J. Symbolic Computation (1992) 13, 175–207

Discriminator Varieties and Symbolic Computation

STANLEY BURRIS

Dept. of Pure Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

(snburris@thoralf.waterloo.edu)

(Received 2 August 1989)

We look at two aspects of discriminator varieties which could be of considerable interest in symbolic computation:

1. discriminator varieties are unitary (i.e., there is always a most general unifier of two unifiable terms), and
2. every mathematical problem can be routinely cast in the form[†]

$$p_1 \approx q_1, \ldots, p_k \approx q_k \text{ implies the equation } x \approx y.$$

Item (1) offers possibilities for implementations in computational logic, and (2) shows that Birkhoff's five rules of inference for equational logic are all one needs to prove theorems in mathematics.

There are seven sections in these notes. Section 1 is a quick tour through the basic concepts from Universal Algebra which we will be using — and some basic properties of discriminator varieties. In section 2 the connection between discriminator terms and simple algebras in discriminator varieties is discussed, and we look at important examples of discriminator varieties from the literature. In section 3 we prove that all discriminator varieties are unitary, give examples of most general unifiers in several of the examples from section 2, and look at the unification problem. In section 4 **Birkhoff's** five rules of inference for equational logic are reviewed, and in section 5 **McKenzie's** reduction of first-order logic to equational logic using discriminator varieties is discussed. In section 6 we look at the **von Neumann-Bernays-Gödel** axioms for set theory, and the ultimate reduction of all mathematics to equational logic. We do not claim any new results in sections 4–6; however it is hoped that bringing together these aspects of equational logic will focus attention on discriminator varieties as a fascinating tool for further computational research. In section 7 there is a proof of the completeness of McKenzie's reduction. (Our terminology will follow that of Burris & Sankappanavar (1981).)

1. Background from Universal Algebra

Universal algebra has focused a great deal of attention on equations, and on the classes of algebras defined by equations — called *equational classes*. A class of algebras

[†]Indeed one could reduce this to the form "$p(x_1, \ldots, x_n) \approx x_1$ implies $x \approx y$" (see Theorem 5.3); it is not clear that one would make any computational gains by such a further reduction.

0747-7171/92/020175+33 $03.00/0

closed under homomorphic images, subalgebras and direct products is called a *variety*. In the mid-thirties Birkhoff proved that equational classes are the same as varieties. (We prefer to use the word variety — it is a shorter than equational class.) The smallest variety $\mathrm{V}(K)$ containing a given class K of algebras is called the *variety generated by* K. The equations true of $\mathrm{V}(K)$ are the same as those true of K.

In the mid-forties Birkhoff showed that every variety V is generated by the class V_{SI} of *subdirectly irreducible algebras*[†] in V. For example the subdirectly irreducible Boolean algebras are the 1- and 2-element Boolean algebras. Consequently an equation is true in Boolean algebras iff it is true in the 2-element Boolean algebra — this fact justifies the use of *truth-tables* in the calculus of classes. In view of Birkhoff's theorem one is justified in thinking of V_{SI} as *generalized truth-tables* for V, a central part of V from which one can recover valuable information about V, e.g., the equational theory of V.

The class V_{S} of *simple* algebras in V is a popular subclass of V_{SI}, and consists of those algebras which have at most two congruences. As an example we mention that the classification of finite simple groups has been an exciting topic in recent years. One-element algebras are called *trivial algebras* — our convention is to include them in the subdirectly irreducible and simple algebras. However many references will exclude them.

FACT 1.1. Given an algebra $\mathbf{A}$ and a congruence θ of $\mathbf{A}$ we have: $\mathbf{A}/\theta$ is a nontrivial simple algebra iff θ is a maximal congruence.

(See Burris & Sankappanavar (1981), p. 59, Th. 8.9.)

In general V_{S} does not generate V, so an equation true of V_{S} need not be true of all algebras in V. However **discriminator varieties V are generated by V_{S}**. (We defer the definition and examples of discriminator varieties till the next section.) If a variety V is such that $V_{\mathrm{SI}} = V_{\mathrm{S}}$ then we say it is *semisimple*.

FACT 1.2. Suppose $\mathbf{A}$ is in a semisimple variety and $a, b \in \mathrm{A}$. Then

$$a = b \qquad \text{iff} \qquad a/\theta = b/\theta$$

for every maximal congruence θ of $\mathbf{A}$.

(This follows from the fact that the intersection of the maximal congruences of $\mathbf{A}$ is the identity relation.)

FACT 1.3. Discriminator varieties are semisimple.

(See Burris & Sankappanavar (1981), p. 165, Th. 9.4.)

The *free algebras* in V also play the role of generalized truth-tables. A V-free algebra freely generated by X, written $\mathbf{F}_V(X)$, is nothing more than an **algebra of normal forms** $\bar{p}$ of terms p over the set X. The key properties of $\mathbf{F}_V(X)$ are:

[†]An algebra $\mathbf{A}$ is *subdirectly irreducible* if it is a one-element algebra or there is a congruence μ of $\mathbf{A}$ which is not the identity relation and has the property $\mu \subseteq \theta$ for every congruence θ of $\mathbf{A}$ which is not the identity relation. (Such a minimum congruence μ among the nonidentity congruences of $\mathbf{A}$ is called the *monolith* of $\mathbf{A}$.)

1. $V \models p \approx q$ iff $\bar{p} = \bar{q}$
2. $f(\bar{t}_1, \ldots, \bar{t}_n) = \overline{f(t_1, \ldots, t_n)}$ for f a fundamental n-ary operation
3. For any algebra $\mathbf{A}$ in V and any map $\alpha : X \to \mathrm{A}$, α extends to a homomorphism $\beta : \mathbf{F}_V(X) \to \mathbf{A}$, i.e., for f a fundamental n-ary operation we have

$$\beta f(\bar{t}_1, \ldots, \bar{t}_n) = f(\beta\bar{t}_1, \ldots, \beta\bar{t}_n).$$

Thus we have

$$\beta(\overline{f(t_1, \ldots, t_n)}) = f(\beta\bar{t}_1, \ldots, \beta\bar{t}_n).$$

Given V and X there may be many ways to choose normal forms for terms over X, but the resulting algebras of normal forms are isomorphic, so we can speak of *the* V-free algebra $\mathbf{F}_V(X)$. And since $\mathbf{F}_V(X)$ is determined up to isomorphism by the cardinality of X we can simply refer to the V-free algebras $\mathbf{F}_V(0), \ldots, \mathbf{F}_V(n), \ldots, \mathbf{F}_V(\omega), \ldots$, i.e., we merely specify the size of the set X of free generators.

For example let V be the variety of semigroups. Then we can let $\mathbf{F}_V(X)$ be X^+, the semigroup of nonempty strings over X. The string xyz would be the normal form for the terms $x(yz)$ and $(xy)z$. For monoids we would use X^*, and for rings we would use polynomials over X with integer coefficients, where the monomials are linearly ordered (to provide unique normal forms). The normal forms in these varieties are finite objects — however when working with general varieties it is not clear how to find a canonical normal form which is a finite syntactic object, so algebraists solve the issue by taking *the set of all terms equivalent to a given term* p as the normal form $\bar{p}$ of p. This may not be convenient for computational purposes, but it is a simple choice.

Birkhoff noted that $V \models p \approx q$ iff $\mathbf{F}_V(X) \models p \approx q$ for any set X whose size is at least as large as the number of variables in the equation $p \approx q$. Consequently we see that the free algebras in V do indeed play the role of generalized truth-tables.

At this point we have sufficient information, when combined with the discussion of simple algebras in discriminator varieties in the next section, to show that discriminator varieties have unitary unification type. But to study the *unification problem* for discriminator varieties we will need a deeper structure theorem for algebras in discriminator varieties. This theorem says roughly that such an algebra decomposes into the study of a Boolean algebra and the simple quotients of the algebra. The next paragraphs give a precise formulation of this structure theorem, and the important consequences.

A subalgebra $\mathbf{A}$ of a direct product $\Pi\mathbf{A}_i$ is a *subdirect product* if for each i and each $a_i \in \mathrm{A}_i$ we have an $a \in \mathrm{A}$ such that $a(i) = a_i$, i.e., $\mathbf{A}$ "touches" every element of every coordinate $\mathbf{A}_i$. Birkhoff proved that every algebra in a variety V is isomorphic to a subdirect product of algebras in V_{SI}. Because of this the members of V_{SI} are regarded as "building blocks" of the variety V. However in practice these building blocks fit together in extremely complicated ways — and (in the study of first-order properties) Birkhoff's theorem seems to offer little information about V from V_{SI} beyond that of the equational theory of V_{SI}.

In 1966 Dauns & Hofmann extracted a highly specialized subdirect product (from the work of Grothendieck concerning sheaves used in the study of algebraic geometry) which in their terminology would be described as *an algebra of global sections of a Hausdorff sheaf over a Boolean space*, and for which Burris & Werner (1979) introduced the simple

phrase *a Boolean product*† — for such products one can analyze not only equations but also *primitive positive* sentences φ, i.e., sentences of the form $\exists \vec{x} \bigwedge p_i(\vec{x}) \approx q_i(\vec{x})$, in terms of the behavior of the stalks. (We note that in the study of Boolean products the coordinate algebras are called *stalks*.)

FACT 1.4. A Boolean product $\mathbf{A}$ satisfies a primitive positive sentence φ iff each stalk of $\mathbf{A}$ satisfies φ.

(See Burris & Werner (1979), p. 272, Lemma 1.1.)

