(Rev. May 12, 1970)

SLAC-PUB-814
September 1970
(MISC)

MAXIMAL MODELS AND REFUTATION COMPLETENESS: SEMIDECISION PROCEDURES IN AUTOMATIC THEOREM PROVING*

by

Lawrence Wos
Argonne National Laboratory
Argonne, Illinois

and

George Robinson
Stanford Linear Accelerator Center
Stanford, California

Presented at a Conference on Decision Problems in Group Theory, University of California, Irvine, September 1969.

---

*Work performed in part under the auspices of the United States Atomic Energy Commission.

1. Introduction

In recent years the idea of using electronic computers to search for proofs of theorems of quantification theory has drawn considerable attention. One of the more successful methods of attack on the problem has stemmed from the work of Quine,[12] Gilmore,[5] Davis and Putnam,[4] and J. Robinson.[16] This paper is concerned with a portion of the theory underlying an extension of this line of development to systems--first-order theories with equality--in which there is a distinguished relation symbol for equality. The field of mathematics upon which we have concentrated our computer experiments in order to study various properties of our procedure is first-order group theory.

To say that a particular property is decidable for some class of statements means that there is a single uniform procedure which will correctly determine whether the property holds when presented with any given statement from the class. Church showed[2] that, for any fixed procedure, there exists a statement of first-order predicate calculus for which the procedure will not be able to answer correctly the question: is this statement a theorem? For group theory, the question of theoremhood is also known to be undecidable as proved by Tarski.[17] The situation is, however, far from hopeless for first-order theories at least. There do exist procedures which will, if presented with a set of first-order axioms for a theory and a first-order statement that happens to be a theorem thereof, correctly identify the statement as a theorem. Such a procedure is called a *semidecision procedure for theoremhood.* The basis for such a procedure is often a set of inference rules (rules for reasoning from the

axioms of the theory to the statement whose theoremhood is in question), in which case it is natural to call the procedure a *proof procedure* since the procedure generates a proof of the chosen statement if it is a theorem. In order for a procedure to be of interest from the computational viewpoint, it must also be reasonably efficient. The indications are that the inference rules given in this paper may provide the basis for an efficient semidecision procedure for theoremhood not only for first-order group theory but for other first-order theories with equality as well. The reasons for expecting efficiency are: equality is treated as a special logical symbol distinguished from the relation symbols of the mathematical theory, the inferences deduced with the rules have a certain important property of generality that causes us to call the rules *conservative*, and finally a number of inferences which are ordinarily obtained immediately in more classical systems are not deducible with these rules. This last property of nondeducibility is referred to as the lack of deduction completeness. For example, not all theorems of the familiar first-order predicate calculus can be deduced from the proposed set of rules. Contrary to intuition, this is often an advantage computationally.

The set, Π, of inference rules to be studied in this paper consists of resolution, factoring, and paramodulation. The first two are generalizations of well known inference rules for the propositional calculus. Paramodulation is a generalization of a substitution rule for equality. The formulation of resolution and factoring is essentially that of J. Robinson[16] while paramodulation originated with the present authors.[13,15]

It is desirable that the underlying set of inference rules have certain logical properties related to semidecidability. The key property

is that of R-refutation completeness. A refutation is a proof of contradiction. It is proved herein that Π is R-refutation complete, i.e., a refutation employing just the rules from Π exists for any set of statements (each of which is in the appropriate logical form) which possesses no equality model. In other words, if one starts with a statement which should lead to a contradiction with the usual meaning of equality, Π is strong enough to yield a contradiction.

The general idea is as follows. To prove that a statement is a theorem of some first-order theory with equality, one proceeds by first denying the statement. Then this denial and the axioms of the theory in question are converted to a particular logical form which may be said to be, loosely speaking, a conjunction of statements each of which is a disjunction. The existential quantifiers have been replaced by Skolem functions and the remaining variables are considered to be universally quantified precisely over the entire conjunct in which they occur. The resulting statements are called clauses. A contradiction is deducible (using the rules of Π) from the set of clauses if and only if the original statement was true in all equality models of the theory. The sufficiency requirement leads to a definition of R-refutation completeness. The necessity leads to a corresponding soundness concept.

Several examples from first-order group theory are discussed in Appendix 1, and refutations (within Π) for two of them are given.

It is not within the scope of this paper to discuss the strategies which lead to a promising semidecision procedure. Intuitively, "strategy" may refer to the order in which the inference rules are applied or to certain constraints placed on their application. In order to have an efficient

semidecision procedure, one needs, in addition to good inference rules, various strategies. The procedure used in Corollary 2 is not one which is recommended. One of the most successful approaches is that which allows only those inferences which are in part traceable to the special hypothesis of the theorem or to the denial of its conclusion. This strategy is a special case of the set of support strategy. The set of support strategy is known to be R-refutation complete when coupled with Π.[21]

To illustrate the general approach, consider the theorem: if $x^2 = e$ for all x in a group G, then G is commutative. In the notation of first-order predicate calculus the axiom for the existence of an identity element (two-sided) is, $(\exists y)(x)[Pyxx \wedge Pxyx]$, where P denotes product. Replacing the existential quantifier by a Skolem function and then putting the results into the desired form (a set of clauses), one has the clauses Pexx and Pxex. Since the computer attempts to find contradiction from assuming the theorem false, one has, in addition to the clauses corresponding to the other axioms, the clauses Pxxe and Pabc and $\overline{P}bac$. The last two clauses correspond to the assumption that the group is not commutative, i.e., that there exists a pair of elements which do not commute.

The system Π uses no logical axioms nor does it contain any of the more familiar classical rules of inference such as universal instantiation. In the example above, the procedure would consist of applying the various rules of Π, resolution, factoring, and paramodulation, in some order until a contradiction was found. Thus we are vitally interested in the property of R-refutation completeness.

The use of Herbrand models throughout the paper rather than the more familiar concrete models provides a convenient tool for proving the theorems underlying the theory of the approach discussed herein. The theorem of logic which permits use of Herbrand models states in effect that a set has an Herbrand model if and only if it has a concrete model.

In Section 2 we give definitions of a number of concepts, such as *interpretation* and *model*, which are necessary for our study. In Section 3 we prove the *Maximal Model Theorem* which states that, given a particular interpretation for a given satisfiable set of disjunctions, there is a model for that set of disjunctions, whose set-theoretic intersection with the particular interpretation is as small as it can be without losing the property of modelhood. In Sections 4 and 5 we give the inference rules. In Section 6 we introduce the terms, *refutation complete*, *deduction complete*, and *conservative* and prove that the inference system (set of inference rules) Π under study has the desired logical property of R-refutation completeness for *functionally reflexive* sets. The usual axiom set for basic group theory[13] and also that for basic ring theory, for example, are functionally reflexive relative to Π.

Although many classical inference systems have the property of refutation completeness, they are also deduction complete and not conservative. The system, Π, on the other hand, is conservative but not deduction complete. In the light of experience with existing computer programs, this seems to be advantageous in using computers to search for proofs of theorems.

2. Definitions

We shall deal with a language having as primitive symbols *individual variables* $x_1,x_2,\ldots,x_n,\ldots$; *individual constants* $a_1,a_2,\ldots,a_n,\ldots$; *predicate constants* (relation symbols) $P_1^1,P_1^2,\ldots,P_k^j,\ldots$; and *function letters* $f_1^1,f_1^2,\ldots,f_k^j,\ldots$ (where superscripts indicate the number of arguments). (The *equality predicate* P_1^2 will frequently be abbreviated to R.) The remaining primitive symbol is a bar for negation. Individual constants and variables are *terms;* a function letter f_i^n applied to n terms $t_1,\ldots,t_n$ is also a *term.* A predicate letter P_i^n applied to n terms is an *atomic formula* or *atom.* A *literal* is an atom q or the negation $\bar{q}$ thereof. If q is an atom, the negation $\bar{\bar{q}}$ of $\bar{q}$ is just taken to be q. The *absolute value* $|h|$ of a literal h is the atom q such that either h is q or h is $\bar{q}$. A *clause* is a finite set of literals. Intuitively, one may view a clause as the disjunction of its literals, universally quantified (over the entire disjunction) on its individual variables. Also intuitively, existential quantification is (in effect) accomplished through the use of (Skolem) function letters. A *ground clause (literal,term)* is one that has no variables occurring in it.

