CANONICAL ALGEBRAIC SIMPLIFICATION IN COMPUTATIONAL LOGIC

Dallas S. Lankford

May 1975 ATP-25

Appeared as report ATP-25, Univ. of Texas, Math. Dept.,
Automatic Theorem Proving Project, Austin, Texas 78712

CANONICAL ALGEBRAIC SIMPLIFICATION IN COMPUTATIONAL LOGIC

by

Dallas S. Lankford
Southwestern University
Department of Mathematics
Georgetown, Texas 78626

May 1975

Revisions appear as footnotes and as the preface.

current address: Louisiana Tech University
Math Department
Ruston, Louisiana 71272

This work was supported in part by NSF Grant DCR 74-12886.

PREFACE

This paper develops three key ideas: (1) the Knuth-Bendix-Slagle approaches to term rewriting are reworked into a straightforward form which has been the model for all subsequent treatments of term rewriting, (2) the algebraic properties of term rewriting systems are systematically developed to the point that the generalization to equivalence class term rewriting systems is clearly called for even if the precise details of how to accomplish that are not evident, and (3) the inclusion of term rewriting systems in general logical systems via blocked immediate narrowing and blocked inference (resolution, chaining, etc.) of other kinds provides the basis of practical systematic treatments of equality in computational logic.

When I wrote this paper I did not appreciate the debt which is owed to the earlier work of Evans (Proc. Camb. Phil. Soc. 47 (1951), 647-649) and Newman (Ann. Math. 43 (1942), 223-243), primarily because the Knuth and Bendix paper does not discuss the evolution of their approach. In retrospect I can now see how profoundly Evans' work influenced Knuth and Bendix, and in turn how profoundly Newman influenced Evans. The genesis of Slagle's sets of simplifiers still remains a mystery to me, but it was from him that I got the idea to consider finite and unique termination separately. Slagle also introduced the notion of a priori fully narrowed sets of clauses, but for reasons known only to him, he did not consider iterated immediate narrowing. Perhaps the reasons are obvious after one examines the efforts required to establish the refutation completeness of blocked immediate narrowing in "Canonical inference." The key idea which unlocks refutation completeness proofs for immediate narrowing is the Knuth and Bendix completion procedure.

Dallas Lankford
Ruston, Louisiana
November 1980

CANONICAL ALGEBRAIC SIMPLIFICATION IN COMPUTATIONAL LOGIC

by Dallas S. Lankford

ABSTRACT

An expansion of the concept of complete set of reductions is developed and a new general approach to the finite termination problem is presented. Through consideration of mathematical structures from logic and universal algebra, a general design approach to the construction of canonical simplification routines is presented which is applicable to many well-known theories, including commutative theories, such as groups, rings, Boolean algebras, and modules over rings. Also based on the concept of reduction, a new refutation complete restriction of paramodulation is described which appears to be a significant advance in computational efficiency.

1. INTRODUCTION

A variety of methods for the treatment of equality in computational logic have been suggested in the past. The case for using equality axioms in automatic deduction has been investigated by Kowalski (8), but most approaches have favored a separate inference process, often a variation of paramodulation which was introduced by Robinson and Wos (14). From a theoretical standpoint, paramodulation is a valuable conceptual asset for theorem prover design, if only because it provides a basic model of a refutationally complete theorem prover for equality. However, as experiments have shown, paramodulation is practically unacceptable because of catastrophic exponential clause explosion. Systems based on the equality axioms are similarly inefficient.

One theme which emerged from the ensuing search for computationally more efficient equality procedures has been the attempt to extend complete restrictions of resolution to analogous complete restrictions of paramodulation, as exemplified by Chang (2), Chang and Slagle (3), Kowalski (8), Lankford (9), and Wos and Robinson (18). However, many complete restrictions of resolution, especially combinations of strategies, fail to have analogous complete restrictions for paramodulation. Furthermore, experimental results have indicated that these kinds of restrictions do little to improve the power and efficiency of theorem provers for equality.

A more promising trend has been the investigation of a number of ideas related to the informal notion of what has been called simplification or reduction. In an early study of a kind of reduction known as demodulation, Wos, et al. (17) recognized the primary obstacle to an adequate theory of simplification, citing the existence of unsolvable word problems as an insurmountable barrier against the development of universal canonical procedures for equality. Still, the unsolvability of the word problem for some theories does not imply the unsolvability of the word problem for all theories; and, indeed, a primary goal of this paper is to illustrate how well-known mathematical solutions to the word problem for specific theories can be systematically used to construct more efficient procedures for equality in computational logic. In addition, experiments such as those of Bledsoe, et al. (1) and Huet (6), which support the folklore of the utility of ad hoc simplification in computational logic, and the theoretical articles of Plotkin (13) and Slagle (15) point to the need for a better understanding of the practical and theoretical role of reduction in theorem proving.

The results of this paper may be viewed as generalizations and extensions of two sources, Knuth and Bendix (7) and Slagle (15), both of which are concerned with decision procedures for equational theories that are based on certain sets of rewrite rules. These two sources grapple with two central problems in the development of a practical theory of reduction, that of determining when an algebraic theory can be axiomatized by such sets of rewrite rules, and that of combining the associated decision procedures with

rules of inference in a refutationally complete manner. Knuth and Bendix (7) explore the former problem, investigating an algorithm which often detects the solvability of the word problem for a given theory by mechanically deriving the appropriate set of rewrite rules, which they call a complete set of reductions, from the axioms of the theory. Their method involves expressing the axioms of the theory as a set of rewrite rules which possess a finite termination property, and deriving additional rewrite rules from the axiom set until the initial and derived rewrite rules satisfy a closure condition. While the importance of their method is attested to by a number of examples, including a solution to the word problem for free groups on no generators and no relations, their method does not apply directly to commutative theories, since a commutative rewrite rule fails to possess the finite termination property. Slagle (15) attacks the latter problem, suggesting ways to combine such sets of rewrite rules, which he calls sets of simplifiers, with resolution and paramodulation, and establishing refutation completeness for some specific canonical inference rules, for example, blocked subsumption hyper-resolution and paramodulation in conjunction with two additional inference rules called identity and fixed paramodulation.

In Section 2, I merge the two variant concepts of canonical reduction, namely, complete sets of reductions and sets of simplifiers. The principal difference between my approach and the approach of Knuth and Bendix (7) is that, borrowing from Slagle (15), I redefine a set of reductions to be a set of rewrite rules which possess a finite termination property, as opposed to a

set of rewrite rules that satisfy a certain order relation. Except for Theorem 4 of Knuth and Bendix (7), all of their theorems which develop the completeness test clearly[1] remain true for this more general concept of reduction. Consequently, my discussion will be cursory, relying where possible on their results. The primary new result of this section is an alternate approach to the finite termination problem which is at least as convenient for application purposes and perhaps conceptually more transparent than the approach of Knuth and Bendix (7).

In Section 3, I begin with a summary of some well-known facts from logic and universal algebra which provide the theoretical framework for a discussion of the algebraic properties of sets of reductions. A thorough development of the requisite preliminaries is beyond the scope of this paper, so I will assume familiarity with portions of Universal Algebra, by G. Gratzer (4) and Cylindric Algebras, Part I, by L. Henkin, J. D. Monk, and A. Tarski (5), which are more than adequate for the task at hand. After the preliminaries, I first partially characterize a set of reductions by showing that a complete set of reductions constitutes an algorithmic realization of the canonical homomorphism from the absolutely free algebra of terms to the free algebra on a countable number of generators for the class of algebras which satisfy the rewrite rules that form the complete set of reductions. Because complete sets of reductions do not exist for all axiom sets, this partial characterization is clearly insufficient. To complete the characterization, I show that the free algebra on a countable number of generators for the class of algebras

1. This is not entirely accurate. At the very least one must free the diamond lemma of the method used to establish uniform termination, which had in fact been done by Newman in 1942. And I still have not been able to determine if the unique termination theorem of Knuth and Bendix is the same as the one in "Canonical inference."

which satisfy the rewrite rules of a set of reductions is the homomorphic image of the algebra that is the homomorphic algorithmic realization by the set of reductions of the absolutely free algebra of terms. I then employ this algorithmic realization of an intermediate algebra between the absolutely free algebra of terms and the free algebra as the basis of a general method for constructing decision procedures for equational theories which contain commutative axioms. Finally, I apply the method to obtain decision procedures for the commutative theory of groups and the commutative theory of rings, and indicate how this approach may be used to obtain decision procedures for other well-known commutative theories, such as Boolean algebras and modules over commutative rings.

In Section 4, I outline a new complete restriction of paramodulation that I call derived reduction, which is essentially the process that Knuth and Bendix (7) used to derive complete sets of reductions from incomplete sets. Their process generates inferences from the axiom set until a closure condition is satisfied, or until an equation is inferred which cannot be expressed as a rewrite rule without destroying finite termination. Derived reduction uses the same generating process until equations which destroy finite termination are inferred, at which point derived reduction continues, using bidirectional substitution of equals into and by the equations which are not representable as rewrite rules. Even though a set of equations may not be extendable to a complete set of reductions, derived reduction is refutation complete. Incidentally, as a corollary to the refutation

completeness of derived reduction, it follows that the functional reflexive axioms are not needed; and, in addition, substitution into variable positions need never be allowed. The proof of refutation completeness is not contained in this paper, but will be the subject of a subsequent article.[2] Here I am mainly concerned with providing an adequate description of derived reduction, through examples, for design and implementation of practical theorem provers for equality.