In other words, we can solve a system of equations in a Boolean product iff we can solve it in every stalk of the Boolean product; and a given sequence of elements is a solution iff it provides a solution on each stalk. This property fails dramatically with typical subdirect products. Much of what we know about discriminator varieties depends on the following result of Bulman-Fleming & Werner (1977):

FACT 1.5. Every member of a discriminator variety is isomorphic to a Boolean product of simple algebras from the variety.

(See Burris & Sankappanavar (1981), p. 165, Th. 9.4.)

FACT 1.6. Given a discriminator variety V and a primitive positive sentence φ we have

$$\mathbf{F}_V(n) \models \varphi \qquad \text{iff} \qquad S_n \models \varphi$$

where S_n is the set of $(\leq n)$-generated simple algebras in V.

(This follows from Facts 1.4 and 1.5 since the stalks of an n-generated Boolean product are $(\leq n)$-generated.)

2. Discriminator Varieties

A ternary term $t(x, y, z)$ is a *discriminator term* for an algebra $\mathbf{A}$ if, for $a, b, c \in A$,

$$t(a,b,c) = \begin{cases} c & \text{if } a = b \\ a & \text{if } a \neq b. \end{cases} \tag{1}$$

A term $s(x, y, u, v)$ is a *switching term* for an algebra $\mathbf{A}$ if, for $a, b, c, d \in A$,

$$s(a,b,c,d) = \begin{cases} c & \text{if } a = b \\ d & \text{if } a \neq b. \end{cases} \tag{2}$$

†A subalgebra $\mathbf{A}$ of a direct product $\Pi_{x \in X} \mathbf{A}_x$ is called a *Boolean product* if

1. it is a subdirect product
2. X is a Boolean space (i.e., a compact Hausdorff space with a basis of closed-open sets)
3. (*Equalizers are Clopen*) for $f, g \in A$ we have $\{x \in X : f(x) = g(x)\}$ is a closed-open subset of X
4. (*Patchwork Property*) given $f, g \in A$ and a closed-open subset N of X, the element h of $\Pi_{x \in X} A_x$ which equals f on N and g on $X \setminus N$ is also in A.

From a discriminator term we can construct a switching term, and vice-versa, by

$$s(x,y,u,v) = t(t(x,y,u),t(x,y,v),v) \quad (3)$$
$$t(x,y,z) = s(x,y,z,x). \quad (4)$$

Finite algebras with a discriminator term are called *quasiprimal* algebras. Quasiprimal algebras with no proper subalgebras and no nontrivial automorphisms are called *primal* algebras.† In the examples below one sees several classical examples of primal algebras: the 2-element Boolean algebra (or Boolean ring), the rings $\mathbf{Z}_p$, and the Post-algebras $\mathbf{P}_n$. Finite fields are the best known examples of quasiprimal algebras.

A variety V of algebras is a *discriminator variety*‡ if there is a class K of algebras which generates V such that there is a ternary term $t(x,y,z)$ which is a discriminator term for every member of K.

The existence of a discriminator term $t(x,y,z)$ for a class K of algebras leads to a powerful Boolean product structure theorem for $\mathrm{V}(K)$, the variety generated by K, and this in turn has been involved in significant results concerning decidability and existentially closed structures. As it turns out, $t(x,y,z)$ *will be a discriminator term for precisely the simple algebras in the variety* $\mathrm{V}(K)$; and the structure theorem says that every member of the variety $\mathrm{V}(K)$ is isomorphic to a Boolean product of simple algebras in the variety. We need such basic properties of discriminator varieties to understand their possibilities in symbolic computation.

Now we will list important examples of discriminator varieties, giving a set of defining equations for each one, a generating set K, a description of all the simple algebras in the variety, and a discriminator term and switching term for K (and hence for the simple algebras in the variety).

Examples of Discriminator Varieties

1. Boolean Algebras $\langle \mathrm{B}, \vee, \wedge, ', 0, 1\rangle$

Axioms:
$$x \vee (y \vee z) \approx (x \vee y) \vee z \qquad x \wedge (y \wedge z) \approx (x \wedge y) \wedge z$$
$$x \vee y \approx y \vee x \qquad x \wedge y \approx y \wedge z$$
$$x \vee (x \wedge y) \approx x \qquad x \wedge (x \vee y) \approx x$$
$$x \wedge 1 \approx x \qquad x \vee 0 \approx x$$
$$x \wedge x' \approx 0 \qquad x \vee x' \approx 1$$
$$x \wedge (y \vee z) \approx (x \wedge y) \vee (x \wedge z).$$

$K = \{\mathbf{2}_{\mathrm{BA}}\}$ where $\mathbf{2}_{\mathrm{BA}}$ is the two-element Boolean algebra $\langle \{0,1\}, \vee, \wedge, ', 0, 1\rangle$

Simples: just the 1- and 2-element Boolean algebras.

$t(x,y,z) = (x \wedge z) \vee (x \wedge y' \wedge z') \vee (x' \wedge y' \wedge z)$

$s(x,y,u,v) = (x \wedge y \wedge u) \vee (x' \wedge y' \wedge u) \vee (x \wedge y' \wedge v) \vee (x' \wedge y \wedge v).$

†At present it appears that it is extremely difficult to test finite algebras for being quasiprimal, or primal; but we have not been able to find a proof of this.

‡The two basic references on discriminator varieties are Chapter IV of "A Course in Universal Algebra" by Burris & Sankappanavar (1981), and "Discriminator Algebras" by Werner (1978).

Boolean algebras, axiomatized in full generality by Huntington in 1904, form the original discriminator variety. Because of the need to work with the propositional connectives in clausal logic they have been the subject of considerable interest in the field of automated deduction.

2. BOOLEAN RINGS $\langle B, +, \cdot, 0, 1\rangle$

AXIOMS:
$$x + (y + z) \approx (x + y) + z \qquad x \cdot (y \cdot z) \approx (x \cdot y) \cdot z$$
$$x + y \approx y + x \qquad x \cdot y \approx y \cdot x$$
$$x + 0 \approx x \qquad x \cdot 1 \approx x$$
$$x + x \approx 0$$
$$x \cdot (y + z) \approx (x \cdot y) + (x \cdot z)$$

$K = \{\mathbf{2}_{\mathrm{BR}}\}$ where $\mathbf{2}_{\mathrm{BR}}$ is the two-element Boolean ring $\langle \{0, 1\}, +, \cdot, 0, 1\rangle$

SIMPLES: just the 1- and 2-element Boolean rings.

$t(x, y, z) = (x + y) \cdot x + (1 + x + y) \cdot z$

$s(x, y, u, v) = (1 + x + y) \cdot u + (x + y) \cdot v.$

Of course Boolean rings are definitionally equivalent to Boolean algebras — but they are not equivalent in all respects, such as term rewriting, so we include them in our list of examples. We do not give Boolean rings the same language as rings because the 'minus' operation is the same as 'plus'.

3. $(x^p \approx x, px \approx 0)$-RINGS $\langle R, +, \cdot, -, 0, 1\rangle$, for p a prime number.

AXIOMS: to the following ring axioms add $x^p \approx x$ and $px \approx 0$:
$$x + (y + z) \approx (x + y) + z \qquad x \cdot (y \cdot z) \approx (x \cdot y) \cdot z$$
$$x + y \approx y + x$$
$$x + 0 \approx x \qquad x \cdot 1 \approx x \approx 1 \cdot x$$
$$x + (-x) \approx 0$$
$$x \cdot (y + z) \approx (x \cdot y) + (x \cdot z) \qquad (x + y) \cdot z \approx (x \cdot z) + (y \cdot z).$$

$K = \{\mathbf{Z}_p\}$ where $\mathbf{Z}_p$ is the ring of integers modulo p.

SIMPLES: just $\mathbf{Z}_p$ and the trivial ring.

$t(x, y, z) = (x - y)^{p-1} \cdot x + (1 - (x - y)^{p-1}) \cdot z$

$s(x, y, u, v) = (1 - (x - y)^{p-1}) \cdot u + (x - y)^{p-1} \cdot v.$

These rings were studied as a generalization of Boolean rings by McCoy & Montgomery in 1937; they proved that every nontrivial ring in this variety is isomorphic to a subdirect power of $\mathbf{Z}_p$.

4. $(x^m \approx x)$-RINGS $\langle R, +, \cdot, -, 0, 1\rangle$, for $m > 1$.

AXIOMS: the ring axioms plus $x^m \approx x$.

K is the set of finite fields which satisfy $x^m \approx x$. (There are, up to isomorphism, only finitely many such fields.)

SIMPLES: precisely the members of K plus the trivial ring.

$t(x,y,z) = (x-y)^{m-1} \cdot x + (1-(x-y)^{m-1}) \cdot z$

$s(x,y,u,v) = (1-(x-y)^{m-1}) \cdot u + (x-y)^{m-1} \cdot v.$

The variety of $x^m \approx x$ rings was proved to be generated by finite fields in the famous 1945 paper of Jacobson — for a more recent account see Herstein (1975), p. 367. In 1948 Arens & Kaplansky developed modified Boolean power structure theorems for such rings, and in 1966 Dauns & Hofmann gave the Boolean product structure theorems.

If we let K be any finite collection of finite fields then the variety $\mathrm{V}(K)$ generated by K will be a discriminator variety, and the simple members of the variety will be the finite fields embeddable in members of K. For the discriminator term one can use the form above (where m is such that $k-1|m-1$ for each k equal to the size of a member of K); and there is a finite set of equations defining $\mathrm{V}(K)$, but this take a bit more work to find.

