DECISION PROCEDURES FOR SIMPLE EQUATIONAL THEORIES WITH COMMUTATIVE AXIOMS:

COMPLETE SETS OF COMMUTATIVE REDUCTIONS

D. S. Lankford and A. M. Ballantyne

March 1977 ATP-35

DECISION PROCEDURES FOR SIMPLE EQUATIONAL THEORIES WITH COMMUTATIVE AXIOMS: COMPLETE SETS OF COMMUTATIVE REDUCTIONS

by D. S. Lankford
Southwestern University
Mathematics Department
Georgetown, Texas 78626

and A. M. Ballantyne
University of Texas
Mathematics Department
Austin, Texas 78712

ABSTRACT

Complete sets of commutative reductions are defined and a unique termination algorithm is established.

This work was supported in part by NSF Grant #DCR74-12866

March 1977

INTRODUCTION

Commutative axioms cannot be used directly as rewrite rules because they allow infinite sequences of immediate reductions. For example, $f(x,y) \longrightarrow f(y,x)$ successively simplifies $f(a,b)$ to $f(b,a)$, $f(a,b)$, $f(b,a)$, ... What is needed are methods which decide finite and unique termination. In this article we develop techniques for combining commutative axioms with rewrite rules to form decision procedures for some simple equational theories. We assume familiarity with the basic results about complete sets of reductions developed by Knuth and Bendix (1), Lankford (2), and Slagle (3).

COMPLETE SETS OF COMMUTATIVE REDUCTIONS

Let $f_1, \ldots, f_N$ be the function symbols and $v_1, v_2, v_3, \ldots$ be the countable number of variable symbols from which terms are constructed. Constants are function symbols of degree zero. A term is a variable symbol, constant, or expression $f_i(t_1,\ldots,t_{d_i})$ where $t_1, \ldots, t_{d_i}$ are terms and d_i is the degree of f_i. Let f be one of the function symbols of degree 2 and let $\approx$ be the equivalence relation defined by $t \approx u$ iff $t = u$ is a consequence of $f(x,y) = f(y,x)$. The equivalence class $\approx(t)$ is finite for any term t. A commutative rewrite rule is an expression $\approx(L) \longrightarrow \approx(R)$ where L and R are terms. We say $\approx(u)$ is an immediate commutative reduction of $\approx(t)$

by $\approx(L) \longrightarrow \approx(R)$ iff there exist a substitution θ, $t' \in \approx(t)$, $u' \in \approx(u)$, $L' \in \approx(L)$, and $R' \in \approx(R)$ such that u' is the result of replacing one occurrence of L' in t' by R'. A commutative reduction is a finite sequence of immediate commutative reductions. An irreducible equivalence class is one that has no immediate reductions. When U is an immediate commutative reduction of T we write $T \longrightarrow U$. Let $\longrightarrow_c$ be the reflexive, transitive completion of $\longrightarrow$. We say that $\approx(t)$ terminates naturally with $\approx(u)$ in case $\approx(t) \longrightarrow_c \approx(u)$ and $\approx(u)$ is irreducible. A set of commutative rewrite rules $\mathcal{R}$ has the finite termination property iff there is no infinite sequence $\approx(t_1) \longrightarrow \approx(t_2) \longrightarrow \approx(t_3) \longrightarrow \dots$. A set $\mathcal{R}$ of commutative rewrite rules has the unique termination property iff for any $\approx(t)$ and any two naturally terminating sequences $\approx(t) \longrightarrow \approx(u_1) \longrightarrow \dots \longrightarrow \approx(u_m)$ and $\approx(t) \longrightarrow \approx(v_1) \longrightarrow \dots \longrightarrow \approx(v_n)$ of immediate commutative reductions, $\approx(u_m) = \approx(v_n)$. A finite set $\mathcal{R}$ of commutative rewrite rules is a complete set of commutative reductions iff $\mathcal{R}$ has the finite and unique termination properties.

Unique Termination Theorem Let $\mathcal{R}$ be a set of commutative rewrite rules with the finite termination property. To decide whether $\mathcal{R}$ has the unique termination property, perform the following steps. Throughout, unification on variables is not permitted.