2. This proof was to have been based on the solution to the functional reflexive problem (10), which contained an error that the author and others were unable to repair. The case when all equality literals occur as units and is discussed in "Canonical inference." Univ. of Texas, Math. Dept., Automatic Theorem Proving Project, Austin, Texas 78712, report ATP-32, December 1975. So, the reduction methods solve the functional reflexive problem when equations occur as units, but it remains open whether the rewrite rule methods can be extended to the non-unit case without the functional reflexive axioms. The functional reflexive problem has been reported solved for ordinary paramodulation in "Proving theorems with the modification method." by D. Brand, Siam J. Comput. 4, 4 (Dec. 1975), 412-430. But we have been unable to understand that development. In any case, the method of proof there seems substantially different from what is needed to extend the rewrite rule methods. An open problem which might immediately shed some light on the rewrite rule question is M. Richter's conjecture of the ground refutation completeness of uniform substitution. Uniform substitution differs from ordinary substitution in that when a substitution is done, not just one occurrence, but all occurrences (throughout the entire clause being substituted into) must be replaced. And if in turn the refutation completeness of ground, uniform, unidirectional substitution could be shown, then the rewrite rule methods could be established for the general case.

2. REDUCTION

A primary motive for using reduction is the desire to replace the equational part of a theory by a decision procedure so that no extraneous inferences with equality are ever made. Although the existence of unsolvable word problems precludes the full achievement of this objective, the soundness of this approach is supported by encouraging experimental results, namely, the widespread application of ad hoc simplification routines as a utilitarian part of the inference process.

The most conspicuous feature of these ad hoc simplification routines is the use of the equational part of a theory as a set of rewrite rules. For the purposes of this informal preliminary exposition, let us assume that our only interest is in proving theorems about groups. Of course, any other well-known algebraic structure could be used in place of group theory in this account, and the reader should find no difficulty in creating alternate examples to suit his taste. From the equivalent axiomatizations of the equational theory of groups, I have arbitrarily selected the so-called right group axioms, $x + 0 = x$, $x + (-x) = 0$, and $(x + y) + z = x + (y + z)$, for my foundation. A set of rewrite rules for group theory is these axioms (or, as we shall see later, a finite set of consequences of these axioms) restricted to unidirectional substitution of equals. Since each equation has two sides, each equation potentially represents two rewrite rules,

for example, $x + 0 = x$ is potentially $x + 0 \longrightarrow x$ or $x \longrightarrow x + 0$, where the arrow indicates that when inferences are formed, the right side replaces the left, and not vice versa. The process of determining which of the two rewrite rules related to a given axiom to choose for the construction of an ad hoc simplification routine is related to the finite termination problem which will be discussed later. So, for the present, let us assume we are given some criterion which determines for each axiom one of the rewrite rules associated with it; and, furthermore, let us assume that this criterion has selected $x + 0 \longrightarrow x$, $x + (-x) \longrightarrow 0$, and $(x + y) + z \longrightarrow x + (y + z)$.

Another intrinsic feature of these simplification routines is that the rewrite rules are successively applied to an expression (term, atom, literal, clause, etc.) until a corresponding "canonical" expression is obtained. To the purist, whether the output of a simplification routine is canonical or not depends upon whether or not the routine realizes a decision procedure for the underlying rewrite rules. But since an ad hoc simplification routine can be idealistically regarded as a heuristic approximation to a decision procedure for the underlying rewrite rules, it is not unreasonable to call its output canonical. To clarify what is meant above by "successively applied to an expression", I will briefly outline one of several possible constructions of an ad hoc simplification routine for the above three rewrite rules of group theory. An <u>immediate reduction</u> I of an expression E (relative to a set of rewrite rules $\mathcal{R}$) is the result of replacing one occurrence of $L\theta$ in E by $R\theta$, where θ is any substitution and $L \longrightarrow R$ is a rewrite rule (of $\mathcal{R}$). It should be noticed that immediate reduction does not depend upon

unification in the full sense of the word, but rather matching of the left side of a rewrite rule with some subterm of an expression. For example, the only immediate reductions of $(x + 0) + (-(x + 0))$ are $x + (-(x + 0))$, $(x + 0) + (-x)$, 0, and $x + (0 + (-(x + 0)))$. As has been said, an ad hoc simplification routine successively produces immediate reductions until a canonical expression is obtained. But in view of the above example, since there are in general many immediate reductions of a given expression, the simplification routine must be designed so that only one immediate reduction is associated with each expression.

One approach to extracting a function from the relation of reduction is to arbitrarily order the rewrite rules so that the routine produces immediate reductions first by the first rewrite rule, or if the first does not apply then by the second, and so on. For our present description, let us choose the order in which the rewrite rules appear above. This still does not entirely solve the function extraction problem, since, as is seen in the example above, a single rewrite rule may produce more than one immediate reduction of a given expression, e.g., $x + (-(x + 0))$ and $(x + 0) + (-x)$ are immediate reductions of $(x + 0) + (-(x + 0))$ by $x + 0 \longrightarrow x$. To complete the function extraction process, let us assume the left-most application of the rewrite rule is distinguished, for example, $x + (-(x + 0))$ is <u>the</u> immediate reduction associated with $(x + 0) + (-(x + 0))$ by the ad hoc simplification routine just completed.

My expository construction of an ad hoc simplification routine above has overlooked two central design problems for reduction routines, finite termination and unique termination. The finite termination problem is intimately related to the earlier problem of determining which of the two potential rewrite rules to associate with a given axiom. For example, it can easily be seen by length considerations that an arbitrary choice, such as $x \longrightarrow x + 0$, leads to infinite sequences of immediate reductions, e.g., $0 \longrightarrow 0 + 0 \longrightarrow 0 + (0 + 0) \longrightarrow \ldots$, and in such a case the ad hoc routine doesn't halt. Since a new general approach to the finite termination problem will presently be described, let us temporarily defer a deeper investigation. To illustrate the unique termination problem, let us take as an example the equation $-0 = 0$ which is a consequence of the group axioms. If the ad hoc simplification routine described above were a decision procedure for group theory then the canonical outputs for the terms -0 and 0 would be identical, which is not the case (since -0 and 0 have no immediate reductions, the canonical outputs are -0 and 0). Knuth and Bendix (7) have discovered a necessary and sufficient closure condition for such routines which possess the finite termination property to have the unique termination property, and I will presently elaborate on one version of their test. It should be recognized that because of unsolvable word problems, for some theories one or both of the finite and unique termination problems must be unsolvable. However, the finite termination problem can always be solved heuristically by the introduction of a depth parameter, and even if the routine does not have the unique termination property the credence is that

simplification is still useful since it not only incorporates a number of equality inferences into a single inference but also obviates retention of the intermediate inferences.

My fundamental goal in this section is to formalize the concept of reduction described above and to present a new approach to the finite termination problem. I cannot claim any great originality for my formalization, since it consists of a redevelopment and enlargment, a merging, if you will, of the concepts of Knuth and Bendix (7) and Slagle (15). For a thorough understanding of all proofs of this section, familiarity with Knuth and Bendix (7) is helpful. However, since the new approach to finite termination is independent of their results, detailed knowledge of their article is not an absolute prerequisite. For the practical minded reader, the theorems of this section can be comprehended without an understanding of their proofs, and therefore easily applied to the design and construction of practical simplification routines.

The term structure of the first order predicate calculus can be inductively defined as follows. The symbols of the language consist of: a countable number of <u>variable symbols</u> $v_1, v_2, v_3, \ldots$; <u>constant symbols</u> $c_1, c_2, c_3, \ldots$; and <u>function symbols</u> $f_1, f_2, f_3, \ldots$, where each function symbol is assigned a positive integer d_i called the <u>degree</u> of f_i . The set T_0 is the set of all variable symbols and constant symbols. The set T_{i+1} is the set T_i together with all strings of the form $f_j(t_1,\ldots,t_{d_j})$, where $t_1, \ldots$, and t_{d_j} are members of T_i . <u>Terms</u> are just those strings which are members of

T_i for some non-negative i. I will often use other symbols in place of variable, constant, and function symbols, and more readable forms for terms, like $(x + 0) + (-(x + 0))$ instead of $f_1(f_1(v_1,c_1),f_2(f_1(v_1,c_1)))$.

