MAX-PLANCK-INSTITUT FÜR INFORMATIK

A Goal Oriented Strategy Based on Completion

Rolf Socher-Ambrosius

MPI–I–92–206 February 1992

Im Stadtwald
66123 Saarbrücken
Germany

A Goal Oriented Strategy Based on Completion

Rolf Socher-Ambrosius

MPI–I–92–206 February 1992

Author's Address

Rolf Socher-Ambrosius
Max–Planck–Institut für Informatik
Im Stadtwald
D-6600 Saarbrücken 11
F. R. Germany
socher@mpi-sb.mpg.de

Publication Notes

This report has been submitted for publication elsewhere

"Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministers für Forschung und Technologie (Betreuungskennzeichen ITS 9103) gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor."

A Goal Oriented Strategy Based on Completion

Rolf Socher-Ambrosius*

Max-Planck-Institut für Informatik, Im Stadtwald,
D-W-6600 Saarbrücken, Germany
email: socher@mpi-sb.mpg.de

Abstract

In this paper, a paramodulation calculus for equational reasoning is presented that combines the advantages of both Knuth-Bendix completion and goal directed strategies like the set of support strategy. Its soundness and completeness is proved, and finally the practical aspects of this method are discussed.

1 Introduction

Knuth-Bendix completion has turned out to be a well suited tool for proving equational theorems. It seems, however, that this approach to equational reasoning does not support backward reasoning or goal directed strategies. Knuth-Bendix completion is a forward chaining method, deriving new consequences from the axioms and previously derived equations, intended to produce a complete set of equations and rewrite rules for the given equational specification. The proof of the theorem actually is more like a "side effect" of completion. The well-known problem with the set of support strategy [17] in combination with a paramodulation calculus [15] is the need to paramodulate into and from variables (see, for instance, [16]), both of which is precluded in Knuth-Bendix completion. The following clause set

$$\{f(a,b)\simeq a,\ a\simeq b,\ f(x,x)\not\simeq x\}$$

which is taken from [16], is unsatisfiable. However, if the third clause is chosen as set of support, then a refutation can be obtained only by paramodulating *into* a variable. Snyder and Lynch [16] have found a way to overcome this difficulty. Their *relaxed paramodulation calculus* employs the set of support strategy, and it is complete without paramodulation into variables.

Still, the relaxed paramodulation calculus cannot solve the problem with paramodulating *from* variables, as the following example (see [18]) shows:

$$\{(x*y)*z\simeq x*(y*z),\ x*e\simeq x,\ x*i(x)\simeq e,\ i(i(a))\not\simeq a\} \qquad (1)$$

---

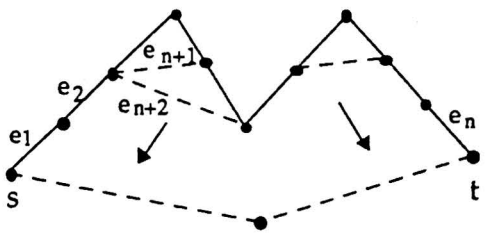
* This work was supported by a grant from the Max-Kade Foundation

If the inequality is chosen as the set of support, then no refutation can be obtained without paramodulating *from* the variable x in the second equation. Paramodulating from variables is incompatible with the most important feature of Knuth-Bendix completion, namely the orientation of equations into rules. In order to keep the advantages of completion when combining it with backward reasoning, the basic feature of orienting equations should be preserved as much as possible.

Considering the success of goal directed approaches, such as set of support (with the goal chosen as the set of support), or SLD-resolution for automated theorem proving, it would be desirable to have a similar strategy for completion based methods in equational reasoning (for a more detailed discussion of this issue, see [18]. Bonacina and Hsiang [4] pointed out that Knuth-Bendix completion, when used to prove equational theorems, does not take into account the existence of a theorem to be proved. Rather, it derives consequences from the axioms regardless whether they contribute to the proof.

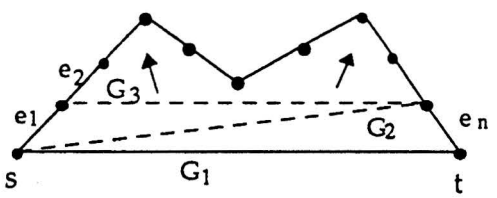
In the following, a goal oriented calculus for equational reasoning is presented that comes as close as possible to Knuth-Bendix completion. This strategy takes advantage of the additional information about the equational theorem to be proved, thus reducing the number of unnecessary consequences derived. It works in the spirit of the set of support strategy, however, relying more on the terms occurring in the goal equation, rather than on the goal itself. In fact, this is just the idea proposed in connection with research problem 3 in [18]. The performance of this method depends on the particular equation to be proved. If the information provided by the structure of the goal is too weak, then our method reduces to ordinary completion.

To get an idea of how this strategy proceeds, we compare the basic approach of completion based theorem proving (fig. 1a) with a paramodulation calculus that employs the set of support strategy (see fig. 1b) to prove an equation $s =_R t$, where $R=\{e_1,\ldots,e_n\}$. Knuth-Bendix completion successively eliminates the peaks in the proof $s =_R t$ by deriving new axioms e_k, such that the proof $s =_{R^\infty} t$ has no peak, and hence is a *rewrite proof*. Goal oriented strategies, on the other hand, start from the theorem to be proved, $G_1 = \{s \approx t\}$, and successively derive new goals G_i by paramodulating some G_j, $j<i$, with one of the axioms. It should be remarked that other goal directed strategies, like E-resolution [13], proceed basically in the same way.



Completion Based Approach

fig. 1a



Paramodulation Based Approach with Set of Support

fig. 1b

The proceeding of *goal directed completion* is illustrated in fig. 2. An equation $e \in R$ is distinguished as a *goal equation*, if it can occur as the leftmost (rightmost) step of the proof $s =_R t$, that is, if there exists a term s' (t') with $s \leftrightarrow_e s'$ ($t' \leftrightarrow_e t$, respectively). New

consequences are derived from the axioms and previously derived equations using two basic derivation rules: Superposition is used to resolve those peaks that comprise at least one goal equation. A very restricted form of paramodulation is used in order to derive new goal equations, thus decreasing the distance between the left-hand (right-hand) side of the goal and the leftmost (rightmost) peak. The basic idea is to allow only the resolution of the leftmost (rightmost) peak of the proof $s=_R t$.