The result $C\theta$ of a *substitution* $\theta = [t_1/u_1,\ldots,t_n/u_n]$ (where the u_i are all distinct) on a clause C (literal,term) is the clause (literal,term) obtained by replacing uniformly and simultaneously each occurrence (if any) of each variable u_i $(i = 1,\ldots,n)$ by the corresponding term t_i. Here the clause (literal,term) $C\theta$ is called an *instance* of C. Suppose that a substitution θ transforms S, some set of literals, into a set $S\theta$. Then if for all literals h_1 and h_2 in $S\theta$, $|h_1| = |h_2|$, we shall call θ a *unifier* (or *match*) of S. Then set S is said to be *unifiable* if such a θ exists.* If θ transforms all members of a set

*When a finite set S of literals is to be unified, we do not usually find it particularly instructive to view that set as a clause, although it is not set-theoretically distinguishable from a clause.

T of terms into a single term, θ is likewise called a *unifier* or *match* of T. The process of finding and applying a unifier or match is called *unification* or *matching*.

If a substitution θ can be obtained by composing two substitutions θ_1 and θ_2 in that order, then θ is an *instance* of θ_1. If one clause (literal,term,substitution) is an instance of a second and the second is also an instance of the first, then each is a *variant* of the other. A finite, non-empty set of clauses (literals,terms,substitutions) that have an instance in common have a *most general common instance* (not necessarily unique), i.e., one that itself has as instances all common instances of the originals. Similarly, if a finite, non-empty set S of literals (terms) is unifiable, then it has a *most general unifier (match)*, i.e., a substitution θ that unifies S and such that every unifier of S is an instance of θ.

If a term t occurs as the i_k-th argument of the i_{k-1}-st argument of ... of the i_1-st argument of a literal h, then the ordered k-tuple $(i_1,\ldots,i_k)$ is called the *position vector* of that occurrence of t in h. Two terms are *in the same position* in their respective literals if they have the same position vectors.

The *Herbrand universe* H_S of a set S of clauses (literals,terms) is the set of all ground terms that can be constructed from the symbols occurring in S. Here if no individual constant occurs in S, a_1 is to be added to the vocabulary. An *Herbrand atom* for a set S of clauses is a predicate letter P_i^n from S applied to n terms from H_S. An *interpretation I of* S is a set of ground *literals* whose absolute values are all the Herbrand atoms for S, such that for each Herbrand atom j exactly one of j and $\bar{j}$ is in I. An interpretation I *satisfies the ground clause* C' if at least one

of the *literals* of C' is in I, i.e., if $I \cap C' \neq \emptyset$. The interpretation I *satisfies the* (possibly non-ground) *clause* C if it satisfies every ground instance of C (over the Herbrand universe under consideration); it *satisfies the set* S *of clauses* if it satisfies each clause in S; it *condemns the ground clause* C' if it contains the negation of every literal of C'; it *condemns a clause* C if it condemns some ground instance of C (over the Herbrand universe under consideration); and it *condemns the set* S *of clauses* if it condemns some clause in S. When the empty set of literals is viewed as a clause *false*, it can be satisfied by no interpretation and is vacuously condemned by every interpretation. Note however that the same set-theoretic object $\emptyset$, when viewed as a set of clauses, is vacuously satisfied by all interpretations.

A *model* M *of the set* S *of clauses* is an interpretation of S that satisfies S; it is an *R-model* of S if, in addition, the set $\{(t_1,t_2) \mid Rt_1t_2 \in M\}$ (i.e., the set $\{(t_1,t_2) \mid P_1^2t_1t_2 \in M\}$) is an equality relation for the terms occurring in M, i.e., if each of the following is true for all terms s and t in H_S and all function letters f_k^n:

(i) $Rtt \in M$.

(ii') If the literal $k \in M$ and $Rst \in M$ and if h is obtained by replacing in k some one occurrence of s by an occurrence of t, then $h \in M$.

It is often more convenient to replace (ii') by the condition

(ii) If an *atom* $k \in M$ and $Rst \in M$ and if h is obtained, etc.

since (i) and (ii) are equivalent to (i) and (ii').

The other familiar properties follow easily from (i) and (ii'):

(iii) If $Rst \in M$, then $Rts \in M$.

(iv) If $Rst \in M$ and $Rtu \in M$, then $Rsu \in M$.

(v) For all $t_0,t_1,\ldots,t_n$ in H_S, if $Rt_jt_o \in M$ and f occurs in some literal in M, then

$$Rf(t_1,\ldots,t_{j-1},t_j,\ldots,t_n)f(t_1,\ldots,t_{j-1},t_0,\ldots,t_n) \in M.$$

3. The Maximal Model Theorem

One can associate with each interpretation T of a set S of clauses a partial ordering relative to T of the set of all interpretations of S, and hence of the set of models of S, given by:

$M_1 \leq_T M_2$ if and only if $M_1 \cap T \supseteq M_2 \cap T$.

The maximal model theorem states that, given a set S of clauses possessing a model and an interpretation T of S, among the models of S there exists one, say M, such that no other model of S has a strictly smaller set-theoretic intersection with T.

Maximal model theorem. If a set S of clauses has a model, then, given any interpretation T of S, S has a maximal model M_o relative to T.

Proof: Let $T' = \{x | \bar{x} \in T\}$, or equivalently the set of the negations of the literals in T. Then T' is an interpretation of S, and $T' \cap T = \emptyset$.

If we show that every simply ordered (relative to T) set of models of S has an upper bound which is itself a model, then an application of Zorn's lemma will suffice. Let the set $\{H_\alpha\}_{\alpha \in \Lambda}$ be a simply ordered (relative to T) set of models of S. Let $M = \bigcup_{\alpha \in \Lambda} (H_\alpha \cap T') \cup (\bigcap_{\alpha \in \Lambda} (H_\alpha \cap T))$. M will be shown to be both an upper bound for $\{H_\alpha\}$ and a model of S.

That M is an upper bound (if it is an interpretation) follows from considering H_β for an arbitrary β in Λ and noting that $H_\beta \cap T' \subseteq \bigcup_\alpha (H_\alpha \cap T') = M \cap T'$, since $T \cap T'$ is empty. Thus $H_\beta \cap T' \subseteq M \cap T'$, hence $H_\beta \cap T \supseteq M \cap T$ and $H_\beta \leq_T M$.

To see that M is an interpretation, consider an arbitrary literal b in T since every Herbrand atom is represented in T either by itself or by its negation. If b is in H_α for all α in Λ, then b is in $\bigcap_\alpha (H_\alpha \cap T)$ so b is in M. If for some α, the literal b is not in H_α, then for that α, $\bar{b}$ is in H_α since each H_α is an interpretation. But $\bar{b} \in T'$, so $\bar{b} \in \bigcup_\alpha (H_\alpha \cap T') \subseteq M$. This shows that at least one of b and $\bar{b}$ is in M. To see that only one of b and $\bar{b}$ is in M assume that b is in M. Then since b is not in T' it cannot be in $\bigcup_\alpha (H_\alpha \cap T')$ and so must be in $\bigcap_\alpha (H_\alpha \cap T)$ and hence in $\bigcap_\alpha H_\alpha$. But with b in every H_α, $\bar{b}$ cannot be in any H_α and hence not in $\bigcup_\alpha (H_\alpha \cap T')$ and hence not in M.

To see that M is a model of S, consider some arbitraily chosen ground instance D of some clause in S.

Case 1. $D \subseteq T'$. Let H_β be arbitrarily chosen. Since H_β is a model of S, $H_\beta \cap D$ is not empty and so contains some literal, say c. Since D is contained in T' by assumption, c is in T', so $c \in T' \cap H_\beta \subseteq \bigcup_\alpha (H_\alpha \cap T') \subseteq M$. Therefore, $c \in M \cap D$.

Case 2. $D \not\subseteq T'$. Then $D \cap T = \{c_1, c_2, \ldots, c_k\}$ for some finite $k > 0$, since D has only a finite number of literals and since all literals of all ground instances of all clauses of S are in $T \cup T'$.