A set of reduction relations is a finite non-empty set of objects of the form $L \longrightarrow R$ where L and R are terms.[3] A term u is an immediate reduction of a term t (with respect to a fixed set of reduction relations $\mathcal{R}$), denoted $t \longrightarrow u$, in case there exists a member $L \longrightarrow R$ of $\mathcal{R}$ and a substitution θ such that u is the result of replacing one occurrence of $L\theta$ in t by $R\theta$. It is convenient to think of immediate reduction as a form of paramodulation restricted to unidirectional substitution, i.e., $L\theta$ does not replace $R\theta$, and one way unification, i.e., $L\theta$ occurs in t. A set of reduction relations $\mathcal{R}$ has the finite termination property in case for each term t there exists no infinite sequence $t \longrightarrow t_1 \longrightarrow t_2 \longrightarrow \ldots$ of immediate reductions. A set of reductions[4] is a set of reduction relations which has the finite termination property. A term u is a reduction of a term t, denoted $t \longrightarrow\!\!* u$, in case $t \longrightarrow t_1 \longrightarrow \ldots \longrightarrow t_n \longrightarrow u$. Included in the definition of reduction are the cases $t \longrightarrow\!\!* u$ when $t \longrightarrow u$, and $t \longrightarrow\!\!* t$ when t is irreducible, that is, when t has no immediate reductions. Thus, $\longrightarrow\!\!*$ is the reflexive, transitive completion of $\longrightarrow$. A set of reductions has the unique termination property in case $t \longrightarrow\!\!* u$ and $t \longrightarrow\!\!* v$, with u and v irreducible, implies that u and v are identical terms. A set of reductions is complete when it has the unique termination property.

3. Because we will consider reduction relations with the finite termination property, we usually assume that each variable symbol which occurs in R also occurs in L.
4. Thus, by definition a set of reductions has the finite termination property.

A central result of Knuth and Bendix (7) is a closure condition which, when applied to a set of reductions, determines whether or not it is complete. Let us now direct our attention to an extension of that test to this more general concept of reduction. The only proof in their development which needs modification is that of the following lattice condition.

Theorem 1 Let $\mathcal{R}$ be a set of reductions. $\mathcal{R}$ is complete iff when $t \longrightarrow u$ and $t \longrightarrow v$ there exists a term w such that $u \longrightarrow^* w$ and $v \longrightarrow^* w$.

Proof Let $\mathcal{R}$ be a complete set of reductions and let $t \longrightarrow u$ and $t \longrightarrow v$. Since $\mathcal{R}$ has the finite termination property, there are irreducible terms i_u and i_v which satisfy $u \longrightarrow^* i_u$ and $v \longrightarrow^* i_v$. Because $\mathcal{R}$ is complete, i_u and i_v are identical terms.

Conversely, suppose that when $t \longrightarrow u$ and $t \longrightarrow v$ there exists a term w such that $u \longrightarrow^* w$ and $v \longrightarrow^* w$. The crux of the matter is that, because of the finite termination property, the depth of the immediate reduction tree of each term is finite. We therefore induct on the depth of the immediate reduction tree of t. For depth 1, u and v are irreducible, hence, by the lattice condition, identical. For the induction step, let the depth be $k + 1$. Since the immediate reduction trees of u and v have depth less than or equal to k, it follows by the induction hypothesis that u and v have the unique termination property. By the

lattice condition, there exists a term w such that $u \longrightarrow^* w$ and $v \longrightarrow^* w$. By finite termination, let $w \longrightarrow^* i_w$ with i_w irreducible. Notice that any irreducible reduction of t that is also a reduction of u and of v is identical to i_w. Since the choice of u and v was arbitrary, it follows that any two irreducible reductions of t are identical terms.

The completeness test of Knuth and Bendix (7) is phrased in terms of concepts which are not widely known to workers in the field of computational logic, so I have translated their test into a more traditional form below.

Theorem 2 Let $\mathcal{R}$ be a set of reductions, let $*$ be any simplification algorithm, that is, any algorithm which associates with each term t a corresponding $\mathcal{R}$-irreducible term t^*, such as the one extracted from the immediate reduction relation for the ad hoc simplification routine above, and let $P(\mathcal{R})$ be the set of all paramodulants $t = u$ of pairs of reductions $L_1 \longrightarrow R_1$ and $L_2 \longrightarrow R_2$ of $\mathcal{R}$ such that $t = u$ is obtained from $L_2 = R_2$ by unifying L_1 on a subterm of L_2 which is not a variable, and replacing the unified occurrence of L_1 in L_2 by R_1.

$\mathcal{R}$ is complete iff t^* and u^* are identical terms for each member $t = u$ of $P(\mathcal{R})$.

Proof When the concept of superposition is rephrased in terms of the

concept of paramodulation, this proof is essentially like the proof to the Corollary to Theorem 5 of Knuth and Bendix (7)[5], and so the details are omitted.

Example 1 For small sets of reductions, the above test can be carried out by hand, so let us apply the test to two sets of reductions, one complete and one incomplete.

For the former, let $\mathcal{R}$ consist of the three reductions

R1. $x + (-x) \longrightarrow 0$,

R2. $(-x) + x \longrightarrow 0$, and

R3. $-(-x) \longrightarrow x$.

It can easily be seen by length considerations that this $\mathcal{R}$ has the finite termination property. By inspection of cases, it can be verified that $P(\mathcal{R})$ consists of

P1. $0 = 0$, by R1 and R1,

P2. $(-x) + x = 0$, by R1 and R3,

P3. $0 = 0$, by R2 and R2,

P4. $x + (-x) = 0$, by R2 and R3,

P5. $x = x$, by R3 and R3, and

P6. $-x = -x$, by R3 and R3.

Notice that in the definition of $P(\mathcal{R})$ it is not necessary to keep the paramodulant of a reduction $L \longrightarrow R$ with itself onto L , since the paramodulant is $R = R$, as typified by P1, P3, and P5 above. But since the

5. The cases (interaction and non-interaction) are the same, but the details of the proof seem different, c f., "Canonical inference."

paramodulants of a reduction $L \longrightarrow R$ with itself onto a proper subterm of L must be tested, the weaker definition of $P(\mathcal{R})$ was used for simplicity. Now let $*$ be the simplification algorithm which was constructed for the ad hoc simplification routine earlier. It automatically follows that t^* and u^* are identical for the equations $t = u$ of P1, P3, P5, and P6, and it is easy to check that t^* and u^* are identical for the equations $t = u$ of P2 and P4. Thus, by Theorem 2, this $\mathcal{R}$ is a complete set of reductions.

For the latter, let $\mathcal{R}$ consist of the three reductions of group theory, R1. $x + 0 \longrightarrow x$, R2. $x + (-x) \longrightarrow 0$, and R3. $(x + y) + z \longrightarrow x + (y + z)$ For the moment, let us assume that this $\mathcal{R}$ has the finite termination property, and let $*$ be the same simplification algorithm as above. This $\mathcal{R}$ is incomplete because $0 + z = x + ((-x) + z)$ is a paramodulant of R1 and R3 on the subterm $x + y$ of R3, and $(0 + z)^*$, which is $0 + z$, and $(x + ((-x) + z))^*$, which is $x + ((-x) + z)$, are not identical terms.

Let us agree that a set of reductions $\mathcal{R}$ satisfies the closure condition relative to a simplification algorithm $*$ in case $P(\mathcal{R})$ has the property that t^* and u^* are identical terms for each equation $t = u$ of $P(\mathcal{R})$. As is shown by Theorem 2, the closure condition solves the unique termination problem, but it is necessarily dependent upon finite termination. So for the closure test to be useful there must be available some general, computationally efficient test for finite termination. One such finite termination test, based on length considerations, has already been alluded to in Example 1. Unfortunately, that test fails for some important axioms, such as associative

axioms. A more general method, which accepts associative axioms, has been investigated by Knuth and Bendix (7). It is based on a class of partial orderings $>$, of the set of terms, which have three properties: (1) is a total ordering of the ground terms, (2) $t > u$ implies $t\theta > u\theta$ for any substitution θ , and (3) if $t > u$ and w is the result of replacing one occurrence of t in v by u then $v > w$. Properties (2) and (3) above guarantee that when $L > R$ for each $L \longrightarrow R$ of $\mathcal{R}$ it necessarily follows that $t \longrightarrow u$ implies $t > u$. This, together with property (1) above, insure finite termination. While the adequacy of their approach is supported by examples, including a solution to the word problem for free groups on no generators and no relations, I have established the following characterization of the finite termination property, and developed a new test for finite termination which seems conceptually more transparent.

Theorem 3 A set of reduction relations $\mathcal{R}$ has the finite termination property iff there exists a function $\|\cdot\|$ defined on all terms, whose range is a subset of the non-negative integers, and which satisfies $t \longrightarrow u$ implies $\|t\| > \|u\|$.

Proof The reverse implication is clearly true, so let $\mathcal{R}$ be a set of reductions, and let $\|t\|$ be the maximum of the sequences of immediate reductions originating with t . Clearly, if $t \longrightarrow u$ then $\|u\| \leq \|t\| - 1$, and so $\|t\| > \|u\|$.

Theorem 3 suggests the desirability of general classes of functions which have the property $t \longrightarrow u$ implies $\|t\| > \|u\|$, and the next theorem helps establish one such class.

