Benchmark Problems in Which Equality Plays the Major Role*

E. Lusk and L. Wos
Mathematics and Computer Science Division
Argonne National Laboratory
Argonne, IL 60439-4801
lusk@mcs.anl.gov (708) 972-7852

ANL/CP--74903

DE92 013058

1. Introduction

We have recently heard rumors that researchers are again studying paramodulation [Wos87] in the context of strategy for its control. In part to facilitate such research, and in part to provide test problems for evaluating other approaches to equality-oriented reasoning, we offer in this article a set of benchmark problems in which equality plays the dominant role. The test problems are taken from group theory, ring theory, Robbins algebra, combinatory logic, and other areas. For each problem, we include appropriate clauses and comment as to its status with regard to provability by an unaided automated reasoning program.

2. Group Theory

A group is a nonempty set G in which multiplication is associative such that a two-sided identity e exists whose product with x is x and for which the two-sided inverse of x exists. To study group theory, one can use the following clauses; throughout the remainder of this paper, we use the notation of William McCune's program OTTER [McCune90, McCune91a], where " | " means or and " - " means not.

EQ(prod(e,x),x).
EQ(prod(x,e),x).
EQ(prod(inv(x),x),e).
EQ(prod(x,inv(x)),e).
EQ(prod(prod(x,y),z),prod(x,prod(y,z))).
EQ(x,x).

The last clause (for reflexivity) is included, for its presence is required when paramodulation is used. When attempting to prove that some set of equalities is an axiom system for group theory, except for the clause for reflexivity, one simply negates the given clauses.

Problem GT1, simple. If the square of every x is the identity, the group is commutative.

EQ(prod(x,x),e).

Problem GT2, moderate. Prove that the following equality (taken from Meredith [Meredith68] is a single axiom for groups in which the square of every x is the identity. In particular, using the single axiom, derive the axioms for groups (given earlier) and the axiom that asserts that the square of every element is the identity.

EQ(f(f(f(x,y),z),f(x,z)),y).

Problem GT3, moderate. Prove that $[[x,y],y] = e$ when the cube of every x is e, where $[x,y]$ is the product of x, y, the inverse of x, and the inverse of y.

EQ(prod(x,prod(x,x)),e).
EQ(com(x,y),prod(x,prod(y,prod(inv(x),inv(y))))).

Problem GT4, moderate. Prove that the following equality axiomatizes group theory. The corresponding theorem, proved by William McCune using OTTER [McCune91b], is a new contribution to the literature.

EQ(f(x,g(f(y,f(f(f(z,g(z)),g(f(u,y))),x)))),u).

Problem GT5, moderate. Prove that the following equality axiomatizes commutative group theory; this new

*This work was supported by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U.S. Department of Energy, under Contract W-31-109-Eng-38, and by the National Science Foundation under grant CCR-8810947

single axiom was found by William McCune using OTTER [McCune91b].

EQ(f(f(f(x,y),z),g(f(x,z))),y).

Problem GT6, never proved in a single run. Prove that the following equality axiomatizes commutative group theory; the result was verified by Ken Kunen [private correspondence] using OTTER as an assistant.

EQ(f(g(f(g(f(g(f(x1,x2)),f(x2,x1))),f(g(f(z,y)),f(z,g(f(f(v,g(x)),g(y))))))),x),v).

Problem GT7, hard. Prove that the Fibonacci group given by the following equalities is the cyclic group of order 29, where, for example, a(x) means the product of *a* and *x* and a1(x) means the product of the inverse of *a and x*. Instead, one can take the axioms for groups and add equalities that assert that, for seven elements, the product of the first two is the third, ..., and the product of the seventh and the first is the second.

EQ(a(b(x)),c(x)).
EQ(b(c(x)),d(x)).
EQ(c(d(x)),e(x)).
EQ(d(e(x)),f(x)).
EQ(e(f(x)),g(x)).
EQ(f(g(x)),a(x)).
EQ(g(a(x)),b(x)).
EQ(a(a1(x)),x).
EQ(b(b1(x)),x).
EQ(c(c1(x)),x).
EQ(d(d1(x)),x).
EQ(e(e1(x)),x).
EQ(f(f1(x)),x).
EQ(g(g1(x)),x).
EQ(a1(a(x)),x).
EQ(b1(b(x)),x).
EQ(c1(c(x)),x).
EQ(d1(d(x)),x).
EQ(e1(e(x)),x).
EQ(f1(f(x)),x).
EQ(g1(g(x)),x).

3. Ring Theory

A ring *R* is a nonempty set in which addition and multiplication are defined such that under addition the set is a group and such that multiplication is associative and multiplication distributes over addition. The following clauses capture the properties of a ring.

EQ(sum(0,x),x).
EQ(sum(x,0),x).
EQ(sum(minus(x),x),0).
EQ(sum(x,minus(x)),0).
EQ(sum(sum(x,y),z),sum(x,sum(y,z))).
EQ(x,x).
EQ(sum(x,y),sum(y,x))
EQ(prod(prod(x,y),z),prod(x,prod(y,z))).
EQ(prod(x,sum(y,z)),sum(prod(x,y),prod(x,z))).
EQ(prod(sum(y,z),x),sum(prod(y,x),prod(z,x))).