If some c_i, say c, is in all H_α, then c is in $\bigcap_\alpha (H_\alpha \cap T) \subseteq M$, and $M \cap D$ is not empty. Otherwise assume that for $1 \leq i \leq k$ no c_i is in all H_α. Therefore, for each c_i, there is an H_α, say H_{α_i}, with $c_i \notin H_{\alpha_i}$. But the H_α are models and hence interpretations, so $\bar{c}_i \in H_{\alpha_i}$ for $1 \leq i \leq k$. Since $c_1, \ldots, c_k$ are in T, $\bar{c}_1, \ldots, \bar{c}_k$ are in T'. Since the family $\{H_\alpha\}_{\alpha \in \Lambda}$

is simply ordered relative to T, there is among $H_{\alpha_1},\ldots,H_{\alpha_k}$ a largest (relative to T) H_{α_i}, say H. From the definition of $\leq_T$, if a literal is in some $H_{\alpha_i} \cap T'$ $(1 \leq i \leq k)$ then it is in $H \cap T'$.

Therefore, $\{\bar{c}_1,\ldots,\bar{c}_k\} \subseteq H \cap T'$. Since H is a model and hence an interpretation of S, no c_i for $1 \leq i \leq k$ is in H. But $H \cap D \neq \emptyset$ since H is a model of S, so $H \cap D$ contains some element a. Since no c_i is in H, $a \neq c_i$ for each i. Therefore, $a \notin T$ since $T \cap D = \{c_1,c_2,\ldots,c_k\}$. So a is in T', and therefore $a \in H \cap T'$. But $H \cap T' \subseteq \bigcup_{\alpha} (H_\alpha \cap T') \subseteq M$, so again $M \cap D \neq \emptyset$.

Since the cases for D are exhausted, and since D was arbitrarily chosen, for each ground instance D of a clause in S, $M \cap D \neq \emptyset$. M is thus a model of S.

The conditions for Zorn's Lemma being satisfied, there exists a maximal (relative to the ordering associated with T) model M_o of S. Q.E.D.

In a similar fashion one can, by applying the maximal model theorem to T', show that there exists a minimal model of S with respect to T.

In a later section we apply the theorem to the case where T is the set of Herbrand atoms for S. In this case, the key point is that if M_o is a maximal model relative to T, for each atom k in M_o we can find a ground instance D of a clause of S with $D \cap M_o = \{k\}$. This in turn allows us certain applications of paramodulation, which we shall define in a later section.

Corollary A. If S is a set of clauses, T an interpretation of S, and M a maximal (relative to T) model of S, then, for each literal b in $M \cap T$, there exists a clause in S having a ground instance (over H_S) D with $D \cap M = \{b\}$.

Proof. Let S be a set of clauses, T be an arbitrary but fixed interpretation, and M be a maximal (relative to T) model of S. Assume by way of contradiction that there exists a literal b in $M \cap T$ falsifying the theorem. Since M is a model of S we can conclude, therefore, that, for any ground instance (over H_S) C of any clause in S, $(M \cap C) - \{b\}$ is not empty. Let M* be obtained from M by replacing b by $\bar{b}$. M*, therefore, has a non-empty intersection with all ground instances (over H_S) of all clauses in S. For each Herbrand atom a, M* will still contain exactly one of a and $\bar{a}$. M* is, therefore, an interpretation of S and, hence, a model of S. Since b is in T by assumption and T is an interpretation, $\bar{b}$ is not in T. Therefore, $M \cap T$ contains $M^* \cap T$ as a proper subset. This contradicts the maximality of M, and the proof is complete.

Corollary A and the maximal model theorem establish that there are models (the maximal models) which possess the intersection property (given in the conclusion of corollary A) upon which the proof of R-refutation completeness rests. That this intersection property, however, does not characterize maximality can be seen from the following example.

Let $S = \{A, B\}$, where $A = \{\bar{P}a, Qa\}$ and $B = \{Pa, \bar{Q}a\}$. Let $T = \{Pa, Qa\}$. Let $M^* = \{Pa, Qa\}$. M* is not maximal relative to T, for the model $M = \{\bar{P}a, \bar{Q}a\}$ is such that $M^* \cap T = \{Pa, Qa\}$ contains $M \cap T = \phi$ as a proper subset. M is itself a maximal model relative to T. M*, however, has the property that, for every literal b in $M^* \cap T$, there exists a clause in S one of whose ground instances intersects M* in exactly b. This example shows that the property of maximality for models is not equivalent to the intersection property.

Depending on S and T, one cannot be assured of the existence of a unique maximal model. In the example just given, M is the only maximal model. But, if one replaces S by $S^* = \{A^*, B^*\}$, where $A^* = \{Pa, Qa\}$ and $B^* = \{\overline{P}a, \overline{Q}a\}$, there are two maximal models. $M_1 = \{Pa, \overline{Q}a\}$ and $M_2 = \{\overline{P}a, Qa\}$ are both maximal models relative to T.

By definition, a model M is maximal relative to T if and only if, whenever $M \cap T$ contains $M^* \cap T$ for a model M^*, $M \cap T = M^* \cap T$. Using the definition of interpretation one can easily show that a model M is maximal relative to T if and only if, whenever $M \cap T$ contains $M^* \cap T$ for a model M^*, M is equal to M^* set-theoretically.

4. Resolution and Factoring

In this section we give a brief account of an inference rule called *resolution*. This rule may be viewed as a generalization of *modus ponens* and hypothetical syllogism.

Definition (Resolution): If A and B are clauses (with no variables in common) with literals k and h respectively such that k and h are opposite in sign (i.e., exactly one of them is an atom) but $|k|$ and $|h|$ have a most general common instance m, and if σ and τ are most general substitutions with $m = |k|\sigma = |h|\tau$, then infer from (any variants A* and B* of) A and B the clause $C = (A - \{k\})\sigma \cup (B - \{h\})\tau$. C is called a *resolvent* of A* and B* and is inferred by *resolution*.**[16][14]

Example 1: Premises {Pabc} and {$\overline{P}$xyz,Pyxz} yield by resolution {Pbac}. In this example resolution has the effect of first instantiating the second premise (which might be thought of as the commutative axiom in a group theory problem) so that one has two premises to which *modus ponens* can be applied, and then applying *modus ponens*. The clauses which serve as premises for the *modus ponens* applications are most general instances of the given premises, most general with respect to permitting application of *modus ponens*.

Example 2: Premisses {Pax,Qx} and {$\overline{P}$yb,Qy} yield by resolution {Qa,Qb}. Syllogistic inference is not directly applicable to the given pair of clauses. The clause {Qa,Qb}, however, can be inferred by instantiation

---

**Note that the premisses A* and B* to which the rule of inference is applied, are not themselves subject to the restriction of having no variables in common. Given an arbitrary A* and B*, in practice, one need merely re-letter B* to obtain a variant B having no variable in common with A* before constructing the resolvent C from A* and B.

followed by a syllogistic inference. By definition, resolution, in effect, seeks most general instances of the clauses which will permit syllogistic inference.

Example 1 shows that resolution yields in the spirit of *modus ponens* an inference from pairs of premisses even though *modus ponens* does not apply directly to those premisses. Example 2 illustrates the corresponding relationship to syllogistic inference. By placing no restriction on the (finite) number of literals in either premise, resolution extends both *modus ponens* and syllogistic inference in yet another way.

Example 3: From {Nx,Px} and {$-\overline{P}$y,Qy} infer by resolution {$\overline{N}$y,Qy}. This is the familiar categorical syllogism, all F are G, all G are H, so all F are H.

Definition (Factoring): If A is a clause with literals k and h such that k and h have a most general common instance m, and if σ is a most general substitution with $k\sigma = h\sigma = m$, then infer the clause $A' = (A - \{k\})\sigma$ from A. A' is called an *immediate factor* of A. The *factors* of A are given by: A is a factor of A, and an immediate factor of a factor of A is a factor of A.

Example 4: From {Qax,Qyb,Quz} infer (among other clauses) as an immediate factor {Qab,Quz}, which in turn has an immediate factor {Qab}.

5. Paramodulation

The concept of equality is ordinarily handled in first-order theories in one of two ways: 1) a set of explicit first-order axioms (or axiom schemata) is given for the equality predicate; 2) a substitution rule is supplied together with the axiom of reflexivity. It is in the context of the second approach that the inference rule paramodulation is to be understood.