As an example, consider the following system: $E = \{(x*e)*y \approx x*(e*y),\ e*h(x) \approx x,\ e*e \approx e,\ h(e) \approx d,\ a \neq h(a)\}$. The goal is the pair $(a,h(a))$. Each occurring equation can be oriented into a rule:

$(e_1)\ (x*e)*y \rightarrow x*(e*y)$

$(e_2)\ e*h(x) \rightarrow x$

$(e_3)\ e*e \rightarrow e$

$(e_4)\ h(e) \rightarrow d$

and we have the following proof of $a =_R h(a)$ (see fig. 2a).

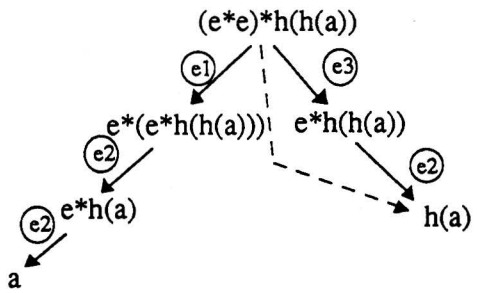


fig. 2a

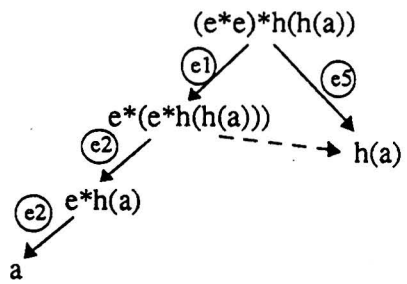


fig. 2b

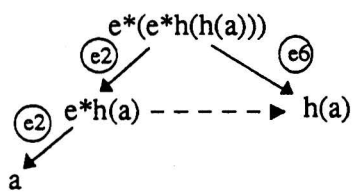


fig. 2c

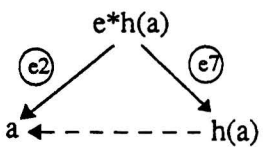


fig. 2d

In our example, it is easy to see that only rule e_2 can occur as the leftmost step of any proof of $a =_R h(a)$. This makes e_2 the only goal equation, and thus disallows any superpositions. In such a situation, the method proceeds by deriving new goal equations performing a paramodulation step on the *left hand side of a goal equation* and the *right hand side of a non-goal equation.* In analogy to a term coined by Dershowitz [6], we call such a step a *forward closure.* In our example, we could paramodulate e_2 with e_3, resulting in the new (goal) equation $(e*e)*h(x) \approx x$, which can be directed into $e_5 = (e*e)*h(x) \rightarrow x$, (see fig. 2a, 2b). The complete proof manages with just this single paramodulation step. Now, (see fig. 2b-d) the rest of the proof proceeds like ordinary completion.

It should be noted that equation e_4 is not used during the derivation. In fact, since the constant d cannot be removed by R, each proof of $a =_R h(a)$ actually is a proof of $a =_{R'}$

$h(a)$, where $R'=R\setminus\{e_4\}$. This particular form of redundancy resembles somewhat the notion of pure clauses (see, for instance, [5]). A derivation using standard completion, however, would infer consequences from e_4, regardless of whether they contribute to the proof.

We just briefly recall some of the basic notions of term rewriting, as they can be found, for instance, in [7] or [10].

For any term t, $\mathcal{P}os(t)$ denotes the set of tree positions of t, with $\lambda\in\mathcal{P}os(t)$ being the root position. We write $p<q$ if p is a prefix of q, and $p \parallel q$, if $p \not< q$ and $q \not< p$ hold. The term t/p is the subterm of t rooted at p, $t[p\leftarrow s]$ denotes replacement of t/p by s. The set of variable positions of t is defined by $\mathcal{VP}os(t) = \{p\in\mathcal{P}os(t) : t/p \in \mathcal{V}\}$, and the set of non-variable positions is defined by $\mathcal{FP}os(t) = \mathcal{P}os(t)\setminus\mathcal{VP}os(t)$. $\mathcal{H}ead(t)$ denotes the function symbol heading t.

For any relation $\rightarrow$, $\leftarrow$ denotes the inverse of $\rightarrow$, $\leftrightarrow$ denotes the symmetric closure of $\rightarrow$, and $\rightarrow^*$ denotes the reflexive, transitive closure of $\rightarrow$. The relation $\rightarrow_e^p$ denotes rewriting on position p using equation (or rule) e, $\rightarrow^\Pi$ denotes parallel rewriting on the set of disjoint positions Π, i.e., $s\rightarrow^\Pi t$, where $\Pi = \{p_1,\ldots,p_n\}$ is a set of mutually parallel (disjoint) positions of s, iff $s\rightarrow^{p_1}\ldots\rightarrow^{p_n} t$. By $\rightarrow^{\|}$, we denote parallel rewriting (without indicating the rewriting positions). finally, we use $\trianglelefteq$ ($\triangleleft$) for the (proper) encompassment ordering, i.e., $s\trianglelefteq t$, iff an instance of s is a subterm of t.

We use $s\approx t$ to denote an equational axiom with one side s and the other side t, while $s\equiv t$ denotes an equational axiom with left hand side s and right hand side t.

2 The Calculus

This section presents the derivation rules of the goal oriented completion calculus, which essentially consists in unfailing completion ([3],[9]), restricted by a certain critical pair condition and an additional *forward closure* rule. The rules operate on triples (E,R,G), where E is a set of equations, R is a set of rules, and G is a set of disequations (the actual goals). In the following we make use of the fact that $s =_E t$ holds iff $s_{gr} =_E t_{gr}$ holds, where $s_{gr}=t_{gr}$ is the skolemized form of s=t. We thus always assume the goals to be ground disequations. A set of goals is called *inconsistent*, iff it contains a disequation $t\neq t$. To each derivation rule, the corresponding proof transformation rule is shown.

We assume a reduction ordering $\succ$ on the set $\mathcal{T}$ of terms, which is total on the set $\mathcal{G}$ of ground terms.

1 Definition *A proof of* $s =_{E\cup R} t$ *is a sequence*

$$s \leftrightarrow_{E\cup R} \ldots \leftrightarrow_{E\cup R} t$$

of ***proof steps****. A proof of the form* $t\leftrightarrow t$ *is called* ***trivial****. A* ***variable overlap*** *is a proof* $s \leftrightarrow^p_{l\approx r} u \leftrightarrow^q_{l'\approx r'} t$, *such that* $p\in\mathcal{VP}os(l')$ *or* $q\in\mathcal{VP}os(l)$. *The subsumption quasi ordering* $\leq$ *on terms carries over to proofs in the following way: Let* $P = w_0 \leftrightarrow^{p_1}_{e_1} \ldots \leftrightarrow^{p_n}_{e_n} w_n$. *Then* $P\leq P'$, *iff* $P' = w'_0 \leftrightarrow^{p_1}_{e_1} \ldots \leftrightarrow^{p_n}_{e_n} w'_n$ *and there exists a substitution* σ *such that* $w'_i=w_i\sigma$ *for* $i=1,\ldots,n$.