Theorem 4 Let $\mathcal{R}$ be a set of reduction relations, let $f_1, \ldots, f_i$ be the function symbols which occur in terms of $\mathcal{R}$, let $F_1, \ldots, F_i$ be functions from the non-negative integers to the non-negative integers such that

(1) the degree of each F_j is the same as the degree of f_j, and

(2) $F_j(x_1,\ldots,x_k,\ldots,x_{d_j}) < F_j(x_1,\ldots,y,\ldots,x_{d_j})$ when $x_k < y$, for $j = 1,\ldots,i$,

and let $\|\cdot\|$ be a non-negative, integer-valued function on the variable and constant symbols which satisfies

(3) $\|f_j(t_1,\ldots,t_{d_j})\| = F_j(\|t_1\|,\ldots,\|t_{d_j}\|)$, for $j = 1,\ldots,i$.

Extend this function to a function on all terms by

(4) $\|f(u_1,\ldots,u_n)\| = \|u_1\| + \ldots + \|u_n\|$ when f is not among the f_j.

If $\|L\theta\| > \|R\theta\|$ for all substitutions θ and all $L \longrightarrow R$ in $\mathcal{R}$ then $\mathcal{R}$ is a set of reductions.

Proof Given $t \longrightarrow u$, it is demonstrated below by induction on the term structure of t that $\|t\| > \|u\|$. When t is a variable or constant symbol, $t \longrightarrow u$ is simultaneously the notation for "u is an immediate reduction of t" and a member of $\mathcal{R}$. Since $\|t\theta\| > \|u\theta\|$ by hypothesis,

using the empty substitution for θ establishes $\|t\| > \|u\|$. For the induction step, let t be $f(t_1,\ldots,t_m)$. On the one hand, if $t \longrightarrow u$ is obtained by applying a reduction relation to one of the t_k of $f(t_1,\ldots,t_m)$ then $f(t_1,\ldots,t',\ldots,t_m) \longrightarrow f(t_1,\ldots,u',\ldots,t_m)$ with $t' \longrightarrow u'$ represents $t \longrightarrow u$. It follows from the induction hypothesis that $\|t'\| > \|u'\|$. By considering the two cases, f is among the $f_1, \ldots, f_i$, and f is not among them, it is clear that $\|t\| > \|u\|$. On the other hand, if $t \longrightarrow u$ is $L\theta \longrightarrow R\theta$ for some θ then $\|t\| > \|u\|$ by hypothesis. Thus, by Theorem 3, $\mathcal{R}$ is a set of reductions.

A rich supply of norm functions $\|\cdot\|$ can easily be generated by selecting the F_j to be polynomial functions in a finite number of arguments over the non-negative integers, whose domains are restricted to the non-negative integers. Formally, a polynomial in a finite number of arguments over a ring is a function F defined by a rule of the form $F(x_1,\ldots,x_m) = \sum_i r_i \prod_j s_{ij}$ where $\prod_j s_{ij}$ are products of the x_k and r_i are members of the underlying ring. For example, $F(x_1,x_2) = 3x_1 + 5x_2 + 7x_1x_2 + 9x_1^2x_2^3$ is a polynomial over the integers in x_1 and x_2. Since condition (2) of Theorem 4 is easily established by induction for all such polynomials, it follows that a corresponding norm function $\|\cdot\|$ can always be constructed inductively by $\|f_j(t_1,\ldots,t_{d_j})\| = F_j(\|t_1\|,\ldots,\|t_{d_j}\|)$ and by $\|f(u_1,\ldots,u_n)\| = \|u_1\| + \ldots + \|u_n\|$ when f is not among the f_j. Thus, for this class of polynomial-based norm functions, if the reduction condition, $\|L\theta\| > \|R\theta\|$ for any substitution θ and any $L \longrightarrow R$ of $\mathcal{R}$, is satisfied then finite termination follows by Theorem 4. This still does not really show that

any such norm functions exist, because a set of reductions does not necessarily satisfy the reduction condition. However, examples are easily constructed, and I will illustrate a solution of the finite termination problem by applying the reduction condition to the polynomial functions below for the three group reductions.

Example 2 Given the three group reductions

$$R1.\quad x + 0 \longrightarrow x\ ,$$

$$R2.\quad x + (-x) \longrightarrow 0\ ,\ \text{and}$$

$$R3.\quad (x + y) + z \longrightarrow x + (y + z)\ ,$$

let F_+ and F_- be the polynomials defined by

$$F_+(x_1,x_2) = 2x_1 + x_2 + 1\ ,\ \text{and}$$

$$F_-(x) = x\ ,\ \text{and}$$

let the norm function $\|\cdot\|$ be defined as above by

$$\|0\| = 1\ ,$$

$$\|v\| = 1 \quad \text{for any variable symbol } v\ ,$$

$$\|t_1 + t_2\| = 2\|t_1\| + \|t_2\| + 1\ ,\ \text{and}$$

$$\|-t\| = \|t\|\ .$$

The reduction condition is verified by checking that

$$(2i + 1) + 1 > i\ ,$$

$$(2i + i) + 1 > 1\ ,\ \text{and}$$

$$2(2i + j + 1) + k + 1 > 2i + (2j + k + 1) + 1$$

for any positive integers i, j, and k.

The results of this section by no means comprise a complete and perfect practical theory of reduction. While a computationally satisfactory test for unique termination has been found, only necessary and not sufficient conditions for finite termination have been given. First of all, both the approach to finite termination of this paper and the approach of Knuth and Bendix (7) assume an arbitrary initial choice of one of the two potential reductions associated with each axiom of a given theory. In the approach of this section, a second arbitrary choice occurs in the selection of the polynomials to associate with the function symbols of the reductions. An analogous arbitrary second choice is assumed in the approach of Knuth and Bendix (7). Now this first arbitrary choice is amenable to computation, since a finite set of equations has finitely many sets of potential reductions. The second is apparently not amenable to computation, since here there are infinitely many sets of corresponding polynomials; and, in the case of Knuth and Bendix (7), infinitely many sets of corresponding weights of function symbols. Ideally, what would be desired in each case is some criterion whereby only a finite subset of the infinite set need be tested. But for the present, trial and error remains a quintessential aspect of the approaches to finite termination.

3. SOME ALGEBRAIC PROPERTIES OF REDUCTION AND THEIR APPLICATIONS

To further delineate the role of reduction in computational logic, I have scanned the mathematical literature in search of mathematical structures to use for the framework of a description of reduction as a decision procedure for the underlying theory.[8] As it turned out, some of the mathematical theory needed for a deeper probe into the nature of reduction has been developed as parts of studies of universal algebra and cylindric algebra. Since a complete development of the necessary parts of those areas is beyond the scope of this paper, I must rely upon the reference texts of Gratzer (4) and Henkin, et al. (5).

A type of algebras τ is a finite sequence, $\tau_1, \ldots, \tau_m$, of non-negative integers. An algebra of type τ is a pair of sets A, F, where F is a finite set $\{f_1,\ldots,f_m\}$ of functions such that the domain of f_i is A^{τ_i}, the set of all τ_i-tuples of A, and the range of f_i is a subset of A. Thus, each f_i is a function of τ_i arguments, and implicit in the definition above is that f_i is a member of A when $\tau_i = 0$. As an example, let us construct the family of absolutely free term algebras, which will be the focus of later attention. To that end, let us define a first order predicate calculus of type τ to be a language based on a countable number of variable symbols and a finite number of constant and function symbols, $s_1, \ldots, s_m$, where the degree of each s_i is τ_i. In other words, this is just a first order predicate calculus with a degree function τ that has a finite domain. The set of all terms T_τ has a natural structure as an

8. Throughout we use an informal notion of algorithm. The normal algorithms of Markov (11) provide a theoretical model when reductions consist only of the associative reduction and ground reductions. However, it is not clear whether normal algorithms are equivalent to reduction systems (though we are rather certain that they are).

algebra of type τ , called the absolutely free algebra of terms, when functions are defined by

$$f_i(t_1,\ldots,t_{\tau_i}) = s_i(t_1,\ldots,t_{\tau_i}) .$$

An algebra homomorphism of type τ is a function θ between algebras A , F and B , G of the same type τ which satisfies

$\theta(f_i(a_1,\ldots,a_{\tau_i})) = g_i(\theta(a_1),\ldots,\theta(a_{\tau_i}))$, when $\tau_i \neq 0$,and

$\theta(f_i) = g_i$, when $\tau_i = 0$.

The lemma below, which will be used later, serves to illustrate these ideas.

Lemma 1 If T_τ is the absolutely free algebra of terms, A , F is an algebra o type τ , and I is any function from the variable symbols of T_τ to A then I extends uniquely to a homomorphism θ from T_τ to A .

Proof Let θ be defined inductively by indetifying the constants of T_τ with the corresponding 0-ary functions of A , and by

$$\theta(s_i(t_1,\ldots,t_{\tau_i})) = f_i(\theta(t_1),\ldots,\theta(t_{\tau_i})) ,$$

where, for clarity, I have used the function symbol s_i to denote the corresponding function of the algebral T_τ .