With the appropriate constraints on the variables in the terms s and t, one version of the substitution rule for equality permits inference of the formula $\rho(t)$ from the formulae Rst (i.e., s=t) and $\rho(s)$. Paramodulation[13,15] extends this substitution rule by permitting inference from the formulae Rst and $\rho(u)$ when the terms s and u, though not identical, have a common instance. (Since the only formulae of interest throughout this paper are clauses, paramodulation is defined only for clauses.)

Definition of Paramodulation: Let A and B be clauses (with no variables in common) such that a literal Rst (or Rts) occurs in A and a term u occurs in (a particular position in) B. Further assume that s and u have a most general common instance $s' = s\sigma = u\tau$ where σ and τ are most general substitutions such that $s\sigma = u\tau$. Where $\hat{B}$ is obtained by replacing by $t\sigma$ the occurrence of $u\tau$ in the position in $B\tau$ corresponding to the particular position of the occurrence of u in B, infer from any variants A* and B* of A and B respectively the clause $C = \hat{B} \cup (A - \{Rst\})\sigma$ (or $C = \hat{B} \cup (A - \{Rts\})\sigma$). C is called a *paramodulant of* A* *and* B* (and also of B* and A*) and is said to be *inferred by paramodulation from* A* *on*

the variant of Rst (or Rts) *into* B* *on (the occurrence in the particular position in* B* *of)* the variant of u. The variant of the literal Rst (or Rts) is called the *literal of paramodulation* and the occurrence of the variant of u is called the *term of paramodulation.****

Before giving examples to aid in one's understanding of paramodulation, we remark that examination of the definition shows that the property of symmetry is inherent in the inference rule. It is also to be noted that, although in much of what follows we assume the presence of the reflexivity axiom, the definition of paramodulation could have been easily extended to obviate the need for this axiom. The extension does not, however, seem to be of practical value.

Example 1: Premisses {Rab} and {Qa} yield {Qb} by paramodulation or by the usual substitution rule for equality.

Example 2: Premisses {Rab} and {Qx} yield {Qb} by paramodulation. They also yield by paramodulation {Qa}. The substitution rule does not apply in its usual form since neither term, a nor b, occurs in the premiss {Qx}. In many systems from {Qx} one could infer {Qa}, which could then have been used as one premiss together with {Rab} and the substitution rule to yield {Qb}.

Example 3: Premisses {Rab} and {Qx,Px} yield among others the clause {Qb,Pa} by paramodulation. Again in many systems {Qa,Pa} could

---

*** See footnote ** for a comment on how separation of variables works out in practice.

have been inferred from {Qx,Px},providing a premiss to which the substitution rule for equality would apply.

Example 4: Premisses {Rxh(x)} and {Qg(y)} yield by paramodulation {Qh(g(y))}. Note that the presence among the constants occurring in the set of clauses under consideration of the individual constants a and b does not lead by paramodulation to either the inference {Qh(g(a))} or {Qh(g(b))}. Paramodulation (in effect) first finds most general instances of the two premisses which permit straightforward application of the substitution rule for equality and then makes the equality substitution.

Example 5: Consider the premises {Rf(xg(x))e} and {Pyf(g(y)z)z}. For intuitive purposes think of P as product, f as product, and g as inverse. The functions f and g are frequently Skolem functions introduced in place of existential quantifiers. For this application of paramodulation let s be f(xg(x)) and u be f(g(y)z). A most general common instance of u and s is f(g(y)g(g(y))). The inference thus made is {Pyeg(g(y))}. Note that both premisses required non-trivial instantiation in order to apply the substitution rule.

Example 6: Premisses {Rf(xg(x))e,Qx} and {Pyf(g(y)z)z,Qz} yield by paramodulation with s and u as in Example 5 {Pyeg(g(y)),Qg(y),Qg(g(y))}. This example illustrates another way in which paramodulation extends the substitution rule, as both premisses in this example contain more than one literal. The substitution rule applies to pairs of formulae one of which is of the form {Rst}.

The extension in the direction illustrated by Example 6 is needed for R-refutation completeness as can be seen by examining the following set of clauses: {{Rab,Qc}, {$\overline{R}$g(a)g(b),Qc}, {Rcd,$\overline{Q}$c}, {$\overline{R}$g(c)g(d),$\overline{Q}$c}, {Rxx}}.

A most crucial property of paramodulation was illustrated by Example 4. From a theorem-proving viewpoint it is important that the inference rules have the property of being conservative (see Section 6).

6. Refutation Completeness

A set S of clauses *implies (R-implies)* a clause C if no model (R-model) of S condemns C, and S *implies (R-implies)* a set T of clauses if it implies (R-implies) each clause in T. We write $S \models C$, $S \models_R C$, $S \models T$, $S \models_R T$ respectively to express these relationships. If for clauses C and D, $\{C\} \models D$, we also write $C \models D$ and say that C *implies* D (and similarly for $\models_R$). If A implies (R-implies) B, then B is a *consequence (R-consequence)* of A. If S has no model (R-model), then S is *unsatisfiable (R-unsatisfiable)*.

A *deduction* D *of a clause* C_n *from* S *in an inference system* Ω is a finite sequence of pairs (C_i, J_i) $i = 1,2,\ldots,n$ where C_i is a clause and J_i is a "justification" of C_i in terms of one of the rules of inference of Ω and previous steps of D or in terms of membership of C_i in S. (By *inference system* we mean here nothing more than a set of rules of inference.) If such a deduction exists we write $S \vdash_\Omega C_n$. Except when the justifications J_i are of particular interest,[21] D will be identified with the sequence $C_1, C_2, \ldots, C_n$. A *refutation* of S in Ω is a deduction of *false* from S in Ω.

Henceforth Π will be the inference system whose rules of inference are paramodulation, resolution, and factoring and Σ will be that whose rules are just resolution and factoring.

A system Ω is *deduction complete (R-deduction complete)* if for any set S of clauses and any clause C, $S \vdash_\Omega C$ whenever $S \models C$ ($S \vdash_\Omega C$ if $S \models_R C$). It is (R-) *refutation complete* if for any (R-) unsatisfiable S, $S \vdash_\Omega$ *false*.

The system Σ is well known to be refutation-complete.[16] (An adaptation to the present formalism of the simple but rather elegant

proof in [16] is given in Appendix 1.) If Σ were deduction complete as well, one could easily prove (see next paragraph) that Π is R-refutation complete (in the presence of the usual reflexivity axiom). (More generally, any deduction complete system Ω, when augmented by paramodulation, becomes R-refutation complete in the presence of reflexivity.) Both Σ and Π would then, however, be virtually worthless for automatic theorem proving algorithms of the type in use today!

Let Ω be any deduction complete system including the inference rule paramodulation and let S be any R-unsatisfiable set of clauses including $\{Rxx\}$. Consider the tautology $\{\overline{R}xy,Rxy\}$, a trivial logical consequence of S. Deduction completeness of Ω would give $S \vdash_\Omega \{\overline{R}xy,Rxy\}$. By paramodulating from this tautology into $\{Rxx\}$, one could show that $S \vdash_\Omega \{\overline{R}xy,Ryx\}$. By paramodulating from this clause into itself, one can deduce $\{\overline{R}xy,\overline{R}zy,Rxz\}$. Since $\{\{\overline{R}xy,Ryx\}, \{\overline{R}xy,\overline{R}zy,Rxz\}\} \models \{\overline{R}xy,\overline{R}yz,Rxz\}$, it follows by deduction completeness of Ω that $S \vdash_\Omega \{\overline{R}xy,\overline{R}yz,Rxz\}$, completing the equivalence relation properties for R. For the predicate substitution property, consider $P_i^j x_1 \ldots x_k \ldots x_j$, for arbitrary choice of i, j, and k. From the deduction completeness of Ω, $S \vdash_\Omega \{\overline{P}_i^j x_1 \ldots x_k \ldots x_j,\ P_i^j x_1 \ldots x_k \ldots x_j\}$. Paramodulating from symmetry into this clause on x_k in $\overline{P}_i^j x_1 \ldots x_k \ldots x_j$, one could show that $S \vdash_\Omega \{\overline{R}x_{j+1}x_k,\ \overline{P}_i^j x_1 \ldots x_{j+1} \ldots x_j,\ P_i^j x_1 \ldots x_k \ldots x_j\}$. For the function substitution property, since $\{Rxx\} \models \{Rf_i^j(x_1 \ldots x_k \ldots x_j) f_i^j(x_1 \ldots x_k \ldots x_j)\}$, deduction completeness of Ω gives $\{Rxx\} \vdash_\Omega \{Rf_i^j(x_1 \ldots x_k \ldots x_j) f_i^j(x_1 \ldots x_k \ldots x_j)\}$.

