A TECHNIQUE FOR ESTABLISHING COMPLETENESS RESULTS IN THEOREM PROVING WITH EQUALITY

Gerald E. Peterson
Department of Mathematical Sciences
University of Missouri at St. Louis
St. Louis, Mo 63121

ABSTRACT

This is a summary of the methods and results of a longer paper of the same name which will appear elsewhere.

The main result is that an automatic theorem proving system consisting of resolution, paramodulation, factoring, equality reversal, simplification and subsumption removal is complete in first-order logic with equality. When restricted to sets of equality units, the resulting system is very much like the Knuth-Bendix procedure. The completeness of resolution and paramodulation without the functionally reflexive axioms is a corollary. The methods used are based upon the familiar ideas of reduction and semantic trees, and should be helpful in showing that other theorem proving systems with equality are complete.

I INTRODUCTION

A. Paramodulation

Attempts to incorporate equality into automatic theorem provers began about 1969 when Robinson and Wos [6] introduced paramodulation and proved that if the functionally reflexive axioms were added to the set of clauses, then resolution and paramodulation constituted a complete set of inference rules. In 1975 Brand [1] showed that resolution and paramodulation are complete even without the functionally reflexive axioms. Unfortunately, the usefulness of these results is limited because unrestricted paramodulation is a weak inference rule which rapidly produces mountains of irrelevant clauses.

B. The Knuth-Bendix Procedure

In 1970 Knuth and Bendix [2], working independently of Robinson and Wos, created a very effective procedure for deriving useful consequences from equality units. Their process used paramodulation, but since it also used simplification and subsumption removal, most of the derived equalities were discarded and the search space remained small. The main defects of this procedure are that each equality must be construed as a reduction, so the commutative law is excluded, and the process works only on equality units, so most mathematical theories, including field theory, are excluded.

C. The Goal

Since resolution and paramodulation constitute a complete set of inference rules, their use will provide a proof of any valid theorem if given sufficient (usually very large) time and space. On the other hand, the Knuth-Bendix process is effective (usually small time and space) on a small class of theorems. We need to combine these two approaches and produce, if possible, an effective, complete prover.

Some progress toward this goal has been reported. For example, the commutative law can be incorporated into the Knuth-Bendix procedure by using associative-commutative unification [4,5]. Also, restricted completeness results (i.e. the set of clauses must have a certain form) have been obtained for systems which appear to be more effective than resolution and paramodulation [3].

D. Contributions of this Paper

An impediment to progress toward the goal has been the lack of an easily used technique for obtaining completeness results. We show here how the use of semantic trees can be generalized to provide completeness proofs for systems involving equality. We use this technique to obtain unrestricted completeness results for a system which is thought to be fairly effective. The verification of effectiveness will require experiments which have not yet been performed.

II METHODS AND RESULTS

A. Semantic Trees

One approach to obtaining completeness theorems is the use of semantic trees. To obtain a semantic tree $T(S)$ for a set S of clauses, we first order the atoms of the Herbrand base $\mathcal{B}$, say $\mathcal{B} = \{B_1, B_2, \ldots\}$. Then we build the binary tree T by giving each node at level k-1 two sons labelled B_k and $\sim B_k$, respectively. There will then be a one-to-one correspondence between the branches of T and the Herbrand interpretations.

If the set S is unsatisfiable, then it will be falsified by every branch of T and as we move down a branch b of T we will come to a node n_b at which it first becomes clear that b does not satisfy S. The node n_b is called a failure node of T. The portion of T(S) from the root and extending up to and including the failure nodes is called the closed semantic tree for S, $\tau(s)$. An inference node of τ is a node whose children are both failure nodes.

Every failure node n_b has an associated clause C_b in S which caused the failure. That is, there is a ground instance $C_b\theta$ of C_b such that if L is a literal of $C_b\theta$, then $\sim L$ occurs on b at or above n_b with one such $\sim L$ occurring at n_b.

It can be shown that the two clauses associated with the children of an inference node will resolve to produce a new clause C which causes failure at or above the inference node and therefore $\tau(S \cup C)$ is smaller than $\tau(s)$. By performing a sequence of such resolutions we can eventually get the closed semantic tree to shrink to a single node and this will imply that the empty clause has been inferred.

B. Incorporating Equality

Problems arise when we attempt to use this process to obtain completeness results for systems involving equality. If S is E-unsatisfiable, then S is falsified by every E-interpretation but not necessarily by every interpretation. Thus it will only be on branches which are E-interpretations that failure nodes will exist in the usual sense. The other branches must be handled in some other manner.

1. Altering Interpretations

The approach we use is to alter an arbitrary interpretation I in a way such that the resulting interpretation I* is an E-interpretation. If I is itself an E-interpretation, then no alteration is needed, I* = I.

The alternation is made as follows. First order $\mathcal{B}$ in a way such that each equality atom occurs before any atom which contains either side of the equality as a subterm. (Other restrictions are also needed on this order.) For an arbitrary interpretation I, define a partial order $\rightarrow$(I) on $\mathcal{B}$ such that $A \rightarrow B$ means essentially that B has been obtained from A by replacing a subterm s of A by a term t and I(s=t) = T. Now define I*(A) as I(A) if A is irreducible and as I*(B) if $A \rightarrow B$.