In the following three definitions we introduce the inference rules of the goal oriented calculus. All rules are given with respect to a restriction Γ, which is a subset of the actual rules and equations, and which will be defined in more detail later on.

2 Definition *Let $e_1,e_2 \in E \cup R$, $\Gamma \subseteq E \cup R$, such that $e_1 \in \Gamma$ or $e_2 \in \Gamma$. The equation $s \approx t$ is called a* ***critical pair*** *of e_1 and e_2 w.r.t. the restriction Γ, if there is a variable overlap free proof $P =$*

$$s \leftrightarrow^{p}_{e_1} u \leftrightarrow^{\lambda}_{e_1} t$$

such that

(i) *P is minimal w.r.t. the subsumption ordering and*

(ii) *neither $s \succ u$ nor $t \succ u$ holds.*

$s \approx t$ is called trivial, if $s = t$. $CP_\Gamma(E,R)$ denotes the set of all non-trivial critical pairs w.r.t. Γ.

3 Definition *Let $e \in E$ and $s \neq t \in G$. Then $s' \neq t$ is called a (goal)* ***paramodulant*** *of e and $s \neq t$, if there is a proof*

$$s' \leftrightarrow^{p}_{e} s$$

such that $s' \not\succ s$ holds. $P(G,E)$ denotes the set of all paramodulants.

4 Definition *Let $\Gamma \subseteq E \cup R$, and let $e,e' \in E \cup R$, $e' \in \Gamma$. Then $s \approx t$ is called a (one step)* ***forward closure*** *of e,e' w.r.t. Γ, if there is a variable overlap free and minimal (w.r.t. $\trianglelefteq$) proof $P =$*

$$s \leftrightarrow^{p}_{e} u \leftrightarrow^{q}_{e} t$$

such that $p=\lambda$ or $q=\lambda$, and $t \not\succ u \not\succ s$ holds.

Let $e_1,\ldots,e_n \in E \cup R$. We define recursively $FC_n(e_1,\ldots,e_n)$ by $FC_0(e_1) = \{e_1\}$, and $FC_n(e_1,\ldots,e_n)$ is the set of all forward closures of elements of $FC_{n-1}(e_1,\ldots,e_{n-1})$ with e_n. Moreover, we define $FC(E,R) = \bigcup_{e,e' \in E \cup R} FC_1(e,e')$ to be the set of all (one-step) forward closures of equations in $E \cup R$.

5 Definition *Let E be a set of equations, R a set of rules, and G a set of disequations. We define the set $\Gamma(G)$ of* ***goal equations*** *to be the set $\{e \in E \cup R : \exists\, t' \in \mathcal{T}, s \neq t \in G:\ t \not\succ t',\ t \prec s,\ t' \leftrightarrow_e t\}$.*

Note that the definition of goal equations implies that $l \approx r \in \Gamma(G)$ iff $r \not\succ l$ and there is $s \neq t \in G$ with $t \prec s$ and $r \trianglelefteq t$.

Example Let $R = \{e_1, e_2, e_3\}$ with $e_1 = (x*y)*z \to x*(y*z)$, $e_2 = x*e \to x$, and $e_3 = x*i(x) \to e$, and let $G = \{a \neq i(i(a))\}$. Then e_2 is a goal equation, because $a*e \to_{e_2} a$ holds, or, equivalently, because $x \trianglelefteq a$ holds. The rules e_1 and e_3 are not goal equations, so we have $\Gamma(G) = \{e_1\}$. Moreover, $CP_\Gamma(R) = \{x*(e*z) \approx x*z,\ x*y \approx x*(y*e)\}$ are critical pairs of e_2 with e_1, and $FC_\Gamma(R) = \{(x*y)*e \approx x*y,\ x*(y*i(y)) \approx x\}$ are the forward closures of e_1 with e_2 and e_3 with e_2, respectively.

6 Definition *Let* $\Gamma=\Gamma(G)$, *and let* $R_E = \{u\sigma \to v\sigma : u \approx v \in E,\ u\sigma \succ v\sigma\}$. *The goal oriented superposition calculus is defined by the following derivation rules and proof transformation rules operating on quadruples* (E,R,G,F), *where* E *is a set of equations,* R *is a set of rules,* G *is a set of goals, and* F *is a subset of* $E \cup R$, *which designates elements derived by forward closure. We use the convention*

$$S[x/y] := \begin{cases} S\setminus\{x\}\cup\{y\}, & \text{if } x \in S \\ S & \text{otherwise} \end{cases}$$

Orientation:

$$\frac{E\cup\{l\approx r\},R,G,F}{E,R\cup\{l\to r\},G,F'} \quad \text{if } l\succ r,\ F'=F[l\approx r/l\to r] \qquad \frac{l \leftrightarrow_{l\approx r} r}{l \to_{l\to r} r}$$

Deduction (Superposition and Forward Closure)

$$\frac{E,R,G,F}{E\cup\{l\approx r\},R,G,F} \quad \text{if } l\approx r \in CP_\Gamma(e,e') \qquad \frac{l \leftrightarrow_{e'} u \leftrightarrow_e r}{l \leftrightarrow_{l\approx r} r}$$

$$\frac{E,R,G,F}{E\cup\{l\approx r\},R,G,F'} \quad \text{if } l\approx r \in FC_\Gamma(e,e'),\ F'=F\cup\{l\approx r\} \qquad \frac{l \leftrightarrow_{e'} u \leftrightarrow_e r}{l \leftrightarrow_{l\approx r} r}$$

Goal Paramodulation

$$\frac{E,R,G,F}{E,R,G\cup\{s'\not\approx t\},F} \quad \text{if } s'\not\approx t \in P(G,E) \qquad \frac{s \leftrightarrow s' \leftrightarrow \ldots \leftrightarrow t}{s' \leftrightarrow \ldots \leftrightarrow t}$$

Deletion

$$\frac{E\cup\{l\approx l\},R,G,F}{E,R,G,F'} \quad F'=F\setminus\{l\approx l\} \qquad \frac{l \leftrightarrow_{l\approx l} l}{\square}$$