5. n-VALUED POST ALGEBRAS $\langle \mathrm{A}, \vee, \wedge, ', 0, 1, \ldots, n-1 \rangle$.

AXIOMS: these are quite involved — we refer the reader to Rosenbloom (1942).

$K = \{\mathbf{P}_n\}$ where $\mathbf{P}_n = \langle \{0,1,\ldots,n-1\}, \vee, \wedge, ', 0, 1, \ldots, n-1 \rangle$ is the n-element n-valued Post algebra, which is a chain under join and meet with $0 < n-1 < n-2 < \cdots < 1$, and $1' = 2$, $2' = 3, \ldots, 0' = 1$.

SIMPLES: just $\mathbf{P}_n$ and the trivial algebra.

$t(x,y,z) = [g(x,y) \wedge x] \vee [g(g(x,y),1) \wedge z]$

$s(x,y,u,v) = [g(g(x,y),1) \wedge u] \vee [g(x,y) \wedge v].$

where $g(x,y) = \left[\bigwedge_{j=1}^{n-1} \left(\bigwedge_{k=1}^{n} x^{(k)} \vee y^{(k)}\right)^{(j)}\right]'$, with $x^{(k)}$ an abbreviation for k applications of the unary operation $'$.

Post introduced n-valued propositional logic in 1921, and Rosenbloom introduced axioms for n-valued Post algebras in 1942. [Rosenbloom used only the fundamental operations $\vee$ and $'$; to simplify the treatment we have expanded his list.] Rosenbloom showed that every finite n-valued Post algebra was isomorphic to a power of $\mathbf{P}_n$. In 1953 Foster studied primal algebras, in particular showing that every n-valued Post algebra is isomorphic to a Boolean power of $\mathbf{P}_n$.

6. *PD*, the **Pure Discriminator** variety of algebras $\langle \mathrm{A}, t \rangle$.

AXIOMS:
$t(x,x,y) \approx y$
$t(x,y,x) \approx x$
$t(x,y,y) \approx x$
$t(x,t(x,y,z),y) \approx y$
$t(u,v,t(x,y,z)) \approx t(u,v,t(t(u,v,x),t(u,v,y),t(u,v,z))).$

$K = \{\langle \omega, t\rangle\}$, where the operation t gives a discriminator term $t(x,y,z)$ on ω.

SIMPLES: all algebras $\langle A, t\rangle$ where $t(x,y,z)$ is a discriminator term on A.

$t(x,y,z)$ is from the fundamental operation

$s(x,y,u,v) = t(t(x,y,u),t(x,y,v),v)$.

The Pure Discriminator variety was introduced by McKenzie in the late 1970's and he showed it had a decidable first-order theory. This result was not published at the time; however it will soon appear in Burris, McKenzie, & Valeriote (1991+).

7. CA_n, the variety of **Cylindric Algebras of Dimension** n. Such algebras are of the form $\langle B, \vee, \wedge, ', 0, 1, c_0, \ldots, c_{n-1}, d_{00}, \ldots, d_{n-1n-1}\rangle$.

AXIOMS: to the axioms for Boolean algebras add the following (where $i,j,k < n$):

$x \leq y \to c_i(x) \leq c_i(y)$ $\qquad$ $d_{ii} \approx 1$

$c_i(c_j(x)) \approx c_j(c_i(x))$ $\qquad$ $c_i(x \wedge c_i(y)) \approx c_i(x) \wedge c_i(y)$

$d_{ik} \approx c_j(d_{ij} \wedge d_{jk})$ if $i \neq j \neq k$

$c_i(d_{ij} \wedge x) \wedge c_i(d_{ij} \wedge x') \approx 0$ if $i \neq j$.

K is the set of all algebras in CA_n such that $x \not\approx 0 \to c_0(c_1(\cdots c_{n-1}(x)\cdots)) \approx 1$ holds.

SIMPLES: precisely the members of K.

$t(x,y,z) = [c(x+y) \wedge x] \vee [c(x+y)' \wedge z]$

$s(x,y,u,v) = [c(x+y)' \wedge u] \vee [c(x+y) \wedge v]$,

where c is the composition $c_0 \circ \cdots \circ c_{n-1}$ and $x+y$ is $(x \wedge y') \vee (x' \wedge y)$.

The theory of cylindric algebras was founded by Tarski in collaboration with Chin and Thompson in the years 1948–52 as an algebraic version of first-order logic. This has been a very rich area of research, with two major textbooks by Henkin, Monk & Tarski (1975/85).

8. MA, the variety of **Monadic Algebras.** Such algebras are of the form $\langle B, \vee, \wedge, ', 0, 1, c\rangle$.

AXIOMS: to the axioms for Boolean algebras add the following axioms (which say that c is a closure operator and the closed elements form a subalgebra of the Boolean algebra):

$x \leq y \to c(x) \leq c(y)$

$c(c(x)) \approx c(x)$ $\qquad$ $c(c(x)') \approx c(x)'$

$c(c(x) \vee c(y)) \approx c(x) \vee c(y)$ $\qquad$ $c(c(x) \wedge c(y)) \approx c(x) \wedge c(y)$.

K is the set of all algebras in MA such that $x \not\approx 0 \to c(x) \approx 1$.

SIMPLES: precisely the members of K.

$t(x,y,z) = [c(x+y) \wedge x] \vee [c(x+y)' \wedge z]$

$s(x,y,u,v) = [c(x+y)' \wedge u] \vee [c(x+y) \wedge v]$.

where $x+y$ means $(x \wedge y') \vee (x' \wedge y)$.

Monadic algebras are essentially cylindric algebras of dimension 1 (the constant d_{00} has been dropped since it equals 1). They give the simplest well-known discriminator variety with infinite simple members. Such varieties are important for the reduction of mathematics to equational logic in section 6.

9. *RA*, the variety of **Relation Algebras** $\langle B, \vee, \wedge, ', 0, 1, \circ, ^{\cup}, \Delta\rangle$.

AXIOMS: take the axioms for Boolean algebras and add

$(x')^{\cup} \approx (x^{\cup})'$ $\quad$ $\Delta^{\cup} \approx \Delta$

$x^{\cup\cup} \approx x$ $\quad$ $x \circ \Delta \approx x \approx \Delta \circ x$

$x \circ (y \circ z) \approx (x \circ y) \circ z$ $\quad$ $(x \circ y)^{\cup} \approx y^{\cup} \circ x^{\cup}$

$(x \vee y)^{\cup} \approx x^{\cup} \vee y^{\cup}$ $\quad$ $(x^{\cup} \circ (x \circ y)') \wedge y \approx 0$

$x \circ (y \vee z) \approx (x \circ y) \vee (x \circ z)$.

K consists of all algebras in RA which satisfy $x \not\approx 0 \to 1 \circ x \circ 1 \approx 1$.

SIMPLES: precisely the class K.

$t(x, y, z) \approx [(1 \circ (x + y) \circ 1) \wedge x] \vee [(1 \circ (x + y) \circ 1)' \wedge z]$

$s(x, y, u, v) = [(1 \circ (x + y) \circ 1)' \wedge u] \vee [(1 \circ (x + y) \circ 1) \wedge v]$,

where $x + y$ means $(x \wedge y') \vee (x' \wedge y)$.

Relation algebras originated with DeMorgan and Peirce as a natural extension of the calculus of classes to the calculus of binary relations. Schröder devoted Volume III of his Algebra der Logik to this subject, and this in turn led to a deep study by Tarski. The modern study of relation algebras, including equational axiomatizations, is due to Tarski. A major new book on relation algebras, "A Formalization of Set Theory without Variables", has been published by Tarski & Givant (1987).

10. $V[\mathcal{F}]$, where V is a discriminator variety and $\mathcal{F}$ is a set of new function symbols, means the variety generated by the simple algebras of V, arbitrarily expanded by functions corresponding to the symbols in $\mathcal{F}$. We say that $V[\mathcal{F}]$ is the *compatible expansion of V by $\mathcal{F}$* because the new functions respect the Boolean product decompositions in V.

AXIOMS: to a set of axioms for V add, for each n-ary function symbol f in $\mathcal{F}$, the following *compatibility axiom*†

$$t(u, v, f(x_1, \ldots, x_n)) = t(u, v, f(t(u, v, x_1), \ldots, t(u, v, x_n))),$$

†The crucial compatibility axioms are motivated by an understanding of the fact that if you want to add functions to a discriminator variety without losing the fact that you have a discriminator variety, and you want to have the original discriminator term work in the new variety, then you need to keep the principal congruences the same. Now the principal congruences in a discriminator variety are equationally defined, namely

$$\Theta(a, b) = \{\langle c, d\rangle : t(a, b, c) = t(a, b, d)\}.$$

So the objective is, for each new n-ary function symbol f, to have

$$\langle c_i, d_i\rangle \in \Theta(a, b) \Longrightarrow \langle f(\vec{c}), f(\vec{d})\rangle \in \Theta(a, b),$$

that is,

$$t(a, b, c_i) = t(a, b, d_i) \Longrightarrow t(a, b, f(\vec{c})) = t(a, b, f(\vec{d})).$$

Thus we see the importance of the term $t(u, v, f(x_1, \ldots, x_n))$ in the condition for axiomatizing $V[\mathcal{F}]$. The precise compatibility condition of McKenzie is no doubt the result of some shrewd intuition.

where t is a discriminator term for the simple algebras in V.

K can be obtained by taking a generating class H for V, and taking all possible expansions of members of H by functions corresponding to symbols in $\mathcal{F}$.