Problem RT1, moderate. Prove that Boolean rings (rings in which the square of every *x* is *x*) are commutative.

EQ(prod(x,x),x).

Problem RT2, hard. Prove that rings in which the cube of every *x* is *x* are commutative.

EQ(prod(prod(x,x),x),x).

Problem RT3, hard. Prove that rings in which the fourth power of *x* is *x* are commutative.

EQ(prod(prod(prod(x,x),x),x),x).

Problem RT4, never proved in a single run unaided. Prove that rings in which the fifth power of *x* is *x* are commutative.

EQ(prod(prod(prod(prod(x,x),x),x),x),x).

4. Robbins Algebra

A Robbins algebra is a nonempty set satisfying the following three axioms, expressed in clause notation, in which the function *o* can be interpreted as plus and the function *n* as negation.

EQ(o(x,y),o(y,x)).
EQ(o(o(x,y),z),o(x,o(y,z))).
EQ(n(o(n(o(x,y)),n(o(x,n(y))))),x).
EQ(x,x).

A Boolean algebra is a nonempty set S with two operations, plus and times, and a 0 and a 1. Each operation is commutative, and each distributes over the other. The 1 is a multiplicative identity, and the 0 is an additive identity. In addition, for every x, the negation of x exists with x plus its negation equal to 1 and x times its negation equal to 0. An alternative axiomatization of Boolean algebra consists of (R1), (R2), and Huntington's axiom (H3) [Huntington33].

(H3) EQ(o(n(o(n(x),y)),n(o(n(x),n(y)))),x).

Whether Robbins implies Boolean is still an open question. What is known is that the addition to the three Robbins axioms of any one of a number of properties of a Boolean algebra suffices to yield Boolean.

Problem RA1, simple. Prove that, if the following axiom is adjoined to the axioms for a Robbins algebra, the resulting algebra is Boolean. We recommend trying to prove Huntington's axiom (H3).

EQ(o(x,0),x).

Problem RA2, moderate. Prove that t addition of the following equality to Robbins yields Boolean.

EQ(o(x,x),x).

Problem RA3, hard. Prove that, where c is a constant, the addition of the following equality to Robbins yields Boolean.

EQ(o(c,c),c).

Problem RA4, never proved in a single run unaided. Where c and d are constants, the addition of the following equality to Robbins yields Boolean.

EQ(o(c,d),d).

Problem RA5, never proved in a single run unaided. Where c and d are constants, the addition of the following equality to Robbins yields Boolean.

EQ(n(o(c,d)),n(d)).

5. Combinatory Logic

Barendregt [Barendregt81] defines combinatory logic as an equational system satisfying the combinators S and K with $((Sx)y)z = (xz)(yz)$ and $(Kx)y = x$. Rather than studying this logic in its entirety, one finds challenging test problems by replacing one or both of S and K by one or more combinators and focusing on questions concerning the possible presence of the strong fixed point property. The set consisting of the combinators under study is called a *basis*, and the set of combinators generated by a basis is called a *fragment*. Where **A** is a given fragment with basis **B**, the strong fixed point property holds for **A** if and only if there exists a combinator y such that, for all combinators x, $yx = x(yx)$, where y is expressed purely in terms of elements of **B**. The problems we offer focus on various subsets of the following combinators.

EQ(a(a(a(B,x),y),z),a(x,a(y,z))).
EQ(a(a(a(C,x),y),z),a(a(x,z),y)).
EQ(a(a(a(H,x),y),z),a(a(a(x,y),z),y)).
EQ(a(I,x),x).
EQ(a(M,x),a(x,x)).
EQ(a(a(a(N,x),y),z),a(a(a(x,z),y),z)).
EQ(a(a(a(S,x),y),z),a(a(x,z),a(y,z))).
EQ(a(a(W,x),y),a(a(x,y),y)).

For each of the following problems, the object is to use the given combinators and no others and prove that the strong fixed point property holds by finding an appropriate object.

Problem CL1, simple. The set consists of B, M, and W.

Problem CL2, hard. The set consists of B and W.

Problem CL3, hard. The set consists of B and N.

Problem CL4, hard. The set consists of B and H.

Problem CL5, hard. The set consists of B, C, I, and S.

To complement the preceding test problems and for those who enjoy the study of open questions, we offer two open questions. Does the set consisting of B and M alone permit the construction of an object that proves that the strong fixed point property holds? Does the set consisting of B and S alone permit the construction of an object that proves that the strong fixed point property holds?

6. Many-Valued Sentential Calculus

The axioms are the following, where the constant T can be interpreted as "true" and the functions i and n as implication and negation, respectively.

EQ(i(T,x),x).
EQ(i(i(x,y),i(i(y,z),i(x,z))),T).
EQ(i(i(i(x,y),y),i(i(y,x),x)).
EQ(i(i(n(x),n(y)),i(y,x)),T).