Simplification

$$\frac{E\cup\{l\approx r\},R,G,F}{E\cup\{l'\approx r\},R,G,F'} \quad \text{if } l \to_R l' \text{ or } l \to_{R_E} l' \text{ by } u\to v \text{ with } u \lhd l;\ F'=F[l\approx r/l'\approx r] \qquad \frac{l \leftrightarrow_{l\approx r} r}{l \to l' \leftrightarrow_{l'\approx r} r}$$

$$\frac{E,R,G\cup\{s\not\approx t\},F}{E,R,G\cup\{s'\not\approx t\},F} \quad \text{if } s \to_R s' \text{ or } s \to_{R_E} s' \text{ by } u\to v \text{ with } u \lhd s \qquad \frac{s \to s' \leftrightarrow \ldots \leftrightarrow t}{s' \leftrightarrow \ldots \leftrightarrow t}$$

$$\frac{E,R\cup\{l\to r\},G,F}{E,R\cup\{l\to r'\},G,F'} \quad \text{if } r \to_{R\cup R_E} r',\ F'=F[l\to r/l\to r'] \qquad \frac{l \to_{l\to r} r}{l \to_{l\to r'} r' \leftarrow_{r\to r'} r}$$

$$\frac{E,R\cup\{l\to r\},G,F}{E\cup\{l'\approx r\},R,G,F} \quad \text{if (i) } l \to_{R\cup R_E} l' \text{ by } u\to v \text{ with } u \lhd l \text{ and (ii) } l\to r \notin F \qquad \frac{l \to_{l\to r} r}{l \to_{l\to r'} r' \leftarrow_{r\to r'} r}$$

We shall write $(E,R,G,F) \vdash\!\!\!\!-_\Gamma (E',R',G',F')$, *if* (E',R',G',F') *is derived from* (E,R,G,F) *by one of the derivation rules. For any two proofs* P,P', *we shall write* $P \Rightarrow P'$, *if* P' *is derived from* P *by one of the transformation rules.*

Note that the system does not allow simplifying the left hand side of any rule that is derived by forward closure. It is thus necessary to keep track of the history of equations generated, which is the only purpose of the set F.

It is easy to check that these inference rules form a sound derivation system:

7 **Lemma** *If there is a derivation* $(E,\emptyset,\{s\neq t\},F) \vdash_\Gamma^* (E',R',G',F')$ *such that* G' *is inconsistent, then* $s =_E t$ *holds.* □

Next, we shall prove that the derivation system is refutation complete under certain restrictions for $\Gamma=\Gamma(G)$. We remark that $(E,R,G,F) \vdash_\Gamma (E',R',G',F')$ implies that for any proof P in $E\cup R$, there exists a proof P' in $E'\cup R'$ with $P \Rightarrow P'$. As a first step, we define an ordering on proofs. The following definition is taken from [2].

8 **Definition** *We assume given a quadruple (E,R,G,F) as in definition 6.*

a) *For any proof step p of the form* $s \to^p_{l\to r} t$ *with* $l\to r \in R\setminus F$, *we define the complexity* $c(p) = (\{s\},s/p,l,t)$. *For any proof step p of the form* $s \to^p_{l\to r} t$ *with* $l\to r \in R\cap F$, *we define the complexity* $c(p) = (\{s\},\bot,\bot,t)$. *The complexity of a step* $s \leftrightarrow^p_{l=r} t$ *is defined by* $(\{s,t\},-,-,-)$. *Finally, the complexity c(P) of a proof P is defined to be the multiset of the complexities of its proof steps.*

b) *The ordering* $\succ^c$ *on pairs c(p) is defined as the lexicographic combination of the multiset extension* $\succ_{mul}$ *of the term ordering* $\succ$, *the subterm ordering, the encompassment ordering, and the term ordering* $\succ$. *The ordering* $\succ_{\mathcal{D}}$ *is defined by* $P \succ_{\mathcal{D}} P'$, *iff* $c(P) \succ^c_{mul} c(P')$. *The element* $\bot$ *is a minimal element of* $\mathcal{T}$ *w.r.t. the encompassment ordering.*

The proof ordering $\succ_{\mathcal{D}}$ is well-founded, and it is easy to verify that the proof transformation rules given in definition 6 decrease the complexity of proofs, that is, $P\Rightarrow P'$ implies $P \succ_{\mathcal{D}} P'$.

So far, the proof of completeness of the calculus is similar to the standard argument. However, the inference rules given in definition 6 restrict the generation of critical pairs, and so there might exist peaks in a given proof P that are not resolvable due to the restriction Γ. So we have to prove that for any non-trivial proof P, there is some proof transformation rule that applies to P. However, we encounter one problem, which occurs with variable overlaps.

Example Let R = {b→a, b→c, fxx→e, fac→c}, let the term ordering $\succ$ be chosen such that c $\succ$ e, and consider the following proof (see fig. 3) P of $c =_R e$:

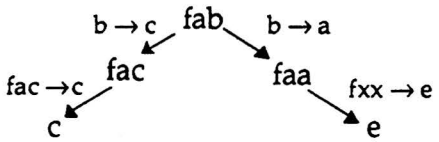


fig. 3

Let G={c≠e}. The (single) peak occurring in this proof cannot be resolved, since the rules b→c and b→a are both non-goal rules. As c $\succ$ e, we have fxx→e $\in \Gamma(G)$. The overlap

between b→a and fxx→e is a variable overlap, hence it cannot be used for deriving a forward closure of b→a. It is easy to verify that fxx→e is the only rule in Γ(G), hence the proof P is irreducible by ⇒.

The solution to this problem is the observation that there is a *top-critical pair* of fxx→e and fac→c, i.e. fxx and fac are top-unifiable.

9 Definition a) *Let $s,t\in\mathcal{T}$. The pair (σ,E) is a* ***top-unifier*** *of s and t, if there exists a proof*

$$s\sigma \leftrightarrow_E^{p_1} \dots \leftrightarrow_E^{p_n} t\sigma$$

with $p_i\in \mathcal{VP}os(s)\cup\mathcal{VP}os(t)$ and $p_i \| p_j$ for $i,j=1,\dots,n$, $i\neq j$. We shall sometimes abbreviate the sequence $\leftrightarrow_E^{p_1} \dots \leftrightarrow_E^{p_n}$ by $\rightarrow^{\|}_E$

b) *Let $e_1,e_2\in E\cup R$, $e_1\equiv l_1\approx r_1$, and let $\Gamma\subseteq E\cup R$, such that $e_1\in\Gamma$ or $e_2\in\Gamma$. $s\approx t$ is called a* ***top-critical pair*** *of e_1 and e_2 w.r.t. Γ, if there is a proof*

$$s \leftrightarrow_{e_2}^{p} v \leftrightarrow_E^{p_1} \dots \leftrightarrow_E^{p_n} w \leftrightarrow_{e_1}^{\lambda} t$$

with $p\notin \mathcal{VP}os(l_1)$, such that neither $s \succ v$ nor $t \succ w$ holds and with $p<p_i$, $p_i\in\mathcal{VP}os(s)\cup\mathcal{VP}os(t)$ and $p_i \| p_j$ for $i,j=1,\dots,n$, $i\neq j$. The equation $s\approx t$ is a ***proper*** *top-critical pair, iff it is not a critical pair of e_1 and e_2 in the sense of definition 2. $\mathrm{TCP}_\Gamma(E,R)$ denotes the set of all proper top-critical pairs w.r.t. Γ.*

In a similar way we define top-paramodulation and top-forward closure, and the sets TP(E,R) of all top-paramodulants, and the set TFC(E,R) of all top-forward-closures.