$t(x, y, z)$ is a discriminator term for the simple algebras in V

$s(x, y, u, v)$ is a switching term for the simple algebras in V.

These axioms are an application of McKenzie's 1975 milestone study of axiom systems for discriminator varieties.† We shall return to his work in section 5 where the varieties $V[\mathcal{F}]$ will be an essential ingredient of the method used. For the interested reader we have given a completeness proof in the Appendix (based on McKenzie's paper, using our notation) from which one can also see that the above axioms indeed perform as required.

3. Discriminator Varieties have Unitary Unification.

In 1965 Robinson proved that the variety of all algebras (of a given type) has unitary unification. Unification is basic to resolution theorem provers and the Knuth-Bendix method for finding rewrite systems. For an excellent survey of the role of unification in computer science see Siekmann (1989). In 1987 Büttner & Simonis proved that Boolean algebras are unitary, and recently this has been extended by Nipkow (1990) to varieties generated by primal algebras. As we mentioned earlier such varieties are finitely generated discriminator varieties, and include examples 1, 2, and 5 from section 2.

Let V be a variety of algebras, and let $p(x_1, \ldots, x_k)$ and $q(x_1, \ldots, x_k)$ be terms in the language of V. A *V-unifier* of p and q is a substitution

$$x_i \leftarrow t_i(u_1, \ldots, u_l) \qquad (1 \leq i \leq k) \tag{5}$$

such that V satisfies

$$p(t_1, \ldots, t_k) \approx q(t_1, \ldots, t_k). \tag{6}$$

p and q are *V-unifiable* if they have a V-unifier; and determining when p and q have a V-unifier is called the *V-unification problem* for V.

A given V-unifier

$$x_i \leftarrow t_i(u_1, \ldots, u_l) \qquad (1 \leq i \leq k) \tag{7}$$

is *more general than* another V-unifier

$$x_i \leftarrow t'_i(v_1, \ldots, v_m) \qquad (1 \leq i \leq k) \tag{8}$$

if there is a substitution

$$u_j \leftarrow t''_j(v_1, \ldots, v_m) \qquad (1 \leq j \leq l) \tag{9}$$

such that V satisfies

$$t'_i(\vec{v}) \approx t_i(t''_1(\vec{v}), \ldots, t''_l(\vec{v})) \qquad (1 \leq i \leq k). \tag{10}$$

The notion of 'more general' establishes a pre-ordering (i.e., a reflexive and transitive relation) on the V-unifiers of p and q, and two V-unifiers are *equivalent* if each is more

†Other axiomatizations can be found in Bloom & Tindell (1983), Mekler & Nelson (1987).

general than the other. Sometimes p and q have, up to equivalence, minimal V-unifiers, called *most general* V-unifiers. If every pair of V-unifiable terms has a most general V-unifier which is more general than all V-unifiers of the pair, then we say the variety has *unitary unification.*

One can describe the property of V having unitary unification in a purely algebraic manner by switching to normal forms, which allows us to replace $\approx$ by $=$. Then in the above discussion of unification the items (5) and (6) become:

$$\bar{x}_i \leftarrow \overline{t_i(u_1, \ldots, u_l)} \qquad (1 \leq i \leq k) \tag{11}$$

$$\overline{p(t_1, \ldots, t_k)} = \overline{q(t_1, \ldots, t_k)}. \tag{12}$$

Now (11) corresponds to a homomorphism $\sigma : \mathbf{F}_V(x_1, \ldots, x_n) \to \mathbf{F}_V(u_1, u_2, \ldots)$ determined by

$$\sigma_1(\bar{x}_i) = \overline{t_i(u_1, \ldots, u_l)} \qquad (1 \leq i \leq k), \tag{13}$$

Without loss of generality we can assume the set of variables $\{u_1, \ldots\}$ is countably infinite. We can state unification in terms of normal forms by saying that $\bar{p}$ and $\bar{q}$ are V-unifiable if there is a homomorphism $\sigma : \mathbf{F}_V(x_1, \ldots, x_n) \to \mathbf{F}_V(u_1, u_2, \ldots)$ such that $\sigma(\bar{p}) = \sigma(\bar{q})$.

Thus, to tidy up the notation a little, unification is essentially concerned with considering a pair a, b of normal forms in $\mathrm{F}_V(n)$ and looking for homomorphisms $\sigma : \mathbf{F}_V(n) \to \mathbf{F}_V(\omega)$ such that $\sigma(a) = \sigma(b)$.

Continuing in this vein we can translate all the syntactic notions regarding unifiers mentioned above, namely for $a, b \in \mathrm{F}_V(n)$ define the set of **unifiers** of a and b to be

$$U_{n,V}(a, b) = \{\sigma \in \mathrm{Hom}(\mathbf{F}_V(n), \mathbf{F}_V(\omega)) : \sigma(a) = \sigma(b)\},$$

where $\mathrm{Hom}(\mathbf{A}, \mathbf{B})$ is the set of homomorphisms from $\mathbf{A}$ to $\mathbf{B}$. We say a and b are **V-unifiable** if $U_{n,V}(a, b) \neq \emptyset$. The problem of determining if any given $U_{n,V}(a, b)$ is empty is called the **unification problem for** V. Define a preorder $\leq$ (called **more general than**) on $U_{n,V}(a, b)$ by $\sigma_1 \leq \sigma_2$ iff there is a $\tau \in \mathrm{Hom}(\mathbf{F}_V(\omega), \mathbf{F}_V(\omega))$ such that $\sigma_2 = \tau \circ \sigma_1$. The homomorphism τ corresponds to the substitution (9), and the equality corresponds to the equations in (10). Two elements σ_1, σ_2 of $U_{n,V}(a, b)$ are **equivalent** if each is $\leq$ to the other. Minimal elements (up to equivalence) of $U_{n,V}(a, b)$ are called **most general unifiers** of a and b. The variety V is **unitary** if for each $\mathbf{F}_V(n)$ and each V-unifiable pair of elements a, b from $\mathbf{F}_V(n)$ there is a μ from $U_{n,V}(a, b)$ such that for every $\sigma \in U_{n,V}(a, b)$ we have $\mu \leq \sigma$.

Since this paper is addressed to an audience of computer scientists, perhaps it will be useful to try to explain what algebra offers in the study of unification. Of course the real justification lies in the results one can obtain. The merits of the algebraic approach can be roughly explained as follows: the syntactic approach to unification focuses on the local behavior of terms under substitution, whereas the algebraic approach (via free algebras) brings in global aspects of how the normal forms of terms relate to one another, i.e, the properties of the free algebras. At first glance this might not seem to be particularly relevant; however our experience has shown this to be a truly powerful tool in the classification of unification types. We will use the structure theorem for (free) algebras in discriminator varieties to prove that they have unitary unification type. Lawrence (1991a,b) has used the occurrence of the Hopf and Schreier properties in free algebras with success — for example he uses them to show that groups have infinitary

unification type. And Albert and Willard have made significant use of the algebraic approach as well. We will use both the syntactic and the algebraic formulations above.

Now we are ready to prove that indeed every discriminator variety is unitary. This offers fascinating possibilities, especially in view of the results of section 6 which show that discriminator varieties can be used to reduce all mathematics to equational logic. Theorem 3.1 can be viewed as an extension of Löwenheim's reproductive solutions of Boolean equations, and this in turn was clearly inspired by Schröder's (1890–1902), Vol. III, p. 161, beautiful work on the general solution of equations in the algebra of all binary relations on a set X.

In the following let $X = \{x_i : 1 \leq i < \omega\}$ and $\hat{X} = \{\hat{x}_i : 1 \leq i < \omega\}$ be two disjoint sets of variables. If p is a term in variables from X let $\hat{p}$ be the term in variables from $\hat{X}$ obtained by replacing each x_i in p by $\hat{x}_i$.

THEOREM 3.1. (DISCRIMINATOR VARIETIES HAVE UNITARY UNIFICATION). *Let V be a discriminator variety with a switching term $s(x,y,u,v)$ on the simple algebras in V. Let $p(x_1,\ldots,x_n)$, $q(x_1,\ldots,x_n)$ be two terms which are V-unifiable, and let $r_1,\ldots,r_n$ be terms in variables from X such that V satisfies $p(r_1,\ldots,r_n) \approx q(r_1,\ldots,r_n)$. Then the substitution*

$$x_i \leftarrow s(\hat{p},\hat{q},\hat{x}_i,\hat{r}_i) \qquad (1 \leq i \leq n),$$

is a V-unifier of p and q which is more general than any other V-unifier of p and q.