(1) For each pair of members $\approx(L_i) \longrightarrow \approx(R_i)$ and $\approx(L_j) \longrightarrow \approx(R_j)$ of $\mathcal{R}$, each L_i' in $\approx(L_i)$, and each L_j' in $\approx(L_j)$, form all paramodulants $x = y$ of $L_i' = R_i$ and $L_j' = R_j$ by left sides into left sides.

(2) For each paramodulant $x = y$ from step 1, fully commutatively reduce $\approx(x)$ and $\approx(y)$ to $\approx(x)^*$ and $\approx(y)^*$.

(3) $\mathcal{R}$ has the unique termination property iff for each paramodulant $x = y$ from step 1, $\approx(x)^* = \approx(y)^*$.

Proof ($\Longrightarrow$) Let $x = y$ be a paramodulant from step 1. It can be shown that there exists an equivalence class $\approx(z)$ such that $\approx(z) \longrightarrow \approx(x)$ and $\approx(z) \longrightarrow \approx(y)$. It follows that $\approx(x)^*$ and $\approx(y)^*$ are equal.

($\Longleftarrow$) This case requires a diamond lemma: if a set of commutative rewrite rules $\mathcal{R}$ has the finite termination property, then $\mathcal{R}$ has the unique termination property iff for each $\approx(t)$ and each pair $\approx(t) \longrightarrow \approx(u)$ and $\approx(t) \longrightarrow \approx(v)$ of immediate commutative reductions of $\approx(t)$, there exists $\approx(w)$ such that $\approx(u) \longrightarrow_c \approx(w)$ and $\approx(v) \longrightarrow_c \approx(w)$. Let t' and t'' be in $\approx(t)$, let $\approx(L_i) \longrightarrow \approx(R_i)$ and $\approx(L_j) \longrightarrow \approx(R_j)$ be in $\mathcal{R}$, let L_i' be in $\approx(L_i)$, let L_j' be in $\approx(L_j)$, let θ_1 and θ_2 be substitutions, let u be the result of replacing one occurrence of $L_i'\theta_1$ in t' by $R_i\theta_1$, and v be the result of replacing one occurrence of $L_j'\theta_2$ in t'' by $R_j\theta_2$. Notice that it is unnecessary to choose members of $\approx(R_i)$ and $\approx(R_j)$ other than R_i and R_j. If t' and t'' are identical, then the methods of Knuth and Bendix (1) and Lankford (2) may be used to complete the proof. If t' and t'' are not identical, then t'' is obtained from t' by a finite number of applications of $f(x,y) = f(y,x)$. Let $t' = t_1, \ldots, t_n = t''$ be the sequence of applications of $f(x,y) = f(y,x)$. Let λ be the substitution such that t_2 is obtained from t_1 by replacing one occurrence of $f(x,y)\lambda$ in t_1 by $f(y,x)\lambda$. If $f(x,y)\lambda$ and $L_i'\theta_1$ do not interact, let u' be the result of replacing that occurrence of $L_i'\theta_1$ in t_2 by $R_i\theta_1$. Since u and u' are in the same equivalence class, we have reduced the problem to the equivalent problem for the shorter deduction of t'' from t_2. If $f(x,y)\lambda$ occurs in $L_i'\theta_1$ in a position that does not correspond to a variable in L_i', then there exists an L_i'' in $\approx(L_i)$ such that $L_i''\theta_1$ is the result of replacing the occurrence of $f(x,y)\lambda$ in $L_i'\theta_1$ by $f(y,x)\lambda$. Here we also have reduced the problem to considering shorter deductions. If $f(x,y)\lambda$ occurs in a position in $L_i'\theta_1$ that corresponds to a variable in L_i', then let $t_1', \ldots, t_n'$ be obtained from $t_1, \ldots, t_n$ by replacing all occurrences of $f(x,y)\lambda$ by $f(y,x)\lambda$. It follows that t_n' is obtained from t_2' by $n - 1$ or fewer applications of $f(x,y) = f(y,x)$ and there exist substitutions θ_1' and θ_2', u' in $\approx(u)$, v' in $\approx(v)$,