It is easy to see from the definition that $s\approx t$ is a top-critical pair of $l_1\approx r_1$ and $l_2\approx r_2$, if there is $p\in\mathcal{FP}os(l_1)$ and a top-unifier (σ,E) of l_2 and l_1/p, such that $s = l_1\sigma[p\leftarrow r_2\sigma]$ and $t = r_1\sigma$.

The notion of top-unification was introduced by Dougherty and Johann [8]. They define top-unification operationally by giving an algorithm. We show that the two definitions basically coincide:

10 Definition *We define derivation systems D_1, D_2. D_1 operates on a set of equations, and D_2 operates on pairs (σ,E), where σ is a substitution and E a set of equations. D_1 is defined by the following two rules*

$$\frac{E\cup\{fs_1\dots s_n = fs_1\dots s_n\}}{E\cup\{s_i=t_i : i=1,\dots,n\}} \quad\text{and}\quad \frac{E\cup\{s=t\}}{\bot}\_ \quad \text{if } \mathcal{H}ead(s)\neq\mathcal{H}ead(t)$$

and D_2 is defined by

$$\frac{\sigma, E\cup\{x=t\}}{\sigma\cup\{x\mapsto t\}, E\{x\mapsto t\}}\_ \quad \text{if } x\notin\mathcal{V}ar(t)$$

The derivation system D_1 is the one given in [8] and [16].

11 Lemma *Let $s,t\in\mathcal{T}$, let σ be a substitution, and E a set of equations. The pair (s,t) is top-unifiable by (σ,E), iff there are maximal derivations $E_0 \Rightarrow^*_{D_1} E'$ with $E_0=\{s=t\}$ and $E'\neq\bot$, and $(\sigma_0,E') \Rightarrow^*_{D_2} (\sigma,E)$ with $\sigma_0 = \mathrm{id}$.*

Proof. First, we remark that (i) $E \Rightarrow_{D_1} E'$ and $s \leftrightarrow^{\shortparallel}_{E} t$ imply $s \leftrightarrow^{\shortparallel}_{E'} t$ and (ii) $(\sigma,E) \Rightarrow_{D_2} (\sigma',E')$ and $s\sigma \leftrightarrow^{\shortparallel}_{E} t\sigma$ imply $s\sigma' \leftrightarrow^{\shortparallel}_{E'} t\sigma'$. If there are maximal derivations $E_0 \Rightarrow^{*}_{D_1} E'$ with $E_0 = \{s=t\}$ and $E' \neq \bot$, and $(\sigma_0,E') \Rightarrow^{*}_{D_2} (\sigma,E)$ with $\sigma_0 = id$, then we have $s\sigma \leftrightarrow^{\shortparallel}_{E} t\sigma$. Due to the construction of the derivation D_1, each position involved is a variable position either in s or in t. If, on the other hand, $E_0 \Rightarrow^{*}_{D_1} E' = \bot$ holds, then there is $p \in \mathcal{FPos}(s) \cap \mathcal{FPos}(t)$, such that $\mathcal{Head}(s/p) \neq \mathcal{Head}(t/p)$, which implies that (s,t) is not top-unifiable. □

12 Definition *A variable overlap* $s \leftrightarrow^{p}_{e} u \leftrightarrow^{q}_{e'} t$ *is called* ***critical*** *(w.r.t. Γ), if* $u \not\succ s$ *and* $t \not\succ u$, $p>q$, *and* $e \notin \Gamma$ *hold.*

13 Lemma Let $s \neq t \in G$ with $s \succ t$. *Let* $R=\{e_1,\ldots,e_n\}$, *and let*

$$P = t_0 \rightarrow_{e_1} t_1 \rightarrow_{e_2} \ldots \rightarrow_{e_n} t_n \equiv t$$

be a proof of $t_0 =_R t_n$, *which is minimal w.r.t.* $\succ$ *and has no critical variable overlaps w.r.t.* $\Gamma(G)$. *If there exists* $e_i \notin \Gamma(G)$, *then there is a proof* P' *with* $P \Rightarrow P'$.

Proof. For each j=1,...,n, let p_j be the rewrite position used in the step $t_{j-1} \rightarrow t_j$.

Let i be the largest number, such that $e_i \notin \Gamma(G)$. Then it follows immediately from the definition of $\Gamma(G)$ that i<n. We have $e_i \notin \Gamma(G)$, and $e_{i+1} \in \Gamma(G)$.

Case 1: If $p_i \shortparallel p_{i+1}$, then we can construct a proof $P' = t_0 \rightarrow_{e_1} \ldots t_{i-1} \rightarrow_{e_{i+1}} t'_i \rightarrow_{e_i} t_{i+1} \rightarrow \ldots \rightarrow_{e_n} t_n$. Continuing this way, we obtain a proof $P^* = t_0 \rightarrow_{e'_1} t'_1 \rightarrow_{e'_2} \ldots \rightarrow_{e'_n} t_n$ with $e'_k \notin \Gamma(G)$, and $e'_{k+1} \in \Gamma(G)$.

Case 2: If $t'_{k-1} \rightarrow_{e'_k} t'_k \rightarrow_{e'_{k+1}} t'_{k+1}$ is a variable overlap, then $p'_k \leq p'_{k+1}$ follows from the assumption that P has no critical variable overlaps. In a way similar to the proof of the critical pair lemma, we can construct a proof $P' = t_0 \rightarrow_{e'_1} \ldots t'_{k-1} \rightarrow^{\shortparallel}_{e'_{k+1}} t''_k \rightarrow_{e_k} t'''_k \shortparallel\leftarrow_{e'_{k+1}} t'_{k+1} \ldots \rightarrow_{e_n} t_n$ with $P' \prec P$, contradicting the minimality of P.