2. Substitutions

For ground substitutions θ, θ' ($\theta = \{v_1 \leftarrow t_1, \ldots, v_k \leftarrow t_k\}$), we write $\theta \rightarrow \theta'$ if θ' is identical to θ except that one term t_j of θ has been replaced by t_j' and $t_j \rightarrow t_j'$. We say that θ is irreducible if every term t_i of θ is irreducible. Suppose $C\theta_1$ and $C\theta_2$ are ground instances of a clause C. If $\theta_1 \rightarrow \theta_2$ then $I^*(C\theta_1) = I^*(C\theta_2)$.

3. Failure Nodes

Let I_b be the interpretation associated with a branch b of T(S). Then I_b^* will be an E-interpretation and will, therefore, be falsified by some clause in S. That is, there will be a ground instance $C\theta$ of a clause C in S such that $I_b^*(C\theta)$ is false and θ is irreducible (I_b). (If θ were reducible we could, by the previous paragraph, reduce it to an irreducible θ' such that $I_b^*(C\theta')$ = F.) Every literal L of $C\theta$ will be falsified by I_b^* and there will exist a failure node n_b such that $\sim L$ occurs on b at or above n_b with one such $\sim L$ occurring at n_b. These failure nodes can be split into two categories as follows. An R failure node, n_b, is one such that the associated clause C is irreducible (I_b) (thus $I_b(C)$ = F) and a P failure node is any failure node which is not an R failure node.

4. Inference Nodes

The two categories of failure nodes lead to two categories of inference nodes. A resolution inference node is a node with two R failure node children and is essentially the same thing as an inference node in a semantic tree for a set without equality. A paramodulation inference node is a P failure node n_b such that every equality node ancestor of n_b has a brother which is an R failure node.

5. Summary of the Completeness Proof

It is easy to show that if S has no E-model, then $\tau(s)$, the closed semantic tree for S, has either a resolution or paramodulation inference node. If $\tau(s)$ has a resolution inference node, then there will be a resolvent C of two clauses of S such that $\tau(S \cup C)$ is smaller than $\tau(s)$.

If $\tau(s)$ has a paramodulation inference node n_b, then there is a clause $C_2\theta$ such that $I_b^*(C_2\theta)$ = F and $C_2\theta$ is reducible (I_b), say $C_2\theta \rightarrow E$. Now $C_2\theta$ reduces to E using some equality s=t such that I_b(s=t) = T. Since s=t occurs in the ordering of $\mathcal{B}$ before the atom in $C_2\theta$ to which it applies, a node labelled s=t occurs on b above n_b. This node has a brother which is an R failure node and hence there is a clause $C_1\theta$ such that s=t is a literal of $C_1\theta$ and if L is any other literal of $C_1\theta$, then $I_b(L) = F$. It follows that $C_1\theta$ and $C_2\theta$ have a paramodulant C'. This ground paramodulation can be lifted to the general level since θ is irreducible and therefore s must start somewhere in C_2. (The lifting lemma for paramodulation holds only in this case.) Thus there is a clause C which is obtained by paramodulating C_1 into C_2 and which has a ground instance C' which is more reduced than $C_2\theta$. This greater reduction can be the basis for an ordering of the closed semantic trees involved and in the sense of this order, the tree for S $\cup$ C will be smaller than the tree for S.

C. Deletion of Unnecessary Clauses

1. Subsumption

Completeness is not lost if subsumed clauses are deleted from S as the proof proceeds.

2. Simplification

If C_1 = (s=t), C_2 contains an instance $s\sigma$ of s as a subterm, and $s\sigma\theta > t\sigma\theta$ for all ground substitutions θ, then the clause $C = C_2[t\sigma]$ is a simplification of C_2 using C_1.

If a clause C has been simplified, then C may be deleted. (Our proof of this fails when the atom simplified is an equality of a certain form, but there are other reasons for believing it is still valid in this case.)

D. The Final Result

A complete system for first-order logic with equality may consist of resolution, paramodulation, factoring, equality reversal, simplification, and subsumption removal with the following restrictions.

1. Simplification and subsumption removal are given priority since they do not increase the size of S.

2. No paramodulation into variables.

3. All paramodulations replace s by t where for at least one ground substitution θ, $s\theta > t\theta$.

4. If s > t then no reversal of the equality (s=t) will be necessary, and if C' is obtained from C by reversing (t=s) then C may be deleted.

REFERENCES

[1] D. Brand, "Proving Theorems with the Modification Method." SIAM J. Comput. 4 (1975) 412-430.

[2] D. E. Knuth and P. B. Bendix, "Simple Word Problems in Universal Algebras." in Leech, J. (ed.), Computational Problems in Abstract Algebras, Pergammon Press, 1970, 263-297.

[3] D. S. Lankford and A. M. Ballantyne, "The Refutation Completeness of Blocked Permutative Narrowing and Resolution." presented at the fourth workshop on automated deduction, February 1-3, 1979.

[4] D. S. Lankford and A. M. Ballantyne, "Decision Procedures for Simple Equational Theories with Commutative-associative Axioms: Complete Sets of Commutative-associative Reductions." Technical Report, Mathematics Department, University of Texas at Austin, August 1977.

[5] G. E. Peterson and M. E. Stickel, "Complete Sets of Reductions for Some Equational Theories." to appear in JACM.

[6] G. A. Robinson and L. Wos, "Paramodulation and Theorem Proving in First Order Theories with Equality." Machine Intelligence 4, American Elsevier, New York, 1969, 135-150.