Paramodulating from symmetry into this clause on x_k would give $\{Rxx\} \vdash_\Omega \{\overline{Rx}_{j+1}{}^{x_k},\ Rf_i^j(x_1 \ldots x_{j+1} \ldots x_j) f_i^j(x_1 \ldots x_k \ldots x_j)\}$. Hence $S \vdash_\Omega S \cup E$ where E is a set of equality axioms strong enough to guarantee that any model of $S \cup E$ would also be an R-model of S. Suppose then, that S is R-unsatisfiable and includes {Rxx}. $S \cup E$ could then have no model; hence $S \cup E \models$ *false* (vacuously). Again applying the hypothesis of deduction completeness of Ω, one could show $S \cup E \vdash_\Omega$ *false* and hence $S \vdash_\Omega$ *false*. Since S was assumed only to be R-unsatisfiable and to include {Rxx}, this shows that Ω is R-refutation complete in the presence of {Rxx}.

To assess the problems involved in proving refutation completeness of systems that, unlike the hypothetical Ω above, are useful for (present techniques of) automatic theorem proving, one must take care to note that *one of the principal reasons that the systems under study here are useful is that they are devoid of the familiar deduction-completeness properties* of many of the customary formulations of quantification theory, hypothesized for the system Ω above. (So that, for example, the Gödel completeness theorem fails to hold for Σ and Π.) The simplest difference is that universal instantiation is not usually possible in Σ or Π. For example, {Rxa} ⊨ {Raa}, but {Raa} is not deducible from {Rxa} in either Σ or Π. The property that makes systems such as Σ and Π valuable for automatic theorem proving is that they are *conservative* in the following sense: Loosely speaking, an inference rule is called *conservative* if it does not allow a *proper* instance C' of a clause C to be inferred from, say A and B, when C itself could have been inferred.[13] (A *proper* instance C' of C is an instance for which C is not an instance of C'.)

Resolution is a conservative inference rule. From {Pf(x),Qa} and $\{\overline{P}y,Qy\}$ one infers by resolution {Qf(x),Qa} and not, for example, {Qf(a),Qa}. The latter, of course, can be inferred by instantiation followed by a syllogistic inference. The instantiation involved therein is, however, not most general among those permitting syllogistic inference. Factoring is also a conservative inference rule, permitting, in effect, only the most general instantiations that allow application of the equivalence of $p \vee p \vee q$ with $p \vee q$.

Paramodulation, in effect, seeks instantiations that are most general among those that permit equality substitutions. In addition to extending substitution in the two ways already discussed, it is important that certain inferences are avoided. (See Example 4 in Section 5.) From a pair of premisses paramodulation yields an inference which could be made after some number (possibly zero) of applications of instantiation to either or both premisses, followed by an application of an equality substitution rule. It combines instantiation with equality substitution in a way such that the property of being conservative is present. Among the possible instantiations which permit application of equality substitution, paramodulation, in effect, allows only the most general instantiations.

The intuitive comparison above of resolution and paramodulation with syllogism (or *modus ponens*) and substitution, respectively, is intended as motivation for, not as a precise characterization of, resolution and paramodulation. For the latter one must refer to the definitions. The intuitive discussion might, for example, lead one to take $C = (A\sigma - \{k\sigma\}) \cup (B\tau - \{h\tau\})$ as the resolvent instead of

$C = (A - \{k\})_\sigma \cup (B - \{h\})_\tau$ as we intend in the present paper. If $A = \{Pa,Px\}$ and $B = \{\overline{P}a,Qb\}$, we intend the resolvent obtained from choosing Px and $\overline{P}a$ as focal literals to be $C = \{Pa,Qb\}$, not $C = \{Qb\}$.

A set S of clauses is said to be *functionally reflexive* if $\{Rxx\}$ is in S, and for each function letter f_k^j occurring in S, $\{Rf_k^j(x_1 \ldots x_j)f_k^j(x_1 \ldots x_j)\}$ is in S. S is said to be *functionally reflexive relative to* an inference system Ω if $S \vdash_\Omega \{Rxx\}$ and for each f_k^j occurring in S, $S \vdash_\Omega \{Rf_k^j(x_1 \ldots x_j)f_k^j(x_1 \ldots x_j)\}$. The principal result of this section is that the inference system Π is R-refutation complete for sets S that are functionally reflexive (or equivalently, for S functionally reflexive relative to Π). For deduction-complete systems Ω, the condition that S be functionally reflexive relative to Ω reduces to the condition that $S \models \{Rxx\}$. Fortunately for efficiency in automatic theorem proving, Π is not deduction complete; consequently, the additional functional reflexivity properties are not so easily eliminated. (When paramodulation was first introduced a few years ago, we conjectured that $S \vdash_\Pi \{Rxx\}$ was sufficient for R-refutation completeness of Π. No counterexample has yet been discovered to this conjecture, but neither has a proof been given that the completeness theorem proved in this section can be replaced by the simpler and somewhat more attractive conjecture. An earlier, weaker version of the completeness theorem was proved in [15]. There it was required that $S \vdash_\Pi \{Rtt\}$ for all terms $t \in H_S$, a much stronger restriction since there are infinitely many such clauses Rtt, but there are only n+1 functional-reflexivity clauses where n is the number of distinct function symbols occurring in S.

Several systems, such as basic group theory [13], satisfy the hypothesis of the present completeness theorem (Corollary 1) but fail to satisfy the hypothesis of the completeness theorem in [15].)

Lemma 1. Let $A \in S$ have a ground instance (over H_S) $A' = A\tau$ such that in A and A' respectively the terms u* in the literal m* and u' in the literal $m' = m^*\tau$ are in corresponding positions. Then there exists a factor E of A and a substitution μ such that (1) $A' = E\mu$ is an instance of E, (2) E and A' have the same number of literals, and (3) there exists a literal m in E and a term u in m with $m\mu = m'$ and with u in m in the corresponding position to u' in m'.

Proof. The substitution τ induces a partition of the clause A given by j equivalent to k if and only if $j\tau = k\tau$, where j and k are literals in A. The sets A_i of literals of the partition are each unifiable. There exists, therefore, a most general single substitution θ which simultaneously unifies each of the A_i. Since τ unifies each set A_i, there exists a substitution μ with $\tau = \theta\mu$. Let $E = A\theta$. Let $m = m^*\theta$, and let $u = u^*\theta$. The clause E has the desired properties.

Lemma 2. Let A' and B' be respectively ground instances (over H_S) of A and B in S, and let C' be a paramodulant of A' and B'. Further assume that the literal, Rs't', of paramodulation is in A' and that the term, u', of paramodulation is in the literal m' of B'. Let the position of u' in m' be given by the vector $(q_1, q_2, \ldots, q_n)$. Let the substitution ρ and the literal m* in B be such that $B\rho = B'$ and $m^*\rho = m'$. Then if there exists a term u* in m* such that u* and u' are in the corresponding position respectively

in m* and m', then there exists a clause C having C' as an instance such that C is a paramodulant of some factors E and F of A and B respectively.

Proof. From Lemma 1 we can conclude the existence of factors E and F of A and B respectively such that E and F have no variables in common, E has A' as an instance, E and A' have the same number of literals, F has B' as an instance, F and B' have the same number of literals, and F contains a literal m containing a term u with u in m and u' in m' in corresponding positions. Furthermore, we can conclude the existence of substitutions τ and θ with $A' = E\tau$, $B' = F\theta$, $m' = m\theta$, and $u' = u\theta$.

E contains a literal Rst with $Rst\tau = Rs't'$, where Rs't' is the literal of paramodulation in A'. Assume without loss of generality that s' is the argument involved in the inference by paramodulation of C'.

Then s and u have a common instance. Let λ and μ be most general substitutions such that $s\lambda = u\mu$ is a most general common instance of s and u.

Let $\hat{F}$ be obtained from $F\mu$ by replacing $u\mu$ in the position $(q_1, q_2, \ldots, q_n)$ in $m\mu$ by $t\lambda$. Let $C = (E - \{Rst\})\lambda \cup \hat{F}$. C is a paramodulant of E and F having C' as an instance, and the proof is complete.