Case 3: If $t'_{k-1} \rightarrow_{e'_k} t'_k \rightarrow_{e'_{k+1}} t'_{k+1}$ is a proper overlap, then $FC(e'_k,e'_{k+1}) \neq \emptyset$. The forward closure rule thus applies to P, hence there is a proof P' with $P \Rightarrow P'$. □

14 Lemma *Let P be a nontrivial ground proof of* $s =_{E \cup R} t$, *which is a minimal proof of* $s =_{E \cup R} t$ *w.r.t.* $\succ$ *and has no critical variable overlaps. Moreover, let* $s \neq t \in G$. *Then there is a proof P' with* $P \Rightarrow P'$.

Proof. As P is a ground proof, P uses only rules in $R' := R \cup R_E$. W.l.o.g. we assume that $s \succ t$. Let

$$P = s \equiv t_0 \leftrightarrow \ldots \leftrightarrow t_n \equiv t$$

If P has a subproof of the form $P_0 = s \rightarrow u$ or $P_0 = t \rightarrow u$, then the goal paramodulation rule applies to P, proving the assertion of the lemma. Otherwise, P contains a peak. Let $P_0 = t_{k-1} \leftarrow t_k \rightarrow t_{k+1}$ be the rightmost such peak, i.e.

$$P = t_0 \leftrightarrow \ldots \leftrightarrow t_{k-1} \leftarrow t_k \rightarrow t_{k+1} \rightarrow \ldots \rightarrow t_n$$

If P_0 does not originate from a critical overlap, then the critical pair lemma [11] implies that there is a proof P'_0 of $s =_{R'} t$ with $P_0 \succ P'_0$, contradicting the minimality assumption on P. If $e_j \notin \Gamma(G)$ holds for some j=k,...,n, then we are done by lemma 13. Otherwise,

$e_k \in \Gamma(G)$, hence P_0 has the form $t_{k-1} \leftarrow_{R'} t_k \rightarrow_{R' \cap \Gamma(G)} t_{k+1}$, and the superposition rule applies to the critical overlap P_0, again proving the assertion of the lemma. □

15 **Lemma** *Let*

$$P = s \leftarrow_{e'}^{q} t_0 \leftrightarrow_{e_1}^{p_1} t_1 \ldots \leftrightarrow_{e_n}^{p_n} t_n \rightarrow_{e}^{p} t$$

be a ground proof with $e=l \rightarrow r$, such that $p_i \in \mathcal{VP}os(l)$ for $i=1,\ldots,n$, and such that there is a critical overlap $s' \leftarrow_{e'}^{q} u \rightarrow_{e}^{p} t'$ of (e,e'). Then there is a proof

$$P' = s \leftrightarrow^{*} u' \leftrightarrow_{s' \approx t'} v' \leftrightarrow^{*} t$$

with $P' \prec P$.

Proof. Let $E:=\{e_1,\ldots,e_n\}$, and let $e'=l' \rightarrow r'$. We can assume e and e' to be variable disjoint, so there is a ground substitution μ with $t_n/p=l\mu$, and $t_0/q=l'\mu$. W.l.o.g. we can assume that $p<q$, i.e. $q=pq'$ for some q'. Let $\sigma=mgu(l/q', l')$, such that σ does not introduce new variables. For each $x \in \mathcal{V}ar(l) \cap \mathcal{V}ar(r)$ with $x\sigma\mu \neq x\mu$, there is a proof $x\sigma\mu =_E x\mu$, and similarly for each $y \in \mathcal{V}ar(l') \cap \mathcal{V}ar(r')$. So we have a proof

$$P' = s \equiv t_0[q \leftarrow r'\mu] \leftrightarrow^{*}_E t_0[q \leftarrow r'\sigma\mu] \leftarrow_{e'}^{q} t_0[q \leftarrow l'\sigma\mu] \equiv t_n[p \leftarrow l\sigma\mu] \rightarrow_{e}^{p} t_n[p \leftarrow r\sigma\mu] \leftrightarrow^{*}_E t_n[p \leftarrow r\mu] \equiv t$$

A straightforward, but tedious argument shows that $P' \prec P$ holds. □

16 **Lemma** *Let P be a minimal nontrivial ground proof of $s =_{E \cup R} t$, let $\Gamma = \Gamma(G)$ with $s \neq t \in G$, and let $TFC_\Gamma(E,R) = TCP_\Gamma(E,R) = TP_\Gamma(E,R) = \emptyset$. Then there exists a proof P' with $P \Rightarrow P'$.*

Proof. Assume to the contrary that P is irreducible w.r.t. $\Rightarrow$. According to lemma 14, we can assume that P has critical variable overlaps. Let

$$P = s \equiv t_0 \leftrightarrow_{e_1}^{p_1} \ldots \leftrightarrow_{e_n}^{p_n} t_n \equiv t$$

with $e_i = l_i \rightarrow r_i$, and let k be the least element of $\{1,\ldots,n\}$, such that $t_{k-1} \rightarrow_{e_k} t_k \rightarrow_{e_{k+1}} t_{k+1}$ is a critical variable overlap.

Case 1: there is $m \in \{1,\ldots,k-1\}$, and $q \in \mathcal{VP}os(l_{k+1})$, such that $p_m < q$. Let $P' = t_{m-1} \leftrightarrow_{e_m}^{p_m} \ldots \leftrightarrow_{e_{k+1}}^{p_{k+1}} t_{k+1}$. Since P' is minimal and irreducible w.r.t. $\Rightarrow$, each peak of P' comprises two rules or equations in $(E \cup R) \setminus \Gamma$. Moreover, there is no critical variable overlap e_i, e_{i+1} for $i<k$, hence $e_i \notin \Gamma$ holds for $i=1,\ldots,k-1$.

