$-complementary. Hence δ is a $\mathcal{T}$-refuting substitution for $\{L_{1,1}\gamma, \ldots L_{n,1}\gamma\}$. Furthermore $(L_{i,1}\gamma)\delta$ is maximal in $(C_i\gamma)\delta$ (*). By the completeness property in the definition of complete set of refuters (def. 3.2) there exists also a more general substitution $\sigma \leq \delta$ such that

$$\{(L_{1,1}\gamma)\sigma, \ldots (L_{n,1}\gamma)\sigma\}$$

is $\mathcal{T}$ -complementary. By lemma (5.5) it follows from (*) that $(L_{i,1}\gamma)\sigma$ is maximal in $(C_i\gamma)\sigma$. Thus we can apply an ordered theory resolution step to $\{C_1\gamma, \ldots, C_n\gamma\}$ with selected literals $L_{1,1}\gamma, \ldots, L_{n,1}\gamma$ yielding the resolvent

$$C' = (R_1\gamma)\sigma \vee \ldots \vee (R_n\gamma)\sigma$$

It remains to show that the given resolvent

$$C\theta = R_1\theta \vee \ldots \vee R_n\theta$$

can be obtained from C' by instantiation. This however follows immediately from the completeness property of most general refuters again, since by that property there exists a substitution σ' such that $\delta = \sigma\sigma'$. Thus

$$\begin{aligned}
&C\theta = \\
&R_1\theta \vee \ldots \vee R_n\theta = \\
&(R_1\gamma)\delta \vee \ldots \vee (R_n\gamma)\delta = \\
&(R_1\gamma)\sigma\sigma' \vee \ldots \vee (R_n\gamma)\sigma\sigma' = \\
&(R_1\gamma\sigma)\sigma' \vee \ldots \vee (R_n\gamma\sigma)\sigma' = \\
&C'\sigma'
\end{aligned}$$

Q.E.D.

THEOREM 5.7
(COMPLETENESS OF ORDERED THEORY RESOLUTION) *Let $\mathcal{T}$ be a theory and M be a $\mathcal{T}$-unsatisfiable clause set.Then there exists an $OTR(\mathcal{T})$-refutation of M.*

PROOF. The proof employs an adapted version of the Skolem-Herbrand-Gödel theorem for theory reasoning. In its basic version the theorem states that a clause set M is unsatisfiable iff there exists a finite set M_g of ground instances of clauses from M which are unsatisfiable; for our purpose however we need the claim for $\mathcal{T}$-unsatisfiability. But the theorem holds for this case too, as can be seen by adding the axioms of the theory $\mathcal{T}$ as first-order clauses to S and applying the basic version.

Thus suppose that M_g is a finite unsatisfiable set of ground instances of clauses from M. By the ground completeness (lemma 5.4) M_g has a refutation. By induction on the length of the refutation and applying the lifting lemma in each step this proof can be carried out on the variable level, using most general $\mathcal{T}$ -refuting substitutions. Q.E.D.

6. Related Work

Related work comes from two sources: the one is *ordering restrictions* and the other is *theory reasoning*. We will discuss both of them.

Early ordered resolution approaches (but not *theory* resolution) are described in ((CL73)). There, in *semantic resolution* the ordering is carried out on the predicate symbols only. As a drawback of comparing the predicate symbols only the restriction is not as effective as could be when “looking inside the literals”. In the same book, OI-resolution is described. It avoids that drawback, but is incomplete. In another approach (OL-Resolution, which is similar to model elimination) the idea of selecting maximal literals only can be imposed on *one* parent clause, but not on both parent clauses.

Recent work in equational reasoning is mainly based on rewriting techniques and the superposition inference rule, which is an order restricted specialization of paramodulation. Ordered inference systems for first order logic with equality were proposed by e.g. ((BG90; HR86; ZK88)). In our calculus, an equationally unsatisfiable literal set is searched by selecting one literal from *multiple* clauses. These whole set is resolved away in the inference step. In contrast to that, the superposition-based calculi “simulate” our inference step by a sequence of more fine-grained superposition steps.

Now we turn to the related work in theory reasoning. Theory reasoning was introduced by M. Stickel within the general, non-linear resolution calculus ((Sti85; Sti83)). There, one main inference rule is called *narrow theory resolution*, which resolves upon a conjunction of theory *literals*. There exists also a variant called *wide theory resolution* which resolves upon a conjunction of *clauses*. Using Stickel's terminology, our theory inference rule is narrow theory resolution.