Problem MV1, simple. Prove that each of the following two equalities hold in many-valued sentential calculus.

EQ(i(n(n(x)),x),T).
EQ(i(x,n(n(x))),T).

Problem MV2, moderate. Prove that the following holds in the calculus.

EQ(i(i(x,y),i(i(z,x),i(z,y))),T).

Problem MV3, moderate. Prove that the following holds in the calculus.

EQ(i(i(x,y),i(n(y),n(x))),T).

Problem MV4, hard. Prove that the following holds in the calculus.

EQ(i(i(i(x,y),i(y,x)),i(y,x)),T).

7. Nonunit Problems

With the intention of spurring research focusing on paramodulation in which nonunit clauses occur, we offer the following problems.

Problem NU1, moderate. The problem asks one to prove the following identity in modular lattices in which a 0 and 1 exist, where ^ is *meet*, v is *join*, and ´ is *complement*.

((A v B)´ v ((A ^ B)´ ^ B)) ^ ((A v B)´ v ((A ^ B)´ ^ A)) = (A v B)´

The following clauses can be used, the first 18 of which capture the properties of a modular lattice.

EQ(meet(0,x),0).
EQ(meet(x,0),0).
EQ(join(0,x),x).
EQ(join(x,0),x).
EQ(meet(1,x),x).
EQ(meet(x,1),x).
EQ(join(1,x),1).
EQ(join(x,1),1).
EQ(meet(x,x),x).
EQ(join(x,x)x).
EQ(meet(x,y),meet(y,x)).
EQ(join(x,y),join(y,x)).
EQ(meet(meet(x,y),z),meet(x,meet(y,z))).
EQ(join(join(x,y),z),join(x,join(y,z))).
EQ(meet(x,join(x,y)),x).
EQ(join(x,meet(x,y)),x).
-EQ(meet(x,z),x) | EQ(meet(z,join(x,y)),join(x,meet(y,z))).
EQ(x,x).
EQ(join(r2,meet(a,b)),1).
EQ(meet(r2,meet(a,b)),0).
EQ(join(r1,join(a,b)),1).
EQ(meet(r1,join(a,b)),0).
EQ(join(r1,meet(a,r2)),b2).
EQ(join(r1,meet(b,r2)),a2).
-EQ(meet(a2,b2),r1).

Problem NU2, moderate. Prove that subgroups of index 2 are normal, using the clauses given earlier to study group theory. O(x) means that *x* is in the subgroup, j(x,y) is an element of the subgroup that exists if *x* and *y* are not in the subgroup (for giving the definition of index 2).

```
O(e).
-O(x) | O(inv(x)).
-O(x) | -O(y) | O(prod(x,y)).
O(x) | O(y) | O(i(x,y)).
O(x) | O(y) | EQ(prod(x,j(x,y)),y).
O(b).
-O(prod(a,prod(b,g(a)))).
```

Problem NU3, hard. In set theory, prove that if two ordered pairs are equal, then they are equal componentwise. In the following clauses, op(x,y) means the ordered pair, up(x,y) means the unordered pair, IN is set membership, and sing is the singleton set consisting of *x*.

```
EQ(x,x).
IN(x,sing(x)).
-IN(x,sing(y)) | EQ(x,y).
IN(x,up(x,y)).
IN(y,up(x,y)).
-IN(x,up(y,z)) | EQ(x,y) | EQ(x,z).
EQ(op(x,y),up(sing(x),up(x,y))).
EQ(op(m1,r1),op(m2,r2)).
-EQ(m1,m2) | -EQ(r1,r2).
EQ(op(x,y),up(sing(x),up(x,y))).
```

References

[Barendregt81] Barendregt, H. P. *The Lambda Calculus: Its Syntax and Semantics*, North-Holland, Amsterdam (1981).

[Huntington33] Huntington, E., "New sets of independent postulates for the algebra of logic, with special reference to Whitehead and Russell's Principia Mathematica", *Trans. of AMS*, **35**, pp. 274-304 (1933).

[McCune90] McCune, W., *OTTER 2.0 Users Guide*, Technical Report ANL-90/9, Argonne National Laboratory, Argonne, Illinois, 1990.

[McCune91a] McCune, W., *What's New in OTTER 2.2*, Mathematics and Computer Science Division Technical Report ANL/MCS-TM-153, Argonne National Laboratory, 1991.

[McCune91b] McCune, W., *Single Axioms for Groups and Abelian Groups with Various Operations*, Mathematics and Computer Science Division Preprint MCS-P270-1091, Argonne National Laboratory, Argonne, Illinois, 1991.

[Meredith68] Meredith, C. A., and Prior, A. N., "Equational logic", *Notre Dame J. Formal Logic*, **9**, pp. 212-226 (1968).

[Wos87] Wos, L., *Automated Reasoning: 33 Basic Research Problems*, Prentice-Hall, Englewood Cliffs, N.J., 1987.

END

DATE
FILMED
6/09/92