Suppose there is some $j \in \{m+1,\ldots,k-1\}$ such that $p_j \notin \mathcal{VP}os(l_{k+1})$, and let j be a maximal such j, that is, $p_{j'} \in \mathcal{VP}os(l_n)$ for $j'=j+1,\ldots,k$. If q is any variable position of l_{k+1}, then $p_j \not< q$, since m was chosen maximal with that property. Hence we have $p_j \parallel p_{j'}$ for $j'=j+1,\ldots,k$. By "moving the step $t_{j-1} \leftarrow_{e_j}^{p_j} t_j$ to the right", we obtain a proof

$$P'' = t_{m-1} \leftrightarrow \ldots \leftrightarrow t_{j-1} \leftrightarrow_{e_{j+1}}^{p_{j+1}} t'_{j+1} \leftrightarrow \ldots \rightarrow_{e_k}^{p_k} t'_k \leftrightarrow_{e_j}^{p_j} t_k \rightarrow_{e_{k+1}}^{p_{k+1}} t_{k+1}$$

If $t_{j-1} \prec t_j$, then we have $t_{j-1} \leftarrow_{e_j}^{p_j} t_j$, and $P' \succ P''$ holds contradicting the choice of P'. So we can assume $t_{j-1} \succ t_j$, hence $t_{j-1} \rightarrow_{e_j}^{p_j} t_j$ and $e_j \notin \Gamma$. Now consider the subproof $P''_0 = t'_k \rightarrow_{e_j}^{p_j} t_k \rightarrow_{e_{k+1}}^{p_{k+1}} t_{k+1}$ of P''. Since $p_j \notin \mathcal{VP}os(l_n)$, P''_0 is not a variable overlap. Moreover, we have

$e_j \notin \Gamma$ and $e_{k+1} \in \Gamma$. Hence there is a forward closure e' of (e_j, e_{k+1}), contradicting the irreducibility of P w.r.t. $\Rightarrow$.

So we can assume that $p_j \in \mathcal{VP}os(l_{k+1})$ holds for all $j \in \{m+1,\ldots,k\}$. We have $e_j \notin \Gamma$ for each $j<k$, in particular $e_m \notin \Gamma$. If $t_{m-1} \succ t_m$, there is a top-forward closure of (e_m, e_{k+1}), and if $t_{m-1} \prec t_m$, there is a top-critical pair of (e_m, e_{k+1}). We only treat the second case, the first one being rather similar. Since $TCP_\Gamma(E,R)=\emptyset$, (e_m, e_{k+1}) defines a proper critical pair. Hence, according to lemma 15, there is a proof $P'' \prec P'$ of $t_{m-1} =_{E \cup R} t_{k+1}$, contradicting the assumption of the lemma.

Case 2: P has no step $t_{m-1} \leftrightarrow^{p_m}_{e_m} t_m$ with $p_m<q$, for some $q \in \mathcal{VP}os(l_n)$. The proof is very similar to the proof of case 1, this time using the fact that $TP_\Gamma(E,R) = \emptyset$. □

17 Definition *A derivation* $(E_0,R_0,G_0,F_0) \vdash_{\Gamma'(E_0,R_0,G_0)} (E_1,R_1,G_1,F_1) \vdash_{\Gamma'(E_1,R_1,G_1)} \ldots$ *is a fair derivation, iff the following conditions are satisfied:*

(i) *If* $TCP_{\Gamma(G_i)}(E_i,R_i) \cup TFC_{\Gamma(G_i)}(E_i,R_i) \cup TP_{\Gamma(G_i)}(E_i,R_i) \neq \emptyset$, then $\Gamma'(E_i,R_i,G_i) = E_i \cup R_i$. *Otherwise,* $\Gamma'(E_i,R_i,G_i) = \Gamma(G_i)$.

(ii) *Each critical pair and each forward closure of* (R^∞,E^∞) *is contained in* $\bigcup_{k \in \mathbb{N}} E_k$

(iii) *Each result of a goal paramodulation step of* $(R^\infty,E^\infty,G^\infty)$ *is contained in* $\bigcup_{k \in \mathbb{N}} G_k$

$R^\infty,E^\infty,G^\infty$ *are the sets of* ***persistent*** *rules, equations, and goals,*

$$R^\infty = \bigcup_{k \in \mathbb{N}} \bigcap_{j \geq k} R_j$$

and similarly for E^∞, G^∞.

18 Theorem *If* $s =_E t$ *holds, and* $(E_0,R_0,G_0) \vdash (E_1,R_1,G_1) \vdash \ldots$ *is a fair derivation with* $E_0 = E$, $R_0 = \emptyset$, *and* $G_0 = \{s \neq t\}$, *then there is* $n \in IN$, *such that* G_n *is inconsistent.*

Proof. We show: If P is a ground proof of $s_i =_{E_i \cup R_i} t_i$ with $s_i \neq t_i \in G_i$, and $s_i \neq t_i$, then there exists $j \geq i$ and a ground proof P' of $s_j =_{E_j \cup R_j} t_j$ with $s \neq t_j \in G_j$ with $P \succ_{\mathcal{D}} P'$.

First, we remark that we can assume P to be a minimal ground proof of $s_i =_{E_i \cup R_i} t_i$. If P uses a non-persistent rule, equation, or goal e_i, then e_i is removed by some step $(E_j,R_j,G_j) \vdash (E_{j+1},R_{j+1},G_{j+1})$, which implies the existence of P' with $P \succ P'$. We thus can assume that P only uses persistent rules, equations, and goals. Since P is a ground proof, only rules of $R=R^\infty \cup R_{E^\infty}$ are used. If a proof transformation step $P \Rightarrow P'$ applies to P, then due to the fairness condition, the corresponding derivation step is performed for some $j>i$, hence $P \succ P'$, where P' is a proof of $s_j =_{E_j \cup R_j} t_j$ with $s_j \neq t_j \in G_j$.

So let us assume that no transformation step applies to P. Then, according to lemma 15, $TCP_{\Gamma(G_i)}(E_i,R_i) \cup TFC_{\Gamma(G_i)}(E_i,R_i) \cup TP_{\Gamma(G_i)}(E_i,R_i) \neq \emptyset$. According to the fairness condition, this implies $\Gamma'(E_i,R_i,G_i)=E_i \cup R_i$. Since no goal paramodulation step applies to P, we can conclude that P contains a peak $s' \leftarrow u \rightarrow t'$, which is a proper overlap. Since $\Gamma'(E_i,R_i,G_i)=E_i \cup R_i$, the critical pair rule applies to P, which is a contradiction.

Now we have proved that for each ground proof of $s_i =_{E_i \cup R_i} t_i$ with $s_i \neq t_i$ and $s_i \neq t_i \in G_i$ there is a proof P' with $P \succ_{\mathcal{D}} P'$. Since the proof ordering $\succ_{\mathcal{D}}$ is well founded, we can

conclude the existence of $n \in \mathbb{N}$, such that $s_n = t_n$ and $s_n \neq t_n \in G_n$, i.e., G_n is inconsistent. □

3 Implementation and Practical Results

In this section several practical aspects of the goal oriented calculus are discussed.