The word problem can be couched in algebraic terms as follows. Suppose we are given a finite set of equations E which are taken as axioms of a theory. Let S be the set of function symbols which occur in E and let T , S be the associated absolutely free algebra of terms. A congruence relation C on an algebra A , F of type τ is an equivalence relation on A which satisfies the substitution property: for $j = 1,..,m$, if a_{1i} C a_{2i}, for $i = 1,\ldots,$ j then $f_j(a_{11},\ldots,a_{1\tau_j})$ C $f_j(a_{21},\ldots,a_{2\tau_j})$. Let us assume

familiarity with the usual notions of[6] structure, interpretation, satisfiability, truth, and model for first order theories. Let $C(E)$ denote the relation on T_τ defined by

$t\ C(E)\ u$ iff $t = u$ is true in every equality model of E.

Because of the Gödel completeness theorem[7] for first order languages, an equivalent definition of $C(E)$ is

$t\ C(E)\ u$ iff $t = u$ is a consequence of E.

It can easily be checked that $C(E)$ is a congruence relation on T_τ, S. My intention is to construct the quotient algebra related to $C(E)$. In general, given an algebra A, F and a congruence relation C on A, let A/C denote the set of all equivalence classes, that is, the family of all sets $C(a) = \{x \mid x\ C\ a\}$. A/C has a natural algebraic structure when functions $\mathcal{F} = \{F_1,\ldots,F_m\}$ are defined by

$$F_i(C(a_1),\ldots,C(a_{\tau_i})) = C(f_i(a_1,\ldots,a_{\tau_i})).$$

The quotient algebra of A, F relative to C is the algebra A/C, $\mathcal{F}$ above. Associated with the quotient algebra is the canonical homomorphism θ from A to A/C which is defined by

$$\theta(a) = C(a).$$

The quotient algebra $T_\tau/C(E)$, $\mathcal{F}$ is called the free algebra for the equational theory E. Its connection with the word problem is that $\theta(t) = \theta(u)$ in $T_\tau/C(E)$ iff $t = u$ is a consequence of E, where is the canonical homomorphism. Thus, the word problem is solvable in case there is an algorithmic realization of the canonical homomorphism from the absolutely free algebra of terms to the free algebra for the theory.

6. consequence
7. I am now no longer sure who to give credit for this result. Walter Taylor in his "Equational logic" survey in the Houston Journal of Mathematics (1979) credits: Birkhoff, G. "On the structure of abstract algebras." Proc. Cambr. Philos. Soc. 31 (1935), 433-454.

My primary objective in this section is to characterize a set of reductions in the algebraic setting described above. As a first approximation, I show that a complete set of reductions $\mathcal{R}$ may be viewed as an algorithmic realization of the canonical homomorphism from the absolutely free algebra of terms to the free algebra for the equational theory $E(\mathcal{R})$, where $E(\mathcal{R})$ denotes the equations $L = R$ such that $L \longrightarrow R$ belongs to $\mathcal{R}$. Next I will show that the free algebra for an equational theory E is isomorphic to the free algebra on a countable number of generators for the class of algebras which are models of E. Through a well-known result of Garrett Birkhoff on equational definability, this characterization leads to a useful tool for identifying unsolvable problems for complete sets of reductions. Birkhoff's condition for equational definability is that the class of models of the theory must be closed under the formation of subalgebras, homomorphic images, and direct products. As an application of this tool, a complete set of reductions does not exist for fields or integral domains, since those classes are not closed under direct products. In addition, a complete set of reductions does not exist for some equational definable algebras, so it is highly desirable to have a mental model of an incomplete set of reductions. I provide such a model by showing that the free algebra for the equational theory $E(\mathcal{R})$ is the homomorphic image of the algebra that is the homomorphic algorithmic realization by $\mathcal{R}$ of the absolutely free algebra of terms. I then employ this model as the basis of a general method for constructing decision procedures for equational theories which contain commutative axioms. Finally, I apply the method to obtain decision procedures for the commutative theory of groups and rings, and indicate how this approach may be used to obtain

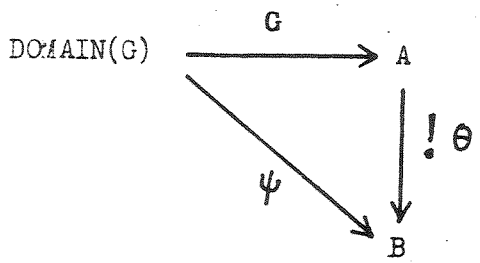
decision procedures for other well-known commutative theories, such as Boolean algebras and modules over commutative rings.

Theorem 5 If $\mathcal{R}$ is a complete set of reductions, and $*$ is any simplification algorithm, that is, any algorithm which associates with each term t a corresponding $\mathcal{R}$-irreducible term t^*, then $C(E(\mathcal{R}))(t) = C(E(\mathcal{R})(u)$ in T_{Γ}, S iff t^* and u^* are identical terms.

Proof If $C(E(\mathcal{R}))(t) = C(E(\mathcal{R}))(u)$ then $t\ C(E(\mathcal{R}))\ u$, that is, $t = u$ is a consequence of $E(\mathcal{R})$. That t^* and u^* are identical is now show by induction on the consequence structure. The initial consequences of $E(\mathcal{R})$, which are the equations of $E(\mathcal{R})$ and substitution instances of them and $x = x$, can be shown to satisfy the identical term condition. For the induction step, the immediate consequences are (1) $u = t$, if $t = u$ was a previous consequence, (2) $t = v$, if $t = u$ and $u = v$ were previous consequences, and $f(t_1,\ldots,t_k,\ldots,t_m) = f(t_1,\ldots,u,\ldots,t_m)$, if $t_k = u$ was a previous consequence. For (1), u^* and t^* are identical because, by the induction hypothesis, t^* and u^* are identical. For (2), use a similar argument. For (3), by the induction hypothesis t_k^* and u^* are identical; hence, by unique termination, $(f(t_1,\ldots,t_k,\ldots,t_m))^*$ and $(f(t_1,\ldots,u,\ldots,t_m)^*$ are identical.

Conversely, if t^* and u^* are identical terms then clearly $t = u$ is a consequence of $E(\mathcal{R})$, so $t\ C(E(\mathcal{R}))\ u$.

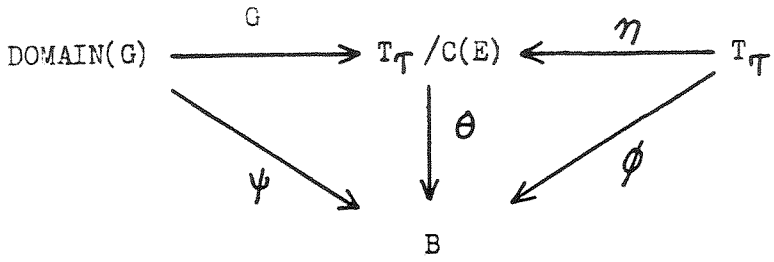
The concept of a free algebra on a countable number of generators which satisfies a set of equations is included in a more general notion of free algebra below, which I have taken from Gratzer (4). Let $\mathcal{K}$ be a class of algebras of a fixed type, let A be a member of $\mathcal{K}$, and let G be a function whose range generates A, that is, A is the smallest subalgebra of A which contains RANGE(G). A is said to be a <u>free algebra over $\mathcal{K}$ with generating family G</u> in case for any algebra B of $\mathcal{K}$ and any function ψ from DOMAIN(G) to B, there is a homomorphism θ from A to B such that $\psi = \theta \circ G$, i.e., ψ is the composition of θ with G. It easily follows that the homomorphism θ is unique, so that, in the jargon of universal algebra, the free algebra over $\mathcal{K}$ with generating family G can be depicted by the diagram



which is called the <u>universal mapping property for A</u>. Next, I show that the free algebra for an equational theory E is isomorphic to a free algebra on a countable number of generators over the class $\mathcal{M}$ of models of E by proving that it satisfies the required universal mapping property.

<u>Theorem 6</u> If E is a set of equations then $T_T/C(E)$ satisfies the universal mapping property for the free algebra over the class $\mathcal{M}$ of models of E with generating family $G = \{(v_i, C(E)(v_i)) \mid v_i \text{ is a variable symbol}\}$.

Proof Let B be any member of $\mathcal{M}$ and let ψ be any function from DOMAIN(G) to B . It is convenient to depict the completed proof by the diagram



where η denotes the canonical homomorphism from T_τ to $T_\tau/C(E)$. Clearly G generates $T_\tau/C(E)$. Since DOMAIN(G) is the set of variable symbols, ψ has a unique extension to a homomorphism ϕ by Lemma 1. Thus by the universal mapping property for quotients it follows that there exists a unique homomorphism θ such that $\phi = \theta \circ \eta$. Now it easily follows that $\psi = \theta \circ G$.