Lemma 3. Let S be a functionally reflexive set of clauses closed under both paramodulation and factoring. If A' and B' are respectively

ground instances (over H_S) of A and B in S and if C' is a paramodulant of A' and B', then there exists a C in S having C' as an instance.

Proof. Assume without loss of generality that the paramodulation is from A' into B'. Let Rs't' in A' be the literal of paramodulation, and let the literal m' in B' contain the term u' of paramodulation in the position $(q_1,q_2,\ldots,q_n)$. Let ρ be such that $B' = B\rho$.

Case 1. There exists a literal m* in B with $m^*\rho = m'$ and with m* containing a term u* in the position $(q_1,q_2,\ldots,q_n)$. Lemma 2 applies to A, B, A', and B' to yield a clause C having C' as an instance. C is in S since S is closed under both paramodulation and factoring.

Case 2. No such m* exists in B. Let m in B be such that $m\rho = m'$. There exists, therefore, a term (a variable) x in m in the position $(q_1,q_2,\ldots,q_k)$ with $k < n$. Thus there exist $p = n-k$ functions f_i such that ρ contains the element $f_1(\ldots f_2(\ldots(\ldots f_p(\ldots u'\ldots))))/x$.

Let $G_1,G_2,\ldots,G_p$ be the functional reflexivity axioms corresponding to $f_1,f_2,\ldots,f_p$. G_i are in S since S is functionally reflexive.

A set containing $G_1,G_2,\ldots,G_p$ and B and closed under paramodulation contains a clause B_p with the following properties: $B_p\theta = B'$ for some substitution θ, B_p contains a literal m_p with $m_p\theta = m'$, and m_p contains a term u_p with u_p in m_p and u' in m' in corresponding position. By applying the argument of Case 1 to A, B_p, A', and B', we can complete the proof.

Theorem 1. If a functionally reflexive set S is closed under paramodulation and factoring, and if S is R-unsatisfiable, then S is unsatisfiable.

Proof. Assume by way of contradiction that S is satisfiable, and hence, has a model. Let P be the set of atoms over H_S. By the maximal model theorem S has a maximal model M relative to P. We shall show that M is an R-model of S thus contradicting the R-unsatisfiability of S.

Since S is functionally reflexive, Rxx ε S. For arbitrary t ε H_S, therefore, Rtt ε M since M is a model of S.

Consider arbitrary s and t in H_S, and assume Rst in M. Let k and h be atoms (in P) such that h is obtained from k by replacing some one occurrence of s by t. Assume k ε M. By Corollary A there exist ground clauses A' and B' such that $A' \cap M = \{Rst\}$ and $B' \cap M = \{k\}$. Let A and B in S be such that A' and B' are respectively instances of A and B.

Let $C' = (A' - \{Rst\}) \cup (B' - \{k\}) \cup \{h\}$. C' is a paramodulant of A' and B'. By Lemma 3 we can conclude that there exists a C ε S having C' as an instance.

Since M is a model of S, the intersection of M and C' is not empty. This intersection, therefore, equals {h}, and so h ε M. Thus M has been shown to be an R-model, and a contradiction is reached. The proof is complete.

The desired proof, in the presence of functional reflexivity, of the R-refutation completeness of Π follows as a corollary from Theorem 1.

Corollary 1. For sets (consisting of clauses) functionally reflexive relative to Π, Π is R-refutation complete.

Proof. Let S be a functionally reflexive relative to Π, R-unsatisfiable set of clauses and let S* be its closure under paramodulation, resolution, and factoring. S* is therefore functionally reflexive. By Theorem 1, S* is unsatisfiable. By the Resolution Theorem (see Appendix 1), *false* must then be in S*. But every clause C in S* is the last line of a deduction $D_1,\ldots,D_n$ of C from S in Π with $D_i \in S^*$ for $i = 1,\ldots,n$. Hence S* contains a refutation of S in Π.

Corollary 2. For finite or effectively enumerable sets S that are functionally reflexive relative to Π, there is a semi-decision procedure for R-unsatisfiability.

The proof of this corollary is a form of a well known argument, whose details are supplied for the reader who may be unfamiliar with the formalism of paramodulation and resolution. The semidecision procedure exhibited in the proof is chosen to make the argument transparent. It is *not* recommended as an efficient theorem-proving procedure; efficient procedures generally require a number of "strategies" for determining the sequence in which inferences are to be made and for suppressing unwise applications of the rules of inference.

Proof. Let S be a finite or effectively enumerable R-unsatisfiable set of clauses functionally reflexive relative to Π. Consider an effective enumeration $C_1,C_2,\ldots$ of S. Let $W^0 = \emptyset$. For $i > 0$, let $W^i = W^{i-1} \cup \{C_i\} \cup Q^i$, where Q^i is the set of all clauses C such that C is a paramodulant, resolvent, or factor of clauses in W^{i-1} and no

lexically earlier variant of C has that property.**** Each W^i is finite and can be effectively generated from W^{i-1}. For each clause C in the set S* defined in the proof of Corollary 1, $\bigcup_i W^i$ contains a variant of C. Hence *false* ε $\bigcup_i W^i$ and thus is in some W^n.

To complete the proof, note that for any clause D, $S \vdash_\Pi D$ only if $S \models_R D$, so that $S \vdash_\Pi$ *false* only if S is R-unsatisfiable.

By examining the proofs of Theorem 1 and Corollaries 1 and 2, we can easily obtain parallel results about an inference system in which paramodulation is subject to the following constraint: allow paramodulation only when the term of paramodulation is contained in a positive literal. Thus, in the clause {Pb, $\overline{Q}$b}, only the first occurrence of b would be admitted as a term of paramodulation. One can then prove, for example: if a functionally reflexive set S is closed under paramodulation (restricted by the constraint above) and factoring, and if S is R-unsatisfiable, then S is unsatisfiable. The proof is identical to that given in Theorem 1. The parallel to Corollary 1 also holds when paramodulation is so constrained. If the semidecision procedure of Corollary 2 is modified by constraining paramodulation as above, the property of semidecidability is not lost.

**** From any pair of clauses that have at least one non-ground resolvent, one can obtain infinitely many resolvents since any non-ground clause has infinitely many variants and any variant of a resolvent of A and B is also a resolvent of A and B. Hence the need to select a single representative in order that Q^i and W^i remain finite. A similar situation exists for paramodulation and factoring.

Appendix 1.

Refutation Completeness Theorem for Σ: If S is any unsatisfiable set of clauses, then $S \vdash_{\Sigma} false$.

Proof. (Adapted from [16].) Let Y be the closure of S under resolution and factoring. Since for each clause C in Y, $\bar{Y}$ contains (all the steps of) a deduction of C from S in Σ, we need only show that $false \in Y$. Consider an ordering $a_1, a_2, \ldots$ of the Herbrand atoms for S. Let $I_o = \emptyset$, and for $j > 0$, let $I_j = I_{j-1} \cup \{a_j\}$ if $I_j \cup \{\bar{a}_j\}$ condemns some clause in Y, otherwise let $I_j = I_{j-1} \cup \{\bar{a}_j\}$. Let $I = \bigcup_j I_j$. I is an interpretation of Y, but cannot be a model of Y, since it would then be a model of S, but S is unsatisfiable. Hence I condemns some clause C in Y, that is, for some ground instance C' of C, the negation of each literal of C' is in I. Since there are only finitely many literals in C', C' must be condemned by I_j for some finite j. Let j be the smallest non-negative integer such that I_j condemns (some ground instance C' of) some clause C in Y. The following establishes that $j = 0$: Suppose that $j > 0$. Then I_{j-1} condemns no clause in Y while I_j condemns some clause C. It must be that $I_j = I_{j-1} \cup \{a_j\}$ since $I_j = I_{j-1} \cup \{\bar{a}_j\}$ only if $I_{j-1} \cup \{\bar{a}_j\}$ condemns no clause of Y, and by hypothesis I_j condemns C. But $I_{j-1} \cup \{\bar{a}_j\}$ must also condemn some (ground instance D' of some) clause D in Y, since otherwise I_j would be $I_{j-1} \cup \{\bar{a}_j\}$. I_{j-1} by hypothesis condemns neither C' nor D'. The literal $\bar{a}_j$ must therefore occur in C' and a_j must occur in D'. Hence there must be a set C* of literals contained in C with a substitution σ such that $C^*\sigma = \{\bar{a}_j\}$ and $(C-C^*)\sigma = C' - \{\bar{a}_j\}$ and a set $D^* \subseteq D$ with a substitution τ such that

$D^*\tau = \{a_j\}$ and $(D-D^*)\tau = D' - \{a_j\}$. Let σ^* and τ^* be most general unifiers for C^* and D^* respectively. Then $C_1 = C\sigma^*$ and $D_1 = D\tau^*$ are factors of C and D respectively and must thus be in Y. The factors C_1 and D_1 then resolve on the literals $C^*\sigma^*$ and $D^*\tau^*$ to give a clause E also in Y. The clause $E' = (C' - \{\bar{a}_j\}) \cup (D' - \{a_j\})$ is a ground instance of E. Since $I_{j-1} \cup \{a_j\}$ condemns C', I_{j-1} must condemn $C' - \{\bar{a}_j\}$; and similarly since $I_{j-1} \cup \{\bar{a}_j\}$ condemns D', I_{j-1} must condemn $D' - \{a_j\}$. Hence I_{j-1} must condemn E, contrary to hypothesis. Hence $J \neq 0$, so j must be zero.