PROOF. Let $\mathbf{F}_V(n) = \mathbf{F}_V(x_1,\ldots,x_n)$, $\mathbf{F}_V(\omega) = \mathbf{F}_V(X)$, and let $\mu : \mathbf{F}_V(n) \to \mathbf{F}_V(\omega)$ be defined by

$$\mu(\bar{x}_i) = s(\bar{p},\bar{q},\bar{x}_i,\bar{r}_i).$$

Let φ be a maximal congruence of $\mathbf{F}_V(\omega)$. Then

$$\begin{aligned}
\mu(\bar{x}_i)/\varphi &= s(\bar{p},\bar{q},\bar{x}_i,\bar{r}_i)/\varphi \\
&= s(\bar{p}/\varphi,\bar{q}/\varphi,\bar{x}_i/\varphi,\bar{r}_i/\varphi) \\
&= \begin{cases} \bar{x}_i/\varphi & \text{if} \quad \bar{p}/\varphi = \bar{q}/\varphi \\ \bar{r}_i/\varphi & \text{if} \quad \bar{p}/\varphi \neq \bar{q}/\varphi \end{cases}
\end{aligned}$$

since $\mathbf{F}_V/\varphi$ is simple by Fact 1.1. Thus

$$\begin{aligned}
\mu(\bar{p})/\varphi &= p(\mu(\bar{x}_1),\ldots,\mu(\bar{x}_n))/\varphi \\
&= p(\mu(\bar{x}_1)/\varphi,\ldots,\mu(\bar{x}_n)/\varphi) \\
&= \begin{cases} p(\bar{x}_1/\varphi,\ldots,\bar{x}_n/\varphi) & \text{if} \quad \bar{p}/\varphi = \bar{q}/\varphi \\ p(\bar{r}_1/\varphi,\ldots,\bar{r}_n/\varphi) & \text{if} \quad \bar{p}/\varphi \neq \bar{q}/\varphi \end{cases} \\
&= \begin{cases} p(\bar{x}_1,\ldots,\bar{x}_n)/\varphi & \text{if} \quad \bar{p}/\varphi = \bar{q}/\varphi \\ p(\bar{r}_1,\ldots,\bar{r}_n)/\varphi & \text{if} \quad \bar{p}/\varphi \neq \bar{q}/\varphi \end{cases} \\
&= \begin{cases} \bar{p}/\varphi & \text{if} \quad \bar{p}/\varphi = \bar{q}/\varphi \\ p(\bar{r}_1,\ldots,\bar{r}_n)/\varphi & \text{if} \quad \bar{p}/\varphi \neq \bar{q}/\varphi. \end{cases}
\end{aligned}$$

Likewise we have

$$\mu(\bar{q})/\varphi = \begin{cases} \bar{q}/\varphi & \text{if} \quad \bar{p}/\varphi = \bar{q}/\varphi \\ p(\bar{r}_1,\ldots,\bar{r}_n)/\varphi & \text{if} \quad \bar{q}/\varphi \neq \bar{q}/\varphi. \end{cases}$$

Consequently $\mu(\bar{p})/\varphi = \mu(\bar{q})/\varphi$. By Facts 1.2, 1.3 we see that $\mu(\bar{p}) = \mu(\bar{q})$, so μ is a V-unifier of $\bar{p}$ and $\bar{q}$.

Next let $\sigma \in U_{n,V}(\bar{p}, \bar{q})$. Then

$$\begin{aligned} \sigma\mu(\bar{x}_i) &= \sigma(s(\bar{p}, \bar{q}, \bar{x}_i, \bar{r}_i)) \\ &= s(\sigma(\bar{p}), \sigma(\bar{q}), \sigma(\bar{x}_i), \sigma(\bar{r}_i)) \\ &= \sigma(\bar{x}_i) \end{aligned}$$

since $\sigma(\bar{p}) = \sigma(\bar{q})$. As $\sigma\mu$ and σ agree on the generators $\bar{x}_i$ of $\mathbf{F}_V(x_1, \ldots, x_n)$ it follows that $\sigma\mu = \sigma$, so μ is a V-unifier of $\bar{p}$ and $\bar{q}$ which is more general than any other. Consequently V has unitary unification. Now the substitution claimed in the statement of the theorem is simply a decoding of $\mu(\bar{x}_i) = s(\bar{p}, \bar{q}, \bar{x}_i, \bar{r}_i)$, plus changing the variables on the right-hand side to make them distinct from the variables on the left side to comply with the conventions of computer science. ∎

Now let us look at a few examples of unification in discriminator varieties. Obviously the potentially difficult aspect is trying to find an initial unifier, which we will represent as the sequence $\langle r_1, \ldots, r_n \rangle$. Note that an initial unifier need not have more than one variable involved; and if there are constants in the language then we can use ground terms for the r_i. The method we used to find initial unifiers in the examples below was to use trial and error on the collection S_1 of one-generated simple algebras in V [or the collection S_0 of zero-generated simples if there are constants present) — see Theorem 3.2(b).

Boolean Algebras. Find a most general unifier for $p = x \vee (y \wedge z)'$ and $q = (x \wedge y) \vee z$.

Solution: We can choose $\langle r_1, r_2, r_3 \rangle = \langle 1, 0, 1 \rangle$. Then a most general unifier of p and q is given by

$$\begin{aligned} x &\leftarrow s(\hat{p}, \hat{q}, \hat{x}, 1) \\ y &\leftarrow s(\hat{p}, \hat{q}, \hat{y}, 0) \\ z &\leftarrow s(\hat{p}, \hat{q}, \hat{z}, 1), \end{aligned}$$

and thus by

$$\begin{aligned} x &\leftarrow \hat{x} \vee \hat{y} \vee \hat{z}' \\ y &\leftarrow \hat{x} \wedge \hat{y} \\ z &\leftarrow \hat{x}' \vee \hat{y}' \vee \hat{z}. \end{aligned}$$

Boolean Rings. Find a most general unifier for $p = x + y + xy + yz + xyz$ and $q = xy + z + xyz$.

Solution: We can choose $\langle r_1, r_2, r_3 \rangle = \langle 0, 0, 0 \rangle$. Then a most general unifier of p and q is given by

$$x \leftarrow s(\hat{p}, \hat{q}, \hat{x}, 0)$$

$$\begin{aligned} y &\leftarrow s(\hat{p},\hat{q},\hat{y},0) \\ z &\leftarrow s(\hat{p},\hat{q},\hat{z},0), \end{aligned}$$

and thus by

$$\begin{aligned} x &\leftarrow \hat{x}\cdot\hat{y} \\ y &\leftarrow \hat{x}\cdot\hat{y}+\hat{x}\cdot\hat{z}+\hat{x}\cdot\hat{y}\cdot\hat{z} \\ z &\leftarrow \hat{x}+\hat{y}\cdot\hat{z}. \end{aligned}$$

$x^m \approx x$-Rings. Find a most general unifier for $p = x^2 + y^2$ and $q = 2x^2 + 1$.

Solution: We can choose $\langle r_1, r_2\rangle = \langle 0, 1\rangle$. Then a most general unifier of p and q is given by

$$\begin{aligned} x &\leftarrow s(\hat{p},\hat{q},\hat{x},0) \\ y &\leftarrow s(\hat{p},\hat{q},\hat{y},1), \end{aligned}$$

and thus by

$$\begin{aligned} x &\leftarrow (1-(1+\hat{x}^2-\hat{y}^2)^{m-1})\cdot\hat{x} \\ y &\leftarrow (1-(1+\hat{x}^2-\hat{y}^2)^{m-1})\cdot\hat{y}+(1+\hat{x}^2-\hat{y}^2). \end{aligned}$$

n-valued Post algebras. Find a most general unifier for $p = x \wedge y'$ and $q = (x \wedge y)'$.

Solution: We can choose $\langle r_1, r_2\rangle = \langle 1, 0\rangle$. Then a most general unifier of p and q is given by

$$\begin{aligned} x &\leftarrow s(\hat{p},\hat{q},\hat{x},1) \\ y &\leftarrow s(\hat{p},\hat{q},\hat{y},0). \end{aligned}$$

(This is complicated to expand out in terms of the basic operations.)

Monadic Algebras. Find a most general unifier for $p = x \vee c(y)$ and $q = c(x \wedge y)$.

Solution: We can choose $\langle r_1, r_2\rangle = \langle 0, 0\rangle$. Then a most general unifier of p and q is given by

$$\begin{aligned} x &\leftarrow s(\hat{p},\hat{q},\hat{x},0) \\ y &\leftarrow s(\hat{p},\hat{q},\hat{y},0), \end{aligned}$$

and thus by

$$\begin{aligned} x &\leftarrow \hat{x} \wedge c(\hat{x} \wedge \hat{y}) \\ y &\leftarrow \hat{y} \wedge c(\hat{x} \wedge \hat{y}). \end{aligned}$$

Relation Algebras. Find a most general unifier for $p = x^{\cup} \circ y$ and $q = x \wedge y'$.

Solution: We can choose $\langle r_1, r_2 \rangle = \langle 0, 0 \rangle$. Then a most general unifier of p and q is given by

$$\begin{aligned} x &\leftarrow s(\hat{p}, \hat{q}, \hat{x}, 0) \\ y &\leftarrow s(\hat{p}, \hat{q}, \hat{y}, 0). \end{aligned}$$

and thus by

$$\begin{aligned} x &\leftarrow [1 \circ (\hat{x}^{\cup} \circ \hat{y} + \hat{x} \wedge \hat{y}') \circ 1]' \wedge \hat{x} \\ y &\leftarrow [1 \circ (\hat{x}^{\cup} \circ \hat{y} + \hat{x} \wedge \hat{y}') \circ 1]' \wedge \hat{y}. \end{aligned}$$

$PD[\{+, \cdot\}]$ (the pure discriminator variety augmented by two compatible binary operations $+, \cdot$). Find a most general unifier for $p = x + (y \cdot y)$ and $q = (y + y) + z$.

Solution: We can choose $\langle r_1, r_2, r_3 \rangle = \langle y + y, y, y \cdot y \rangle$. Then a most general unifier of p and q is given by

$$\begin{aligned} x &\leftarrow s(\hat{p}, \hat{q}, \hat{x}, \hat{y} + \hat{y}) \\ y &\leftarrow s(\hat{p}, \hat{q}, \hat{y}, \hat{y}), \\ z &\leftarrow s(\hat{p}, \hat{q}, \hat{z}, \hat{y} \cdot \hat{y}). \end{aligned}$$

Now let us look at the unification problem for such varieties.