The above result is peripheral to my development, but as I have said, it does lead to a helpful test for identifying futile lines of work, namely attempting to find complete sets of reductions when the task is impossible. Even for theories which are equational definable the word problem may be unsolvable, so it is exigent to have guidelines for that situation. Toward that end, let $\mathcal{R}$ be a set of reductions, let $*$ be any simplification algorithm which satisfies

$$(f_i(t_1,\ldots,t_{\tau_i}))^* = (f_i(t_1^*,\ldots,t_{\tau_i}^*))^* ,$$

and let T_τ^* be the image of T_τ under $*$. T_τ^* has a natural algebraic structure of type τ when operations o_i are defined by

$$o_i(t_1,\ldots,t_{\tau_i}) = (f_i(t_1,\ldots,t_{\tau_i}))^* ;$$

and, moreover, $*$ is clearly a homomorphism from T_τ , S onto T_τ^* , O .

Theorem 7 If E is an equational theory, $\mathcal{R}$ is a set of reductions, and $E(\mathcal{R})$ is a set of consequences of E then the free algebra for E is a homomorphic image of $T_\tau{}^*$, O.

Proof It can be seen that $T_\tau/C(E)$ is a homomorphic image of $T_\tau/C(E(\mathcal{R}))$ by showing that $\theta(C(E(\mathcal{R}))(t)) = C(E)(t)$ defines a homomorphism θ onto $T_\tau/C(E)$. That $T_\tau/C(E(\mathcal{R}))$ is a homomorphic image of $T_\tau{}^*$, O follows by observing that $\psi(t) = C(E(\mathcal{R}))(t)$ defines a homomorphism ψ onto $T_\tau/C(E(\mathcal{R}))$.

Theorem 7 forms the basis of an approach to the word problem for theories with commuatative axioms. Given an equational theory E, isolate a subset of E and the well-known consequences of E which can be expressed as reductions, form the intermediate algebra $T_\tau{}^*$, and by some, as yet unspecified, method construct the free algebra for E. The next example will serve to illustrate this approach.

Example 3 The equational theory of commutative groups is the set E consisting of the four axioms

G1. $x + 0 = x$,

G2. $x + (-x) = 0$,

G3. $(x + y) + z = x + (y + z)$, and

G4. $x + y = y + x$.

Let $\mathcal{R}$ be the three reductions

R1. $(x + y) + z \longrightarrow x + (y + z)$,

R2. $-(-x) \longrightarrow x$, and

R3. $-(x + y) \longrightarrow (-x) + (-y)$,

and notice that finite termination is assured by the polynomial functions F_+ and F_- defined by

$$F_+(x_1,x_2) = 2x_1 + x_2 + 1 \text{ , and}$$

$$F_-(x) = 2x \text{ .}$$

An analysis of the structure of $T_T{}^*$ shows that its members are (1) 0 , (2) any variable symbol v_i , (3) $-I$ where I is an irreducible of the first or second kind, and (4) a finite sum of irreducibles of the first three kinds, which I will denote by $\sum$. It is well-known that the free algebra for group theory is isomorphic to the countable direct sum of the group of integers Z , $+$, which I denote by $\bigoplus Z$, $+$ and which is the group of functions from the natural numbers N to Z such that $\{ n \mid f(n) = 0 \}$ is finite, where the group operation is pointwise addition of functions. With the above characterization of the free algebra for group theory as $\bigoplus Z$, $+$, it is easy to "see" the homomorphism from $T_T{}^*$ onto the free algebra for group theory. Let $Occ(e,\Sigma)$ denote the number of occurrences of e in Σ , and define the homomorphism θ from $T_T{}^*$ onto the free algebra by

$$\theta(\Sigma)(n) = \begin{cases} Occ(0,\Sigma) - Occ(-0,\Sigma) & \text{, when } n = 1 \text{ ,and} \\ Occ(v_{n-1},\Sigma) - Occ(-v_{n-1},\Sigma) & \text{, when } n \geq 2 \text{ .} \end{cases}$$

Clearly θ is a homomorphism and onto. In addition, $\bigoplus Z$, $+$ is clearly isomorphic to a subgroup of $T_T{}^*$, so there is a representation of the free algebra within $T_T{}^*$.

To illustrate the process of mapping to $T_\tau{}^*$ described above, let us examine a sample proof of $x + ((-(x + z)) + y) = x + ((-x) + (y + (-z)))$. First, the irreducibles of both sides are produced,

$$x + ((-x) + ((-z) + y)) \quad \text{and} \quad x + ((-x) + (y + (-z))) \ .$$

Mapping to $\bigoplus Z$, $+$ identifies both irreducibles with the same element

$$f(n) = \begin{cases} 0 \ , \text{ if } n = 1 \ , \text{ because there are no occurences of } 0 \ , \\ 0 \ , \text{ if } n = 2 \ , \text{ because the occurrences of } x \text{ and } -x \text{ cancel,} \\ 1 \ , \text{ if } n = 3 \ , \text{ because there is one occurrence of } y \ , \\ -1 \ , \text{ if } n = 4 \ , \text{ because there is one occurrence of } -z \ , \text{ and} \\ 0 \ , \text{ otherwise, because there are no other variable symbols.} \end{cases}$$

The image of this f above back in $T_\tau{}^*$ is $y + (-z)$, so the image of the original equation is $y + (-z) = y + (-z)$.

This example may seem contrived unless I explain my selection of $\mathcal{R}$ and then show how free algebras for other theories can be constructed. The primary impetus for my approach is found in the following assemblage of ideas. By theorem 7 it follows that the free algebra for commutative groups is a homomorphic image of the free algebra for groups. The free algebra for groups has been shown by Knuth and Bendix (7) to be realized by the complete set of reductions which consists of

KB1. $x + 0 \longrightarrow x$,

KB2. $0 + x \longrightarrow x$,

KB3. $x + (-x) \longrightarrow 0$,

KB4. $(-x) + x \longrightarrow 0$,

KB5. $(x + y) + z \longrightarrow x + (y + z)$,

KB6. $-0 \longrightarrow 0$,

KB7. $-(-x) \longrightarrow x$,

KB8. $-(x + y) \longrightarrow (-y) + (-x)$,

KB9. $x + ((-x) + y) \longrightarrow y$, and

KB10. $(-x) + (x + y) \longrightarrow y$.

It can be shown that the homomorphism from the free algebra for group theory to $\bigoplus Z$, + has the same definition as the previous homomorphism θ . By inspection of KB1 - KB10 it can be seen that the effect of the reductions KB1, KB2, KB3, KB4, KB6, KB9, and KB10 is duplicated by θ , which leaves only KB5, KB7, and KB8 as essential reductions. Because the image algebra of the algebra of irreducibles $T_{\mathcal{T}}{}^{*}$ is commutative, I replaced KB8 by R3.

Although I have justified my selection of $\mathcal{R}$ in this case, how should one proceed when the free algebra for a suitable subtheory is not conveniently available? Examination of KB1 - KB10 indicates that many of the reductions of the desired set will correspond to well-known identities of the theory. For theories composed of "random" axioms, the algorithm of Knuth and Bendix (7) should be a powerful aid in the search for the free algebra. Let us trace the "by hand" approach below through a construction of the free algebra for commutative ring theory.

Example 4 Examination of the axioms and elementary identities of commutative ring theory cause me to postulate that the set of reduction relations consisting of

CR1. $x + 0 \longrightarrow x$,

CR2. $0 + x \longrightarrow x$,

CR3. $x + (-x) \longrightarrow 0$,

CR4. $(-x) + x \longrightarrow 0$,

CR5. $(x + y) + z \longrightarrow x + (y + z)$,

CR6. $x \cdot 1 \longrightarrow x$,

CR7. $1 \cdot x \longrightarrow x$,

CR8. $(x \cdot y) \cdot z \longrightarrow x \cdot (y \cdot z)$,

CR9. $x \cdot (y + z) \longrightarrow (x \cdot y) + (x \cdot z)$,

CR10. $(x + y) \cdot z \longrightarrow (x \cdot z) + (y \cdot z)$,

CR11. $-0 \longrightarrow 0$,

CR12. $-(-x) \longrightarrow x$,

CR13. $-(x + y) \longrightarrow (-x) + (-y)$,

CR14. $x \cdot 0 \longrightarrow 0$,

CR15. $0 \cdot x \longrightarrow 0$,

CR16. $x \cdot (-y) \longrightarrow -(x \cdot y)$, and

CR17. $(-x) \cdot y \longrightarrow -(x \cdot y)$

has the finite termination property. In preparation for Theorem 4, let the norm function $\|\cdot\|$ be defined by

$$\|v_i\| = \|0\| = \|1\| = 2 \ ,$$

$$\|t + u\| = 2\|t\| + \|u\| + 3 \ ,$$

$$\|t \cdot u\| = \|t\|^2 \|u\| + 1 \ , \text{ and}$$

$$\|-t\| = 2\|t\| + 1 \ .$$

It is easily checked that the reduction condition of Theorem 4 is satisfied. The irreducibles of T_τ^* in this case can be seen to be (1) 0 , (2) 1 ,