Thus $I_o = \emptyset$ condemns some clause C_o in Y. But this is possible only if $C_o = \textit{false}$. Hence $\textit{false} \in Y$, completing the proof.

To see that factoring is needed for refutation completeness one need only consider the example where S is composed of the clauses $\{Q_a\}$, $\{Px,Py\}$, and $\{\bar{P}w,\bar{P}z\}$. Then the closure under resolution alone includes only variants of clauses in this unsatisfiable set S and of $\{\bar{P}u,Pt\}$, but does not include *false*, so no refutation can be obtained.

Resolution is sometimes defined in such a way as to subsume factoring. While this appears to simplify a few definitions and theorems, it leads, we feel, to undesirable consequences when applied to automatic theorem proving. Hence we prefer the older formulation in terms of separate rules for resolution and factoring.

Alternate Proof of the Maximal Model Theorem:

Since the proof of the maximal model theorem appearing in Section 3 was first given,[19] a number of other proofs have been found. The simplest we have been able to work out to date is obtained by adapting

an idea found in the proof of Lindenbaum's Lemma that is given in [11]. Briefly, the alternative proof thus obtained runs as follows: Consider an enumeration $a_1, a_2, \ldots$ of the interpretation T. Let $b_j = \bar{a}_j$ if $S \cup \{\{b_1\}, \ldots, \{b_{j-1}\}, \{\bar{a}_j\}\}$ is satisfiable, otherwise let $b_j = a_j$. Let $Q_o = \emptyset$ and for $j > 0$ let $Q_j = \{\{b_1\}, \ldots, \{b_j\}\}$. Let $M = \{b_1, b_2, \ldots\}$. Then M is an interpretation, and furthermore if M is a model from the manner of construction, it must be a maximal one relative to T. To see that M is a model, suppose that it is not. Then, since M is an interpretation, it must be that M condemns some clause in S. Hence the set $W = S \cup \{\{b_1\}, \{b_2\}, \ldots\}$ must be unsatisfiable. Some finite subset W' of W must then be unsatisfiable. Without loss of generality, let W' be such that no proper subset of W' is unsatisfiable. Then let $W' \cap \{\{b_1\}, \{b_2\}, \ldots\} = \{\{b_{j_1}\}, \ldots, \{b_{j_k}\}\}$ and let n be the largest subscript on b appearing therein. Now for $j > 0$, $S \cup Q_j$ must be satisfiable if $S \cup Q_{j-1}$ is, since otherwise both $S \cup Q_{j-1} \cup \{a_j\}$ and $S \cup Q_{j-1} \cup \{\bar{a}_j\}$ would be unsatisfiable. Hence, since S is satisfiable, $S \cup Q_n$ must by induction be satisfiable, which contradicts the unsatisfiability of its subset W'. Thus M condemns no clause in S and must be a model of S, hence a maximal one relative to T.

The construction given in the proof of the Refutation Completeness theorem for Σ can also be made to yield a maximal model. Failure to recognize the importance of closure of Y under resolution and factoring to the modelhood of I has, however, led to a number of spurious "constructive" proofs of the maximal model theorem. Indeed, if S is finite and the closure restriction is satisfied by S itself, I is a model, and an effective means of enumerating I is provided by the technique given in the refutation completeness proof

for Σ. Luckham has successfully applied a similar technique[9] to a useful special case involving a finite set of ground clauses. It appears[10] that attempts to give a proof of the general Maximal Model theorem embodying an effective enumeration of the maximal model may be foredoomed by the existence in first-order logic of sentences for which there exist no effectively enumerable models, hence no such maximal ones.

Appendix 2.

In this section we list a number of theorems from group theory and some from ring theory whose proofs have been obtained by one of our theorem-proving programs. One should not draw conclusions from comparing the times given to obtain proof, since the times obtained were affected materially by the fact that a number of different programs were employed. We shall also include the proofs of two theorems from group theory to illustrate the inference rule, paramodulation.

1. In a group, if $xy = e$, then $yx = e$. The time required to obtain a proof was 815 milliseconds.

2. If $xy = e$ and $zy = e$, $x = z$. 5.3 sec.

3. The right inverse of the right inverse of $x = x$. 326 milliseconds.

4. $(x^{-1})^{-1} = x$. 998 milliseconds.

Example 3 differs from example 4, obviously, by not treating inverse as known to be two sided.

5. A non-empty subset H of a group is a subgroup if and only if, for every x and y in H, xy^{-1} is in H. The necessity was proved in 1.4 sec., assuming that identity and inverse of the subgroup were that of the group as a whole. The sufficiency was proved as a series of lemmas. That H contains e required 89 milliseconds to prove; for every x in H, H contains x^{-1}, 222 milliseconds; H is closed under multiplication, 32 sec.

6. Exponent 2 implies commutativity. 8 sec.

7. The axioms of right identity and right inverse are dependent axioms. The proofs were obtained respectively in 2 seconds and 3+ seconds.

8. Exponent 2 implies commutativity, but using only those axioms sufficient to obtain a proof. 460 milliseconds. The addition of various

lemmas, just as the presence of a full axiom set, can help or hurt the proof-search. In problems of some depth, lemmas will almost certainly be needed.

9. Subgroups of index 2 are normal. The theorem was proved by dividing it into two cases, as is often done in the standard proof. The proof for invariance with respect to elements of the subgroup was obtained in 4.9 sec., for elements outside the subgroup in 27.5 sec.

10. Boolean rings have characteristic 2. 41.7 sec.

11. Boolean rings are commutative, assuming the lemma of characteristic 2. 12+ min.

12. If a set is closed under an associative operation, contains an element e with $e^2 = e$, and every element has (with respect to e) at least one left inverse and at most one right inverse, then the set is a group. A proof was obtained in 5.8 sec.

Of the examples given, example 12 is clearly the most difficult, mathematically speaking. The theorem is not easy to prove for the average student of group theory.

On the other hand, 18.2.8 on page 322 of [7] has not been provable by a computer. The lemma states that exponent 3 implies ((a,b),b) = e for all a and b in the group. The proof is shortened considerably (see appendix of [13]) by the addition of paramodulation to the most successful of the previous theorem-proving systems. That previous system was based on resolution.

We now give a proof, employing paramodulation, of the theorem that the axiom of right inverse is dependent on the set consisting of left identity, left inverse, and associativity. In the following, R is read

as equals, f as product, g as inverse, and e as the left identity. Assume by way of contradiction that there exists an element a lacking a right inverse. The resulting clause is $\{\overline{R}f(ay)e\}$. Braces and commas will be omitted in the following.

Proof.

1. $Rf(ex)x$. Left identity
2. $Rf(g(x)x)e$. Left inverse
3. $Rf(xf(yz))f(f(xy)z)$. Associativity
4. $\overline{R}f(ay)e$.
5. $\overline{R}f(f(ea)y)e$, paramodulate 1 on x into 4 on a.
6. $\overline{R}f(f(f(g(w)w)a)y)e$, 2 into 5 on first occurrence of e.
7. $\overline{R}f(f(g(w)f(wa))y)e$, 3 into 6 on $f(f(g(w)w)a)$.
8. $\overline{R}f(f(g(g(a))e)y)e$, 2 into 7 on $f(wa)$.
9. $\overline{R}f(g(g(a))f(ey))e$, 3 into 8 on $f(f(g(g(a))e)y)$.
10. $\overline{R}f(g(g(a))y)e$, 1 into 9 on $f(ey)$.
11. *false*, resolve 10 and 2.