THEOREM 3.2. *Let V be a discriminator variety, and let S_n be the class of $(\leq n)$-generated simple algebras in V For any terms $p(x_1, \ldots, x_n)$ and $q(x_1, \ldots, x_n)$ in the language of V let $\bar{p}$ and $\bar{q}$ be the corresponding elements of $\mathbf{F}_V(n)$. Then we have*

(a) $U_{n,V}(\bar{p}, \bar{q}) \neq \emptyset \iff S_1 \models \exists x_1 \cdots \exists x_n [p(x_1, \ldots, x_n) \approx q(x_1, \ldots, x_n)]$
$\iff \mathbf{F}_V(1) \models \exists x_1 \cdots \exists x_n [p(x_1, \ldots, x_n) \approx q(x_1, \ldots, x_n)]$.

(b) $x_i \leftarrow r_i$, $1 \leq i \leq n$, *is a V-unifier of p and q iff* $S_1 \models p(r_1, \ldots, r_n) \approx q(r_1, \ldots, r_n)$.

If there are constants in the language of V then we also have

(a') $U_{n,V}(\bar{p},\bar{q}) \neq \emptyset \iff S_0 \models \exists x_1 \cdots \exists x_n [p(x_1,\ldots,x_n) \approx q(x_1,\ldots,x_n)]$
$\iff \mathbf{F}_V(0) \models \exists x_1 \cdots \exists x_n [p(x_1,\ldots,x_n) \approx q(x_1,\ldots,x_n)]$.

(b') $x_i \leftarrow r_i$, $1 \leq i \leq n$, *where the* r_i *are ground terms, is a* V*-unifier of* p *and* q *iff* $S_0 \models p(r_1,\ldots,r_n) \approx q(r_1,\ldots,r_n)$.

PROOF. (a) Suppose $U_{n,V}(\bar{p},\bar{q}) \neq \emptyset$. Choose $\sigma \in \mathrm{Hom}(\mathbf{F}_V(n), \mathbf{F}_V(\omega))$ such that $\sigma(\bar{p}) = \sigma(\bar{q})$. For $\mathbf{A} \in S_1$ choose any $\sigma' \in \mathrm{Hom}(\mathbf{F}_V(\omega), \mathbf{A})$. Then, with $\sigma'' = \sigma' \circ \sigma$, we have $\sigma''(\bar{p}) = \sigma''(\bar{q})$, so $p(\sigma''(x_1),\ldots,\sigma''(x_n)) \approx q(\sigma''(x_1),\ldots,\sigma''(x_n))$. Thus one can solve $p \approx q$ in any member of S_1.

Next suppose one can solve $p(x_1,\ldots,x_n) \approx q(x_1,\ldots,x_n)$ in any member of S_1. Then by Fact 1.6 we see that $p \approx q$ can be solved in $\mathbf{F}_V(1)$ as well. As $\mathbf{F}_V(1)$ embeds into $\mathbf{F}_V(\omega)$ (in any variety) it follows that $p \approx q$ can be solved in $\mathbf{F}_V(\omega)$, so $U_{n,V}(\bar{p},\bar{q}) \neq \emptyset$.

(b) This is just the argument of part (a) cast in syntactic form.

Essentially the same proof applies, when constants are present, with S_1 replaced by S_0, and $\mathbf{F}_V(1)$ replaced by the initial algebra $\mathbf{F}_V(0)$. ∎

COROLLARY 3.3. *Let* V *be discriminator variety such that the decision problem for satisfiability of equations in* S_1 *[*S_0 *if there are constants present] is solvable, i.e., there is an algorithm to determine whether or not any given equation can be solved in all members of* S_1 *[or* S_0*]. Then the unification problem for* V *is decidable.*

PROOF. By Theorem 3.2 two terms p and q are V-unifiable iff $p \approx q$ is solvable for each member of S_1 [or S_0]. ∎

As a special case of the above corollary we have:

COROLLARY 3.4. *Let* V *be discriminator variety such that* S_1 *[*S_0 *if there are constants present] has only finitely many algebras in it, all of which are finite, and the number of function symbols is finite. Then the unification problem for* V *is decidable. (This applies to Examples 1–9 from section 2.)*

Let us look at some examples.

Boolean Algebras. Terms p and q are unifiable iff the equation $p \approx q$ has a solution in the 2-element Boolean algebra.

Boolean Rings. Terms p and q are unifiable iff the equation $p \approx q$ has a solution in the 2-element Boolean ring.

$x^m \approx x$-Rings. Terms p and q are unifiable iff the equation $p \approx q$ has a solution in the prime fields $\mathbf{Z}_n$ for primes n such that $n-1 | m-1$.

n-valued Post algebras. Terms p and q are unifiable iff the equation $p \approx q$ has a solution in the Post algebra $\mathbf{P}_n$.

Monadic Algebras. Terms p and q are unifiable iff the equation $p \approx q$ has a solution in the 2-element monadic algebra.

Relation Algebras. Terms p and q are unifiable iff the equation $p \approx q$ has a solution in the relation algebras with at most 4 elements.

So far we have looked at unifying a single pair of terms $\langle p, q\rangle$. One can generalize the results of this section to the simultaneous unification of several pairs of terms. Let $\Pi = \{\langle p_1, q_1\rangle, \ldots, \langle p_k, q_k\rangle\}$ be k pairs of terms in the variables $x_1, \ldots, x_n$. A substitution is a *V-unifier* of Π if it is a V-unifier of each pair $\langle p_i, q_i\rangle$ in Π. Then we can speak of *more general* and *most general* V-unifiers of Π. We say that V has *unitary generalized unification* if each V-unifiable Π has a most general V-unifier which is more general than all V-unifiers of Π.

With the help of the next lemma we can show discriminator varieties have unitary generalized unification. As we have seen, the first powerful property of discriminator varieties is the "Boolean product of simples" structure theorem which allows us to switch the study of unification to the simple algebras — indeed this is how we discovered the most general unifier. The next remarkable property of discriminator varieties is the fact that when working with nontrivial simple algebras we can replace quantifier-free formulas by equations. This is a big part of the secret in section 5 of reducing first-order logic to equational logic via discriminator varieties.

LEMMA 3.5. *Let V be a discriminator variety, let V_S be the class of simple algebras in V, and let V_S^+ be the class of nontrivial (i.e., the universe has more than one element) simple algebras in V. Then we have the following reductions on quantifier-free formulas:*

$$V_S \models x \approx y \vee u \approx v \leftrightarrow s(x,y,u,v) \approx u \tag{14}$$
$$V_S \models x \not\approx y \wedge u \not\approx v \leftrightarrow s(x,y,u,v) \not\approx u \tag{15}$$
$$V_S \models x \approx y \wedge u \approx v \leftrightarrow t(x,y,u) \approx t(y,x,v) \tag{16}$$
$$V_S \models x \not\approx y \vee u \not\approx v \leftrightarrow t(x,y,u) \not\approx t(y,x,v) \tag{17}$$
$$V_S \models x \approx y \wedge u \not\approx v \leftrightarrow s(x,y,u,v) \not\approx v \tag{18}$$
$$V_S \models x \not\approx y \vee u \approx v \leftrightarrow s(x,y,u,v) \approx v \tag{19}$$
$$V_S^+ \models x \not\approx y \leftrightarrow \forall u \forall v\, s(x,y,u,v) \approx v. \tag{20}$$

PROOF. Let us verify the first reduction (14). Suppose $\mathbf{S} \in V_S$. Then we ask for the precise conditions on a, b, c, d which ensure

$$s(a,b,c,d) = c \tag{21}$$

holds in $\mathbf{S}$. From the definition (2) of the switching function, and since s is a switching function on the members of V_S, we see that if $a = b$ then (21) becomes $c = c$, and if $a \neq b$ then (21) becomes $d = c$. We can sum this up by saying that $s(a,b,c,d) = c$ holds in $\mathbf{S}$ iff $a = b$ or $c = d$.

One can give a similar argument for all of the reductions except the last, (20), which we look at now We want to know, given $\mathbf{S} \in V_S^+$, precisely when $\forall u \forall v\, s(a,b,u,v) \approx v$ holds in $\mathbf{S}$. If $a = b$ then, using (2), this statement becomes $\forall u \forall v\, (u \approx v)$, i.e., $\mathbf{S}$ is a trivial algebra. But since algebras in V_S^+ are nontrivial, this possibility cannot occur. Next we consider the case that $a \neq b$. Then the statement becomes $\forall u \forall v\, (v \approx v)$, which

is always true. We can sum this up by saying that $\forall u \forall v\, s(a,b,u,v) \approx v$ holds in $\mathbf{S}$ iff $a \neq b$. ∎

THEOREM 3.6. *Let V be a discriminator variety. Then V has unitary generalized unification.*

PROOF. Let $\Pi = \{\langle p_1, q_1\rangle, \ldots, \langle p_k, q_k\rangle\}$ be given. Then, making repeated use of (16) in Lemma 3.5, we see that the open formula $p_1 \approx q_1 \wedge \cdots \wedge p_k \approx q_k$ is equivalent on V_S^+ to an open formula of the form $p \approx q$. Then using Fact 1.5 we see that Π and the pair $\langle p, q\rangle$ have the same V-unifiers. Thus, by Theorem 3.1, V has unitary generalized unification. ∎

The *generalized unification problem* for a variety V, formulated in algebraic terms, is to determine whether or not any given finite set of (quantifier-free) equations $p_i \approx q_i$ $(1 \leq i \leq n)$ can be simultaneously solved in $\mathbf{F}_V(\omega)$.