The need to weaken the restriction Γ to the whole set $E \cup R$ of equations and rules, whenever top-critical pairs, top-forward closures, or top-goal paramodulations occur, raises the question of practical relevance. As $\Gamma = E \cup R$ makes the restriction void, and reduces the goal oriented calculus to standard completion, a frequent occurrence of this top-unifiable structures in practice would strongly reduce the value of the whole approach. On the other hand, the need to extend the restriction Γ only arises in connection with critical variable overlaps. Several examples of group theory and the theory of Ternary Boolean Algebra haver been considered so far (see fig. 4), and never did critical variable overlaps occur.

In fig. 4, we provide a statistics for GOC, an actual implementation of the goal oriented completion method, in terms of the number of equations generated for several examples. These results are compared to the results of the theorem prover OTTER [14], and the SbReve system [1]. Wos_2 and Wos_5 are among six equality problems published by Lusk and Overbeek [12], Wos_2 stating that the inverse function in a group is an involution. Wos_5 is a lemma for axiom independence in the theory of Ternary Boolean Algebra and reads as follows:

$f(f(v,w,x),y,f(v,w,z)) \approx f(x,y,f(v,w,z))$

$f(y,x,x) \approx x$

$f(x,x,y) \approx x$

$f(x,y,x) \approx x$

$f(gy,y,x) \approx x$

$f(a,ga,b) \not\approx b$

$GrpComm_1$ and $GrpComm_2$ are two problems of group theory. $GrpComm_1$ states that a group G with the usual axiomatization is commutative, if $i(xy) \approx i(x)i(y)$ holds for all $x,y \in G$. $GrpComm_2$ states that G is commutative, if $x(y(xy)) \approx x(x(yy))$ holds for all $x,y \in G$.

	GOC	OTTER	SbReve
Wos_2	17	37	49
Wos_5	578	12,386	410
$GrpComm_1$	112	585	184
$GrpComm_2$	376	505	370

fig. 4: Number of Equations Generated for Several Problems

Another aspect of the goal oriented approach concerns the concentration on goal equations during the derivation process. The basic idea of the whole strategy is the importance of goal equations, and this should be reflected in the search strategy. Usually, the weighting of equations, which serves to select the next equation to deal with, is based on the size of the terms occurring in the equation. In the examples shown in fig. 4, goal equations were preferred over equations of slightly smaller size, i.e., a weighting function $\omega(e)$ for $e = l \approx r$ was employed, which had the form

$$\omega(e) = |l| + |r| - k*\gamma(e),$$

with $\gamma(e)=1$ if $e \in \Gamma$, and $\gamma(e)=0$ otherwise, and $1 \leq k \leq 3$.

Acknowledgment

This work was performed during a one year research stay at Stony Brook University. I would like to thank Jieh Hsiang and the people from Max Kade-Foundation, who made this visit possible. I also appreciated the discussions with Jieh Hsiang on the subject, which have contributed much to the contents and the presentation of this paper.

References

[1] S. Anantharaman, J. Hsiang, and J. Mzali. SbReve2: A Term Rewriting Laboratory with (AC)Unfailing Completion. In: D. Plaisted (Ed.)*Proc. of 3rd International Conference on Rewriting Techniques and Applications*, Chapel Hill, N.C. (1989). Springer LNCS , 348-360.

[2] L. Bachmair, N. Dershowitz and J. Hsiang. Orderings for Equational Proofs. *Proc. of 1st Workshop on Logic in Computer Science* (1986), 346-357.

[3] L. Bachmair, N. Dershowitz, and D. Plaisted. Completion without Failure. *Coll. On the Resolution of Equations in Algebraic Structures*, Austin (1987), Academic Press.

[4] M. P. Bonacina and J. Hsiang. On Fairness of Completion-Based Theorem Proving Strategies. In: R. V. Book (Ed.). *Proc. of 4th Int. Conf. on Rewriting Techniques and Applications*, Como (1991). Springer LNCS 488, 348-360.

[5] C.L. Chang and R.C.T. Lee. Symbolic Logic and Mechanical Theorem Proving. Academic Press, New York 1973.

[6] N. Dershowitz. Termination of Linear Rewriting Systems. In: S. Even, O. Kariv (Ed.). *Proc. of 8th Int. Coll. on Automata, Languages and Programming*, Acre, Israel (1981). Springer LNCS 115, 448-458.

[7] N. Dershowitz, J.-P. Jouannaud. Notations for Rewriting. *Bulletin of the EATCS*, **43** (1991), 162-173.

[8] D. J. Dougherty and P. Johann. An Improved General E-Unification Method. In: M. E. Stickel (Ed.). *Proc. 10th International Conference on Automated Deduction*, Kaiserslautern (1990). Springer LNCS 449, 261 - 275.

[9] J. Hsiang and M. Rusinowitch. On Word Problems in Equational Theories. In: Th. Ottmann (Ed.). *Proc. of 14th Int. Colloquium on Automata, Languages and Programming*, Karlsruhe (1987). Springer LNCS 267, 54-71.

[10] J. W. Klop. Term Rewriting Systems. In: S. Abramsky, D. M. Gabbay, and T. S. E. Maibaum (Eds.). *Handbook of Logic in Computer Science*, Oxford University Press, Oxford, to appear.

[11] D.E. Knuth and P.B. Bendix. Simple Word Problems in Universal Algebra. In: J. Leech (Ed.). *Computational Problems in Universal Algebra*, Pergamon Press (1970).

[12] E. Lusk and R. Overbeek. A short problem set for testing systems that include equational reasoning. Argone National Laboratory, 1984.

[13] J. Morris. E-Resolution: Extensions of Resolution to Include the Equality Relation. In: *Proc. of 1st Int. Joint Conf. on Artificial Intelligence*, Washington, D.C. (1969), 287-294.

[14] W. McCune. *OTTER User's Manual*, Argonne Report ANL-88-44 (1988).

[15] G. Robinson and L. Wos. Paramodulation and theorem-proving in first-order theories with equality. In: B. Meltzer, & D. Michie, (Eds.): *Machine Intelligence* 4. Edinburgh, (1969), 135 - 150.

[16] W. Snyder and Ch. Lynch. Goal Directed Strategies for Paramodulation. In: R. V. Book (Ed.). *Proc. of 4th Int. Conf. on Rewriting Techniques and Applications*, Como (1991). Springer LNCS 488, 150-161.

[17] L. Wos, D. Carson, and G. Robinson. Efficiency and Completeness of the set of support strategy in theorem proving. *Journal of the ACM* **12**, (1965), 698-709.

[18] L. Wos. Automated Reasoning. 33 Basic Research Problems. Prentice Hall, Englewood Cliffs1988.