(3) -1 , (4) v_i , (5) $-v_i$, (6) finite products of variable symbols which I will denote by Πv_i , (7) $-\Pi v_i$, and (8) finite sums of irreducibles of types (2) through (7) which I will denote by $\mathcal{S}$. The free algebra for ring theory is well-known to be the polynomials over the ring of integers Z , $+$, in a countable number of variable symbols X_1 , X_2 , X_3 , ... which I denote by $Z(X_i)$. Let me briefly remind the reader of the construction of $Z(X_i)$. For this reminder, let N be the set of natural numbers $\{0 , 1 , 2 , \ldots\}$ and define

$$\mathcal{F} = \{f : N \longrightarrow N \mid \{n \mid f(n) \neq 0\} \text{ is finite}\} .$$

One thinks of $\mathcal{F}$ as being the finite products

$$\prod_{f(n) \neq 0} X_n^{f(n)} .$$

Let $Z(X_i) = \{ F : \mathcal{F} \longrightarrow Z \mid \{f \mid F(f) \neq 0\} \text{ is finite} \}$, and notice that $Z(X_i)$ is a ring when operations $\oplus$ and $\odot$ are defined by

$$(F \oplus G)(f) = F(f) + G(f) \text{ , and}$$

$$(F \odot G)(f) = \sum_{g\#h = f} F(g) \cdot G(h)$$

where $\#$ is defined by

$$(f\#g)(n) = f(n) + g(n) .$$

A decision procedure for $Z(X_i)$ is constructed as follows. Let the finite products ΠX_i of $Z(X_i)$ be identified with corresponding products $c(\Pi X_i)$ containing the same occurrences of symbols in their natural order, e.g., $c(X_1X_3X_1X_2)$ is $X_1X_1X_2X_3$. Since elements of $Z(X_i)$ are finite sums of finite strings $\pm\Pi X_i$, two members of $Z(X_i)$ are equal iff the number of occurrences of products corresponding to $c(\Pi X_i)$ minus the number of occurrences of expressions $-\Pi X_i$ corresponding to $-c(\Pi X_i)$ is identical in both members of each $c(\Pi X_i)$. A decision procedure can now be errected within $T_\pi{}^*$ by

ordering the products $c(\prod X_i)$ through $c(\prod X_i) < c(\prod Y_i)$ iff the length of $c(\prod X_i)$ is shorter than the length of $c(\prod Y_i)$, or at the first position from the left that they differ, the corresponding variable symbols v_j and v_k satisfy $j < k$. Let us illustrate this decision procedure with a sample proof below.

To prove in a commutative ring that $(x + y)^2 = x^2 + 2xy + y^2$, let us assume that the equation is given as

$$(v_1 + v_2) \cdot (v_1 + v_2) = (v_1 \cdot v_1) + ((v_1 \cdot v_2) + ((v_1 \cdot v_2) + (v_2 \cdot v_2)) .$$

The left side of this identity reduces to the sum $\mathcal{S}$, which is

$$(v_1 \cdot v_1) + ((v_2 \cdot v_1) + ((v_1 \cdot v_2) + (v_2 \cdot v_2))) ,$$

in $T_\tau{}^*$, and the right is already irreducible. Both of these map to the same term in the realization of $Z(X_i)$ in $T_\tau{}^*$, namely

$$(v_1 \cdot v_1) + ((v_1 \cdot v_2) + ((v_1 \cdot v_2) + (v_2 \cdot v_2))) .$$

These two examples help us believe that the methods suggested by this section will result in decision procedures for many other well-known commutative theories, such as Boolean algebras and modules over a commutative ring. The methods were developed to overcome the difficulty of expressing a commutative axiom as a reduction, but it should not be overlooked that some more general notion of reduction could not only be acceptable for commutative axioms but also more efficient for computational purposes. However, the fact that some important theories, such as integral domains and fields, do not have decision procedures for their equational parts suggests that methods based on Theorem 7 will continue to be important in that they offer a conceptual guide for the construction of heuristic approximations to unavailable decision procedures.

4. DERIVED REDUCTION

The existence of unsolvable word problems, and the fact that the construction of the free algebra must be done theory by theory are inconveniences which help direct our attention to combining sets of reductions with inference rules in a refutationally complete manner. Some work has been done in this direction by Plotkin (13) and Slagle (15), but since derived reduction is based on ideas of Knuth and Bendix (7), I will not mention their work here. The best use of the algorithm which often derives complete sets of reductions from incomplete sets would seem to be combining that algorithm with a simultaneous search for a refutation, which is what derived reduction does. The refutation completeness of derived reduction, which is related to a solution of the functional reflexive problem by Lankford (10),[9] will be the subject of a subsequent article. Here I am primarily concerned with providing and adequate description of derived reduction, through examples, for design and implementation of practical theorem provers for equality. Also, I will mention a refutationally complete restriction of derived reduction which may significantly improve its effectiveness.

For a thorough discussion of the algorithm, consult Knuth and Bendix (7). I will content myself here with a brief statement of comparison between my approach and theirs. Both algorithms are based on Theorem 2, but my approach treats finite termination through Theorem 4, whereas theirs is dependent upon the class of partial orderings mentioned earlier.

9. Reference (10) did not withstand careful analysis. See page 6 of this paper for comments.

Example 5 Let us assume that our interest is in proving theorems about groups, and, in particular, that we would prove if $x + x = 0$ then the group is commutative. Derived reduction begins in the usual manner with the axiom(s) and the denial of the conclusion,

$$A.\quad x + x = 0 \quad , \text{ and}$$

$$D.\quad a + b \neq b + a \quad .$$

Of course, elimination of quantifiers through Skolemization is assumed. It would be senseless to rederive a complete set of reductions for group theory, so the presence of the reductions KB1 - KB10 is assumed. Furthermore, let us assume the presence of two functions, F_+ and F_-, defined by

$$F_+(x,y) = 2x + y + 1 \quad , \text{ and}$$

$$F_-(x) = 2^x \quad ,$$

and a corresponding norm function $\|\cdot\|$ defined as in Example 2. It is easy to check that the norm function $\|\cdot\|$, by Theorem 4, establishes the finite termination of KB1 - KB10. Now the aim of reduction is to replace the equational part of a theory with a decision procedure, so a reasonable beginning for derived reduction is to attempt to incorporate the axiom(s) as a reduction into the complete set of reductions $\mathcal{R}$ consisting of KB1 - KB10, and to try to extend the resulting set of reductions to a complete set. The first step is easily accomplished by noticing that $i + i > 1$ for any positive integer i, so that the axiom A may be added to $\mathcal{R}$ as

$$R1.\quad x + x \longrightarrow 0$$

without destroying finite termination. Next, the test for unique termination is applied to this new set of reductions $\mathcal{R}_1$.[10] Let me assume the existence

10. $\mathcal{R}_1$ consists of KB1-KB10 and R1.

of an algorithm which produces the members $t = u$ of $P(\mathcal{R})$ of Theorem 2, sequentially testing that[11] t^* and u^* are identical. The algorithm that I have in mind will be reflected in the order that I discuss its outputs below.

O1. $0 = 0$ is produced by KB1 and R1, and so is deleted.

O2. $0 = 0$ is produced by KB2 and R1, and so is deleted.

O3. $0 + z = x + (x + z)$ is produced by KB5 and R1, and since $(0 + z)^*$ is z and $(x + (x + z))^*$ is $x + (x + z)$, the set of reduction $\mathcal{R}_1$ is not complete. The heuristic of Knuth and Bendix (7) is to determine if $z = x + (x + z)$ can be added to the current set of reductions, in this case $\mathcal{R}_1$, without destroying finite termination. In this case,

$$R2.\quad x + (x + z) \longrightarrow z$$

qualifies as a reduction because

$$2i + (2i + j + 1) + 1 > 1$$

for all positive integers i and j. After this is done, the new set of reductions $\mathcal{R}_2$ consisting of $\mathcal{R}_1$ and R2 is tested for completeness. The next reduction generated by this process is

$$R3\quad -x \longrightarrow x\ .$$

After each new reduction is generated, the previous reductions are tested to see if they collapse further under the new reduction. For R2 this did not happen, but for R3 a number of the previous reductions are transformed, resulting in some deletions. For example, KB3 is transformed into $x + x \longrightarrow 0$ which is redundant because of R1, and thus deleted, while KB8 is transformed into $x + y = y + x$. As is seen, this process will sometimes result in equations which cannot be expressed as reductions, but in such a case derived reduction continues, allowing paramodulation into and by both sides of these

11. whether

equations. After the reduction process using R3 is completed, the new set of reductions $\mathcal{R}_3$ is

R'1. $x + 0 \longrightarrow x$,

R'2. $0 + x \longrightarrow x$,

R'3. $(x + y) + z \longrightarrow x + (y + z)$,

R'4. $x + (x + z) \longrightarrow z$, and

R'5. $x + x \longrightarrow 0$,

and the equation list is

E1. $x + y = y + x$.