Our work distinguishes from Stickels in several aspects: for the first, we have lifted our inference rules to full first order logic, while the original work defines a ground calculus only; for the second, and more important, our calculus is fully ordered.

Since Stickel's pioneering work, the scheme was ported to many calculi. It was done for matrix methods ((MR87)), for the connection method ((Pet90)), for connection graphs ((OS91)) and for model elimination ((Bau91)). However, no ordering restrictions are applied in these calculi.

7. Conclusions

In the preceeding text we have presented a resolution calculus for ordered theory reasoning and proved its completeness. Furthermore we showed that theory reasoning can be instantiated to rigid E-unification.

Further work should be done on crucial notions in theorem proving such as “subsumption” and “simplification”. In practice, the inference steps may become “too large” for certain theories due to long computations by the theory reasoner. It may turn out to be more appropriate to simulate a theory resolution step as defined in the text by some “smaller” inference steps. For example, one might say that a rigid E-unification step can be simulated by a sequence of paramodulation steps. For that purpose we are currently working on variant of the calculus that includes a *partial* ordered theory resolution rule.

Acknowledgements: I want to thank U. Furbach and W. Hower for critical comments.

References

(AB70) R. Anderson and W. Bledsoe. A linear format for resolution with merging and a new technique for establishing completeness. *J. of the ACM*, 17:525–534, 1970.

(And81) P. Andrews. Theorem Proving via General Matings. *J.ACM*, 28(2):193–214, 1981.

(Bau91) P. Baumgartner. A Model Elimination Calculus with Built-in Theories. Fachbericht Informatik 7/91, Universität Koblenz, 1991. (Submitted to CADE 11).

(BFL83) R. Brachmann, R. Fikes, and H. Levesque. KRYPTON: a functional approach to knowledge representation. *IEEE Computer*, 16(10):67–73, October 1983.

(BG90) L. Bachmair and H. Ganzinger. Completion of First-Order Clauses with Equality by Strict Superposition. In *Proc. Second Int. Workshop on Conditional and Typed Rewrite Systems, LNCS*. Springer, 1990.

(CL73) C. Chang and R. Lee. *Symbolic Logic and Mechanical Theorem Proving*. Academic Press, 1973.

(Der87) Nachum Dershowitz. Termination of Rewriting. *Journal of Symbolic Computation*, 3(1&2):69–116, February/April 1987.

(GNPS90) J. Gallier, P. Narendran, D. Plaisted, and W. Snyder. Rigid E-unification: NP-Completeness and Applications to Equational Matings. *Information and Computation*, pages 129–195, 1990.

(HR86) J. Hsiang and M. Rusinowitch. A New Method for Establishing Refutational Completeness in Theorem Proving. In *Proc. 8th CADE*, pages 141–152. Springer, 1986.

(MR87) N. Murray and E. Rosenthal. Theory Links: Applications to Automated Theorem Proving. *J. of Symbolic Computation*, 4:173–190, 1987.

(OS91) H.J. Ohlbach and J. Siekmann. The Markgraf Karl Refutation Procedure. In J.L. Lassez and G. Plotkin, editors, *Computational Logic — Essays in Honor of Alan Robinson*, pages 41–112. MIT Press, 1991.

(Pet90) U. Petermann. Towards a connection procedure with built in theories. In *JELIA 90*. European Workshop on Logic in AI, Springer, LNCS, 1990.

(Rob65) J.A. Robinson. A machine-oriented logic based on the resolution principle. *JACM*, 12(1):23–41, January 1965.

(Sti83) M.E. Stickel. Theory Resolution: Building in Nonequational Theories. SRI International Research Report Technical Note 286, Artificial Intelligence Center, 1983.

(Sti85) M. E. Stickel. Automated deduction by theory resolution. *Journal of Automated Reasoning*, pages 333–356, 1985.

(ZK88) H. Zhang and D. Kapur. First-Order Theorem Proving Using Conditional Rewrite Rules. In E. Lusk and R. Overbeek, editors, *Lecture Notes in Computer Science: 9th International Conference on Automated Deduction*, pages 1–20. Springer-Verlag, May 1988.