L_i'' in $\approx(L_i)$, and L_j'' in $\approx(L_j)$ such that u' is the result of replacing one occurrence of $L_i''\theta_1'$ in t_1' by $R_i\theta_1'$ and v' is the result of replacing one occurrence of $L_j''\theta_2'$ in t_n' by $R_j\theta_2'$. Because t_1' and t_2' are identical, we have again reduced our problem to considering shorter deductions. If $L_i'\theta_1$ occurs in $f(x,y)\lambda$ in a position that corresponds to a variable in $f(x,y)$, then there exists u' in $\approx(u)$ such that u' is the result of replacing one occurrence of $L_i'\theta_1$ in t_2 by $R_i\theta_1$. Again we have reduced to consideration of shorter deductions. If $L_i\theta_1$ occurs in $f(x,y)\lambda$ in a position that does not correspond to a variable position in $f(x,y)$, then there exists L_i'' in $\approx(L_i)$ such that u results from t_2 by replacing one occurrence of $L_i''\theta_1$ in t_1 by $R_i\theta_1$. This also reduces the deduction length. This completes the proof of the unique termination theorem.

The set $\mathcal{R}$ consisting of

1. $\{x \cdot 1, 1 \cdot x\} \longrightarrow \{x\}$,
2. $\{x \cdot (x^{-1}), (x^{-1}) \cdot x\} \longrightarrow \{1\}$, and
3. $\{1^{-1}\} \longrightarrow \{1\}$

is a complete set of commutative reductions relative to $x \cdot y = y \cdot x$. The finite termination of $\mathcal{R}$ is established by an argument based on decreasing the number of symbols. The unique termination of $\mathcal{R}$ is established by generating the paramodulants

4. $1^{-1} = 1$ by 1 and 2, and

5. $1 \cdot 1 = 1$ by 2 and 3 .

It can be seen that $\approx(1^{-1})^* = \approx(1)^*$ and $\approx(1 \cdot 1)^* = \approx(1)^*$.

CONCLUSIONS

Treating reduction by equivalence class methods can solve some of the problems of including commutativity in the notion of complete set of reductions. It can be shown that these methods carry over to the case of any finite number of commutative equations. It can also be shown that complete sets of commutative reductions can be combined with the narrowing methods of Lankford (2) and Slagle (3) to form refutation complete resolution systems.

The primary difficulty with the approach of complete sets of commutative reductions is that associatiativity cannot be treated as a rewrite rule. For if it could, then we would have either

A. $\{(xy)z\ ,\ (yx)z\ ,\ z(xy)\ ,\ z(yx)\} \longrightarrow \{x(yz)\ ,\ x(zy)\ ,\ (yz)x\ ,\ (zy)x\}$ or

B. $\{x(yz)\ ,\ x(zy)\ ,\ (yz)x\ ,\ (zy)x\} \longrightarrow \{(xy)z\ ,\ (yx)z\ ,\ z(xy)\ ,\ z(yx)\}$,

both of which produce infinite sequences of immediate commutative reductions, for example $\approx((xy)z) \longrightarrow \approx((yz)x) \longrightarrow \approx((xy)z) \longrightarrow \approx((yz)x) \longrightarrow \dots$. Since associative equations also result in finite equivalence classes, perhaps associativity can be treated by equivalence class methods.

ACKNOWLEDGEMENT

We thank Professor W. W. Bledsoe for his support through the Automatic Theorem Proving Project, Depts. Math. and Comput. Sci., Univ. of Texas.

REFERENCES

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

2. Lankford, D. S. Canonical inference. Automatic Theorem Proving Project, Depts. Math. and Comput. Sci., Univ. of Texas, report # ATP-32, Dec. 1975.

3. Slagle, J. R. Automated theorem proving for theories with simplifiers, commutativity, and associativity. JACM 21, 4 (Oct. 1974), 622-642.