Let us assume alternate rounds of extending to[12] a complete set followed by a round of the special paramodulation described above. In this case, the first round of paramodulation produces the contradiction

C. $a + b \neq b + a$

by E1 and D.

Example 6 In a group, $x + x = 0$ implies $(x + y) + ((-x) + (-y)) = 0$. This example is like the last, except that the denial of the conclusion is

D. $a + (b + ((-a) + (-b)) \neq 0$.

Derived reduction proceeds as before, until after the production of R3 of Example 5 the denial is replaced by

D'. $a + (b + (a + b)) \neq 0$.

Paramodulating E1 (of Example 5) into D' produces $a + (b + (b + a)) \neq 0$ which reduces by R'4 and R'5 (of Example 5) to the contradiction

C. $0 \neq 0$.

---

12. toward

Example 7 In group, $x + (x + x) = 0$ implies $(((x + y) + ((-x) + (-y))) + y) + (-((x + y) + ((-x) + (-y))) + (-y)) = 0$. This example, introduced by Robinson and Wos (14) for a comparison of resolution and paramodulation, has received repeated attention as a "difficult" theorem for equational provers. Nevins (12), using paramodulation and prohibiting substitution into variable positions, which has been shown complete by Lankford (10), has reported a computer proof in which 415 new formulas were generated. With the methods of this section, a much more efficient proof is believed possible.[13] Again I begin with the ten group reductions, the hypothesis expressed as a reduction

$$\text{R1.}\quad x + (x + x) \longrightarrow 0 \quad ,$$

and the Skolemized, reduced, denial of the conclusion

$$\text{D.}\quad a + (b + ((-a) + (b + (a + ((-b) + ((-a) + (-b))))))) \neq 0 \quad .$$

The first round of reduction production is similar to that of Example 5, resulting in

$$\text{R2.}\quad x + (x + (x + z)) \longrightarrow z \quad .$$

The next round of reduction production results in output

$$\text{O.}\quad -x = x + x \quad .$$

Strictly speaking, this does not qualify as a reduction, since 2^i and $3i + 1$ are incomparable. I could handle this situation through a redefinition of the norm function, but the above output is a definition of the function $-$ in terms of the function $+$, and as such represents a meta-reduction which only needs to be used for one round of reduction. So I will assume that derived reduction is further restricted to identify and use meta-reductions. Other

13. Our belief was subsequently confirmed, see W. W. Bledsoe's "Non-resolution theorem proving," In Proc. IJCAI-75, and also in Artif. Intell. 9 (1977), 1-35.

techniques, similar to meta-reduction, were discovered by Knuth and Bendix (7) as extensions of their basic algorithm. After the round of meta-reduction, the reduction list, equation list, and reduced denial are

$$R'1.\quad x + 0 \longrightarrow x \quad ,$$

$$R'2.\quad 0 + x \longrightarrow x \quad ,$$

$$R'3.\quad (x + y) + z \longrightarrow x + (y + z) \quad ,$$

$$R'4.\quad x + (x + (x + z)) \longrightarrow z \quad ,$$

$$R'5.\quad x + (x + x) \longrightarrow 0 \quad ,$$

$$E1.\quad x + (y + (x + y)) = y + (y + (x + x)) \quad , \text{ and}$$

$$D'.\quad a + (b + (a + (a + (b + (a + (b + (b + (a + (a + (b + b))))))))))) \neq 0 \ .$$

On the second round of derived reduction, E1 paramodulates into D' to yield

$$DR2A.\quad a + (b + (a + (a + (b + (a + (a + (b + a))))))) \neq 0 \quad ,$$

and R'3 into R'5 produces

$$DR2B.\quad x + (y + (x + (y + (x + y)))) \longrightarrow 0 \quad .$$

On the third round R'3 and DR2B produce

$$DR3.\quad x + (y + (z + (x + (y + (z + (x + (y + z))))))) \longrightarrow 0 \quad ,$$

which reduces DR2A to the contradiction

$$C.\quad 0 \neq 0 \quad .$$

Even though these examples make derived reduction seem attractive as a step towards more efficient theorem provers for equality, there are some design problems which have yet to be solved. Implicit in the workings of derived reduction is an algorithmic realization of the finite termination test. As long as the norm function is realized by polynomials, an easy

algorithm exists, based on treating $\|t\|$ and $\|u\|$ of the members $t = u$ of $P(\mathcal{R})$ which are not of the form $x = x$ as elements of a ring, and computing the members $\mathcal{S}$ of $Z(X_i)$ which corresponds to $\|t\| - \|u\|$. If the non-zero coefficients of $\mathcal{S}$ are all negative then $t = u$ corresponds to the reduction $u \longrightarrow t$; if the non-zero coefficients of $\mathcal{S}$ are all positive then $t = u$ becomes $t \longrightarrow u$; otherwise, t and u are incomparable. However, if some of the underlying functions for the norm function are not polynomials then I know of no algorithmic test. I would hope that either effective algorithms for the non-polynomial case could be developed, or perhaps some effective process for identifying when a theory can be treated with polynomials alone. To illustrate my lack of information in this direction, I do not know whether the ten group reductions can be identified as reductions through a norm function based on polynomials. Knuth and Bendix (7) do have an algorithmic realization of the finite termination test which was used to derive the ten group reductions, so it may be that their approach is more favorable for computation purposes.

It might be thought that some of the problems raised in this paper are not very important to computational logicians if it is assumed that unsolvable word problems are pathological creatures far removed from the usual terrain of theorem proving. Early examples of unsolvable word problems would support this view, since they involved large numbers of equations. However, a recent example of a theory with an unsolvable word problem consisting of the associative axiom and seven ground equations, not dissimilar to the defining relations for the group of rigid translations of the square, given by Trakhtenbrot (16), indicates the importance of a firmer foundation for the theory of reduction.

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to Woody Bledsoe for his encouragement and support through the Automatic Theorem Proving Project at the University of Texas. I owe a very special thanks to Mike Ballantyne for introducing me to many of the ideas developed in this paper and for his help in programming derived reduction. I would also like to thank Mark Moriconi for his assisting me with many aspects of computer science and for his help in locating related work.

REFERENCES

1. Bledsoe, W.W., Boyer, R.S., and Henneman, W.H. Computer proofs of limit theorems. Artificial Intelligence 3(1972), 27-60.

2. Chang, C.L. Renamable paramodulation for automatic theorem proving with equality. Artificial Intelligence 1(1970), 247-256.

3. Chang, C.L., and Slagle, J.R. Linear refutation for theories with equality. J. ACM 18, 1(January 1971), 126-136.

4. Gratzer, G. Universal Algebra, D. Van Nostrand Company, Inc., 1968.

5. Henkin, L., Monk, J.D., and Tarski, A. Cylindric Algebras, Part I, North-Holland Publishing Company, 1971.

6. Huet, G.P. Experimental results with an interactive prover for logic with equality. Jennings Computing Center, Case Western Reserve University, report 1106.

7. Knuth, D.E., and Bendix, P.B. Simple word problems in universal algebras. Computational Problems in Abstract Algebra, J. Leech, Ed., Pergamon Press, 1970, pp. 263-297.

8. Kowalski, R.A. The case for using equality axioms in automatic demonstration. Lecture Notes in Mathematics, vol. 125, Springer-Verlag, 1968, pp. 112-127.

9. Lankford, D.S. Equality atom term locking. Ph.D. Dissertation, The University of Texas at Austin, Austin, Texas, 1972.

10. Lankford, D.S. The functional reflexive problem. (To appear in the J.ACM). **(withdrawn, see comments on pages 6 and 36)**

11. Markov, A.A. Theory of Algorithms, published for The National Science Foundation, Washington, D.C., and The Department of Commerce, U.S.A., by The Israel Program for Scientific Translations, 1961.

12. Nevins, A.J. A human oriented logic for automatic theorem proving. J. ACM 21, 4(October 1974), 606-621.

13. Plotkin, G.D. Building-in equational theories. Machine Intelligence 7, B. Meltzer and D. Michie, Eds., Edinburgh University Press, 1972, pp. 73-89.

14. Robinson, G., and Wos, L. Paramodulation and theorem proving in first-order theories with equality. Machine Intelligence 4, B. Miltzer and D. Michie, Eds., Edinburgh University Press, 1969, pp. 135-150.

15. Slagle, J.R. Automated theorem-proving for theories with simplifiers, commutativity, and associativity. J. ACM 21, 4(October 1974), 622-642.

16. Trakhtenbrot, B.A. Algorithms and Automatic Computing Machines, Survey of Recent East European Mathematical Literature, A project conducted by A.A. Putnam and I. Wirszup, Department of Mathematics, The University of Chicago, under a grant from The National Science Foundation, D.C. Heath and Company, 1965.

17. Wos, L., Robinson, G.A., and Carson, D.F. The concept of demodulation in theorem proving. J. ACM 14, 4(October 1967), 698-709.

18. Wos, L., and Robinson, G. Paramodulation and set of support. Proc. IRIA Symposium on Automatic Demonstration, Versailles, France, 1968, Springer-Verlag, 1970, pp. 276-310.