COROLLARY 3.7. *Let V be discriminator variety such that the decision problem for satisfiability of equations in S_1 [S_0 if there are constants present] is solvable, i.e., there is an algorithm to determine whether or not any given equation can be solved in all members of S_1 [or S_0]. Then the generalized unification problem for V is decidable.*

PROOF. The conversion of a finite set of equations to a single equation given in the proof of Theorem 3.6 is effective, so we can use Corollary 3.3. ∎

And from this we have the following specialization.

COROLLARY 3.8. *Let V be discriminator variety such that S_1 [S_0 if there are constants present] has only finitely many algebras in it, all of which are finite, and the number of function symbols is finite. Then the generalized unification problem for V is decidable. (This applies to Examples 1–9 from section 2.)*

A variety V is *trivial* if it satisfies $\forall x \forall y (x \approx y)$, i.e., the only algebras in V are trivial.

THEOREM 3.9. *Let V be a nontrivial discriminator variety satisfying the hypotheses of Corollary 3.8 and such that there are two ground terms a, b which are distinct in every nontrivial algebra in V. Then the generalized unification problem for V is NP-complete. (This applies to Examples 1–5 and 7–9 from section 2.)*

PROOF. It is well-known that 3-SAT, the problem of determining the satisfiability of a finite number of propositional clauses, each clause involving exactly three distinct literals, is NP-complete (see Garey & Johnson (1979)). We will show how to convert, in polynomial time, such a set $\mathcal{C}$ of propositional clauses into a set of equations Σ such that $\mathcal{C}$ is satisfiable iff Σ is V-unifiable.

Suppose $\mathcal{C}$ involves the propositional variables $P_1, \ldots, P_n$. Then replace each P_i by the formula $x_i \approx y_i$ to obtain a set of quantifier-free clauses $\mathcal{C}^*$, where each clause is in one of the four forms:

$$(x_i \approx y_i) \vee (x_j \approx y_j) \vee (x_k \approx y_k) \tag{22}$$

$$\neg(x_i \approx y_i) \vee (x_j \approx y_j) \vee (x_k \approx y_k) \tag{23}$$

$$\neg(x_i \approx y_i) \vee \neg(x_j \approx y_j) \vee (x_k \approx y_k) \tag{24}$$

$$\neg(x_i \approx y_i) \vee \neg(x_j \approx y_j) \vee \neg(x_k \approx y_k). \tag{25}$$

We claim that $\mathcal{C}$ is satisfiable in the propositional calculus iff $\mathcal{C}^\star$ is satisfiable in V_{S}^+. For take an assignment of truth values that satisfies $\mathcal{C}$ and let $\mathbf{S} \in V_{\mathrm{S}}^+$. If a given P_i is assigned true then use the assignment

$x_i \leftarrow a$

$y_i \leftarrow a$;

else use the assignment

$x_i \leftarrow a$

$y_i \leftarrow b$.

This assignment for $1 \leq i \leq n$ will clearly make $\mathcal{C}^\star$ true in $\mathbf{S}$ since the ground terms a, b are distinct in $\mathbf{S}$.

Conversely suppose $\mathcal{C}^\star$ is satisfiable in V_{S}^+. Then take a particular assignment of the variables to values in some $\mathbf{S} \in V_{\mathrm{S}}^+$ so that $\mathcal{C}^\star$ holds, and then assign P_i the value "true" iff x_i and y_i are assigned the same value in $\mathbf{S}$. This shows $\mathcal{C}$ is satisfiable.

Before continuing note that since we have the two terms a and b which are distinct in V_{S}^+ we can replace (20) in Lemma 3.5 by

$$V_{\mathrm{S}}^+ \models x \not\approx y \quad \leftrightarrow \quad s(x, y, a, b) \approx b. \tag{26}$$

Now we apply the reductions of Lemma 3.5, replacing (20) by (26), to the four possible clauses (22)–(25) above to obtain the following possibilities:

$$s(s(x_i, y_i, x_j, y_j), x_j, x_k, y_k) \approx x_k \tag{27}$$

$$s(s(x_i, y_i, x_j, y_j), y_j, x_k, y_k) \approx x_k \tag{28}$$

$$s(t(x_i, y_i, x_j), t(y_i, x_i, y_j), x_k, y_k) \approx y_k \tag{29}$$

$$s(t(t(x_i, y_i, x_j), t(y_i, x_i, y_j), x_k), t(t(y_i, x_i, y_j), t(x_i, y_i, x_j), y_k), a, b) \approx b. \tag{30}$$

Let the resulting set of (quantifier-free) equations be $\mathcal{C}^{\star\star}$, and observe from the above set of equations (27)–(30) that indeed there is a linear bound on the size of $\mathcal{C}^{\star\star}$ in terms of the size of $\mathcal{C}$.

From Lemma 3.5 we see that $\mathcal{C}^{\star\star}$ is satisfiable in V_{S}^+ iff $\mathcal{C}^\star$ is satisfiable in V_{S}^+, and hence iff $\mathcal{C}$ is satisfiable in the propositional calculus. Now $\mathcal{C}^{\star\star}$ is satisfiable in V_{S}^+ iff it is satisfiable in every member of V by Fact 1.4 and by noting that it is satisfiable in a trivial algebra. Thus $\mathcal{C}$ is satisfiable implies $\mathcal{C}^{\star\star}$ is satisfiable in $\mathbf{F}_V(\omega)$, and hence it is V-unifiable. For the converse, if $\mathcal{C}^{\star\star}$ is V-unifiable then it is satisfiable in $\mathbf{F}_V(\omega)$, and hence in some member of V_{S}^+ by Facts 1.4 and 1.5. But we have already argued that this happens iff $\mathcal{C}$ is satisfiable. Thus $\mathcal{C}$ is satisfiable in the propositional calculus iff $\mathcal{C}^{\star\star}$ is V-unifiable.

From the NP-completeness of 3-SAT we now see that the generalized unification problem for V must be NP-hard. However the generalized unification problem for V is in NP since, given any finite set of quantifier-free equations, we can verify in polynomial time whether a particular assignment of values in a member of S_1 [or S_0] satisfies the equations; and there are only finitely many algebras in S_1 [or S_0] by assumption. Thus the generalized unification problem for V is actually NP-complete. ∎

REMARK. We originally thought we could prove the generalized unification problem to be NP-complete under the hypotheses of Corollary 3.8, that is, without the requirement that there be two ground terms a, b which are distinct on the nontrivial algebras in V. The referee found our argument dubious, and indeed it seems necessary to add some

extra condition such as that on the ground terms — but we have no proof that extra conditions are necessary.

REMARK. We do not know if the unification problem (for single equations) is NP-complete under the hypotheses of Theorem 3.9 — it is clearly in NP. If we try to use the same argument and collapse the system of equations $\mathcal{C}^{**}$ to a single equation using Lemma 3.5 the size seems to explode in general.

4. A Review of Equational Logic

Systems of equations $p_i(\vec{x}) \approx q_i(\vec{x})$ are commonly used with two distinct meanings: the first is the system of universally quantified sentences $\forall \vec{x}\, p_i(\vec{x}) \approx q_i(\vec{x})$, which is the meaning in this section; and the second is the quantifier-free formula $\bigwedge_i p_i(\vec{x}) \approx q_i(\vec{x})$ which we used in the algebraic discussion of unification when we spoke of solving a system of equations. Hopefully the context makes it clear which meaning is intended.

Garrett Birkhoff introduced equational logic in 1935. He showed that five easy rules of inference which everyone learns in high school give a complete set of rules for deriving equations. We would like to suggest that equational logic is an appropriate place to focus attention because of its simplicity, its well-developed model theory (universal algebra), the wealth of decidability results that are known, and the expressiveness of equations.

BIRKHOFF'S RULES OF INFERENCE FOR EQUATIONAL LOGIC

RULE	NAME	EXAMPLE
$\dfrac{}{p \approx p}$	Reflexive	$\dfrac{}{x^2 \approx x^2}$
$\dfrac{p \approx q}{q \approx p}$	Symmetric	$\dfrac{x \approx x^2}{x^2 \approx x}$
$\dfrac{p \approx q, q \approx r}{p \approx r}$	Transitive	$\dfrac{x \approx x^2, x^2 \approx x + x}{x \approx x + x}$
$\dfrac{p(x_1, \ldots, x_n) \approx q(x_1, \ldots, x_n)}{p(t_1, \ldots, t_n) \approx q(t_1, \ldots, t_n)}$	Substitution	$\dfrac{xy \approx yx}{(x+y)z \approx z(x+y)}$
$\dfrac{p \approx q}{r(\cdots p \cdots) \approx r(\cdots q \cdots)}$	Replacement	$\dfrac{xy \approx yx}{xy + xy \approx xy + yx}$

A **proof** of an equation E from a set of equations Σ is a finite sequence of equations $\mathrm{E}_1, \ldots, \mathrm{E}_n$ such that each E_i

1. is an **axiom** from Σ, or
2. is the result of **applying one of the rules of inference** to some members of $\mathrm{E}_1, \ldots, \mathrm{E}_{i-1}$; and
3. E_n is the equation E.

Birkhoff's rules can be used to prove that one can view a set of equations Σ as a set of **two way rewrite rules** in order to obtain all equational consequences of Σ. Thus to show that $p \approx q$ follows from Σ one can start with p and apply rewrite rules:

$$p \to p_1 \to \cdots \to p_n \to q.$$

From the computational point of view it is valuable to take the symmetry into consideration and start from both p and q.