For the 12th example, the following clauses were input to the theorem-proving program, PG5, [20] and the proof which follows the input (in essence) was obtained in 5.8 sec.

1. $Pxyf(xy)$. Closure
2. $Peee$.
3. $Pg(x)xe$. Left inverse
4. $\overline{P}xyu\ \overline{P}yzv\ \overline{P}uzw\ Pxvw$. } Associativity
5. $\overline{P}xyu\ \overline{P}yzv\ \overline{P}xvw\ Puzw$. }
6. $\overline{P}xye\ \overline{P}xze\ Ryz$.
7. Rxx. Reflexivity

8. $\overline{P}xyu\ \overline{P}xyv\ Ruv.$ — Well-definedness
9. $\overline{R}xy\ \overline{R}yz\ Rxz.$ — Transitivity
10. $\overline{R}uv\ \overline{P}uxy\ Pvxy.$
11. $\overline{R}uv\ \overline{P}xuy\ Pxvy.$
12. $\overline{R}uv\ \overline{P}xyu\ Pxyv.$
13. $\overline{R}uv\ Rf(ux)f(vx).$
14. $\overline{R}uv\ Rf(xu)f(xv).$
15. $\overline{R}uv\ Rg(u)g(v).$

(Substitutivity of equals)

16. $\overline{P}eaa.$ — e is not a left identity

Proof.

1. $Peee.$
2. $\overline{P}xyu\ \overline{P}yzv\ \overline{P}uzw\ Pxvw.$
3. $\overline{P}xye\ \overline{P}xze\ Ryz.$
4. $\overline{R}uv\ \overline{P}xyu\ Pxyv.$
5. $\overline{P}eaa.$
6. $Pg(x)xe.$
7. $Pxyf(xy).$
8. $\overline{R}f(xy)v\ Pxyv$, resolve clause 4 on second literal against clause 7.
9. $\overline{R}f(ea)a$, 8_2 (i.e., second literal of clause 8) vs. 5.
10. $\overline{P}yzv\ \overline{P}f(xy)zw\ Pxvw$, 2_1 vs. 7.
11. $\overline{P}f(xe)ew\ Pxew$, 10_1 vs. 1.
12. $Pxef(f(xe)e)$, 11_1 vs. 7.
13. $\overline{P}g(z)ye\ Ryz$, 3_2 vs. 6.
14. $\overline{P}g(a)f(ea)e$, 13_2 vs. 9.
15. $\overline{P}yzv\ \overline{P}ezw\ Pg(y)vw$, 2_1 vs. 6.
16. $\overline{P}ezw\ Pg(y)f(yz)w$, 15_1 vs. 7.

17. $Pg(y)f(ye)e$, 16_1 vs. 1.

18. $Rf(xe)x$, 13_1 vs. 17.

19. $\overline{P}ezv\ \overline{P}f(f(xe)e)zw\ Pxvw$, 2_1 vs. 12.

19'. $\overline{P}ezv\ \overline{P}xzw\ Pxvw$, demodulate [Reference 20] second literal of 19 with 18.

20. $\overline{P}xzw\ Pxf(ez)w$, $19'_1$ vs. 7.

21. $Pg(x)f(ex)e$, 20_1 vs. 6.

22. *false*, 21 vs. 14.

PG5 gave as output only 1 thru 5, 9, 12, 14, and 21.

The following proof is also of example 12 but employing paramodulation as an additional inference rule.

Proof.

1. $Rf(ee)e$.

2. $Rf(xf(yz))f(f(xy)z)$.

3. $\overline{R}f(xy)e\ \overline{R}f(xz)e\ Ryz$.

4. $Rf(g(x)x)e$.

5. $\overline{R}f(ea)a$.

6. $\overline{R}f(g(z)y)e\ Ryz$, resolve 3_2 vs. 4.

7. $\overline{R}f(g(a)f(ea))e$, resolve 6_2 and 5.

8. $Rf(g(y)f(yz))f(ez)$, paramodulate 4 into 2 on $f(xy)$.

9. $Rf(g(y)f(ye))e$, paramodulate 1 into 8 on $f(ez)$.

10. $Rf(ye)y$, resolve 6_1 and 9.

11. $Rf(xf(ez))f(xz)$, paramodulate 10 into 2 on $f(xy)$.

12. $Rf(g(x)f(ex))e$, paramodulate 4 into 11 on $f(xz)$.

13. *false*, resolve 7 and 12.

Acknowledgements

The authors are deeply indebted to W. F. Miller for his support and encouragement of the automatic theorem-proving effort over the years; to D. L. Luckham and C. C. Green; and most particularly to W. W. Boone for his extensive and invaluable assistance in the preparation and criticism of this paper. An important portion of the fundamental research underlying the results in this paper was supported by the Computer Science Department of the University of Wisconsin and the University of Wisconsin Computing Center during 1966-67. The current work is supported by the U. S. Atomic Energy Commission.

BIBLIOGRAPHY

1. Church, A., Introduction to Mathematical Logic I, Princeton (1956).

2. Church, A. , A Note on the Entscheidungsproblem, J. Symb. Logic, 1, pp. 40-41, 101-102 (1936).

3. Darlington, J., Automatic theorem proving with equality substitutions and mathematical induction. Machine Intelligence 3 (ed. D. Michie), Edinburgh: Oliver and Boyd (1968).

4. Davis, M. and Putnam, H., A computing procedure for quantification theory, J. Assn. Comput. Mach. 7, pp. 201-215 (1960).

5. Gilmore, P. C., A proof method for quantification theory: its justification and realization, IBM Journal, (Jan. 1960), pp. 28-35.

6. Green, C., Theorem-proving by Resolution as a Basis for Question-answering Systems, Machine Intelligence 4, (ed. B. Meltzer and D. Michie)(1969).

7. Hall, Marshall, The Theory of Groups, Macmillan Co., New York (1959), p. 322.

8. Kowalski, R. and Hayes, P., Semantic trees in automatic theorem proving, Machine Intelligence 4 (ed. B. Meltzer and D. Michie)(1969).

9. Luckham, D., Some tree-paring strategies for theorem-proving, Machine Intelligence 3, (ed. D. Michie), Edinburgh Univ. Press, Edinburgh (1968).

10. Luckham, D., Personal communication (1969).

11. Mendelson, E., Introduction to Mathematical Logic, Van Nostrand (1964).

12. Quine, W., A proof procedure for quantification theory, J. Symb. Logic, 20, pp. 141-149 (1955).

13. Robinson, G.and Wos, L., Paramodulation and theorem-proving in first-order theories with equality. Machine Intelligence 4 (ed. B. Meltzer and D. Michie)(1969).

14. Robinson, G., Wos., L., and Carson, D., Some theorem-proving strategies and their implementation. AMD Tech. Memo No. 72, Argonne National Laboratory (1964).

15. Robinson, G. and Wos, L., Completeness of paramodulation, J. Symb. Logic 34, p. 160 (abstract)(1969).

16. Robinson, J., A machine-oriented logic based on the resolution principle, J. Assn. Comput. Mach. 12, pp. 23-41 (1965).

17. Tarski, A., Mostowski, A., and Robinson, R., Undecidable theories, North Holland (1953).

18. Wos, L., Carson, D., and Robinson, G., The unit-preference strategy in theorem proving. AFIPS Conference Proceedings 26, Spartan Books, Washington, D. C., pp. 615-621 (1964).

19. Wos, L., and Robinson, G., The maximal model theorem, J. Symb. Logic 34, pp. 159-160 (abstract)(1969).

20. Wos, L., Robinson, G., Carson, D., and Shalla, L., The concept of demodulation in theorem proving, J. Assn. Comput. Mach. 14, pp. 698-709 (1967).

21. Wos, L. and Robinson, G., Paramodulation and Set of Support, IRIA Symposium on Automatic Demonstration, Versailles, 1968. (Proceedings forthcoming in Lecture Notes in Mathematics, Springer-Verlag.)