For a simple example let Σ be the three popular equations defining groups:

(1) $x \cdot (y \cdot z) \approx (x \cdot y) \cdot z$
(2) $x \cdot 1 \approx x$
(3) $x \cdot x^{-1} \approx 1$

The following is a proof of $1 \cdot x \approx x$, using the three axioms as two-way rewrite rules:

1	$1 \cdot x$	$\rightarrow$	$1 \cdot (x \cdot 1)$	axiom 2
2		$\rightarrow$	$1 \cdot (x \cdot (x^{-1} \cdot (x^{-1})^{-1}))$	axiom 3
3		$\rightarrow$	$1 \cdot ((x \cdot x^{-1}) \cdot (x^{-1})^{-1})$	axiom 1
4		$\rightarrow$	$1 \cdot (1 \cdot (x^{-1})^{-1})$	axiom 3
5		$\rightarrow$	$(1 \cdot 1) \cdot (x^{-1})^{-1}$	axiom 1
6		$\rightarrow$	$1 \cdot (x^{-1})^{-1}$	axiom 2
7		$\rightarrow$	$(x \cdot x^{-1}) \cdot (x^{-1})^{-1}$	axiom3
8		$\rightarrow$	$x \cdot (x^{-1} \cdot (x^{-1})^{-1})$	axiom 1
9		$\rightarrow$	$x \cdot 1$	axiom 3
10		$\rightarrow$	x	axiom 2

The KNUTH-BENDIX Completion Procedure attempts to convert a set of equations into (one-way) rewrite rules which can be used to find normal forms for terms. One of the first and best known applications was to groups, resulting in the following 10 rewrite rules:

$1 \cdot x$	$\rightarrow$	x	1^{-1}	$\rightarrow$	1
$x^{-1} \cdot x$	$\rightarrow$	1	$(x^{-1})^{-1}$	$\rightarrow$	x
$(x \cdot y) \cdot z$	$\rightarrow$	$x \cdot (y \cdot z)$	$x^{-1} \cdot (x \cdot y)$	$\rightarrow$	y
$x \cdot 1$	$\rightarrow$	x	$x \cdot (x^{-1} \cdot y)$	$\rightarrow$	y
$x \cdot x^{-1}$	$\rightarrow$	1	$(x \cdot y)^{-1}$	$\rightarrow$	$y^{-1} \cdot x^{-1}$

This brief glance at equational logic will suffice for our immediate purposes. However for a more accurate picture of the breadth of the subject see McNulty (1989) and Taylor (1979); and for the Knuth-Bendix algorithm see Dershowitz (1989).

5. Reducing First-Order Logic to Equational Logic.

Much recent work in computer science has been focused on reducing first-order logic to two-sorted rewrite systems and equational logic, the first sort referring to the original first-order language, and the second sort to Boolean connectives (see the comments at the end of section 6). In this section we show how one can reduce first-order logic with equality to the traditional one-sorted equational logic. At this point we do not know that this will offer computational advantages, but the possibility is too good to ignore given the simplicity of the reduction.

We start by considering a first-order language $\mathcal{L}$. One of the fundamental achievements of Gödel was to show that the semantic notion $\Sigma \models \sigma$ could be captured by a syntactic notion $\Sigma \vdash \sigma$. In computer science, however, there has been a strong preference for simpler syntactic systems, in particular systems that avoid having to manipulate quantifiers. The usual procedure is to observe that $\Sigma \models \sigma$ holds iff $\Sigma \cup \{\neg\sigma\}$ is not

satisfiable. For τ a sentence let τ^* denote its Skolemized form. Then $\Sigma \cup \{\neg\sigma\}$ is not satisfiable iff $\Sigma^* \cup \{(\neg\sigma)^*\}$ is not satisfiable, where Σ^* is the set of Skolemized forms of sentences in Σ (of course one chooses new Skolem functions for every sentence to be Skolemized). This reduces the syntactic level to universally quantified sentences. Such sentences are easily expressed as conjunctions of clauses (i.e., universally quantified disjunctions of atomic and/or negated atomic formulas), so we have $\Sigma \models \sigma$ iff a certain set of clauses is not satisfiable. Robinson's resolution rule is complete for unsatisfiable sets of clauses (i.e., one can always derive the empty clause), provided one does not have equality in the language. If equality is present then other rules, such as paramodulation, must be introduced.

We will also use the reduction to clauses. But after this the approach via discriminator varieties is radically different from other approaches in the literature. Let us examine this crucial difference in more detail.

LEMMA 5.1. *Let V be a discriminator variety. Then for V_S^+ we see that every quantifier free formula is equivalent to a formula of the form $p \approx q$ or of the form $\forall u \forall v\, p \approx q$. Consequently universally quantified statements, in particular clauses, are equivalent to (universally quantified) equations in V_S^+.*

PROOF. Given a quantifier-free formula, repeatedly apply any of the reductions in Lemma 3.5 except for the last one until one has obtained an equivalent quantifier-free formula of the form $p' \approx q'$ or $p' \not\approx q'$. In the former case we have a quantifier free reduced form; in the latter case apply the last reduction of Lemma 3.5 to obtain $\forall u \forall v\, s(p', q', u, v) \approx v$. ■

Now we give the detailed steps (which are a slight variation of the analysis of axioms for discriminator varieties due to McKenzie (1975)) to derive a set of equations which can be used to analyze $\Sigma \models \sigma$ when we are working with a first-order language *with* equality. [We remark that for this and the next section, finitely generated discriminator varieties, such as examples 1–5, are of lesser importance. The major applications depend on having a discriminator variety with *infinite* simple members, such as examples 6–10].

The reader should keep in mind that we are motivated by trying to carry out the reduction so it works on the simple algebras in a discriminator variety (STEPS 1–6); then we want to make sure that the equations so obtained do not spoil the original discriminator variety — this is achieved by the compatibility equations (STEP 7).

THE MCKENZIE REDUCTION TO EQUATIONS

GIVEN: A set Σ of first-order sentences, and σ, a first order sentence.

STEP 1: Choose a discriminator variety V such that the generalized spectrum† of Σ is a subset of the generalized spectrum of V. Choose a discriminator term $t(x, y, z)$ and a switching term $s(x, y, u, v)$ for the simple algebras in V.

STEP 2: Put all sentences in $\Sigma \cup \{\neg\sigma\}$ in prenex form.

†The spectrum of a set of sentences, or of a class defined by such, is the set of sizes of the finite models. The generalized spectrum includes the sizes of the infinite models as well.

STEP 3: Skolemize the set of sentences from STEP 2.

STEP 3′: (optional) Break the sentences from STEP 3 into clauses.

STEP 4: Replace all atomic subformulas in the universal sentences resulting from STEP 3 (or STEP 3′) of the form $r(t_1, \ldots, t_n)$, r a relation symbol, by $f_r(t_1, \ldots, t_n) \approx t_1$. [$f_r$ is a new function symbol corresponding to r — this approach to encoding relations as functions can be found in Ackermann (1954), page 98.]

STEP 5: Use the following reductions (along with the commutativity of the connectives $\wedge$ and $\vee$) to replace each of the universal sentences obtained in STEP 4 by a sentence of the form $\forall x_1 \cdots \forall x_n (p \approx q)$, or of the form $\forall x_1 \cdots \forall x_n (p \not\approx q)$:

$$
\begin{array}{lcl}
x \approx y \vee u \approx v & \longrightarrow & s(x,y,u,v) \approx u \\
x \not\approx y \wedge u \not\approx v & \longrightarrow & s(x,y,u,v) \not\approx u \\
x \approx y \wedge u \approx v & \longrightarrow & t(x,y,u) \approx t(y,x,v) \\
x \not\approx y \vee u \not\approx v & \longrightarrow & t(x,y,u) \not\approx t(y,x,v) \\
x \approx y \wedge u \not\approx v & \longrightarrow & s(x,y,u,v) \not\approx v \\
x \not\approx y \vee u \approx v & \longrightarrow & s(x,y,u,v) \approx v.
\end{array}
$$

STEP 6. Replace each of the sentences resulting from STEP 5 which are of the form $\forall x_1 \cdots \forall x_n (p \not\approx q)$ by $\forall u \forall v \forall x_1 \cdots \forall x_n \, s(p,q,u,v) \approx v$.

Let $\mathrm{Red}_V(\Sigma, \sigma)$ be the set of equations resulting from steps 1–6 applied to Σ and σ.

STEP 7. Let $\mathcal{F}$ be the set of function symbols occurring in $\mathrm{Red}_V(\Sigma, \sigma)$ which do not already appear in V, and let $\mathrm{Ax}(V[\mathcal{F}])$ be a set of equational axioms $\mathrm{Ax}(V)$ for V plus the following *compatibility axiom* for each n-ary function symbol f in $\mathcal{F}$:

$$t(u,v,f(x_1,\ldots,x_n)) \approx t(u,v,f(t(u,v,x_1),\ldots,t(u,v,x_n))).$$

Now we are ready for the key result of this section.

THEOREM 5.2. $\Sigma \models \sigma$ *iff*

(a) *one-element models of* Σ *are models of* σ, *and*

(b) $Ax(V[\mathcal{F}]) \cup Red_V(\Sigma, \sigma) \models \forall x \forall y (x \approx y)$.

If Σ *has no one-element models, then we can drop condition (a).*

PROOF. The technical details are in the Appendix; we will make a few comments here. We have simply taken a standard reduction (STEPS 1–4) of $\Sigma \models \sigma$ to a set of unsatisfiable universal sentences and meshed this with McKenzie's analysis of satisfiability in the simple algebras of a discriminator variety. Steps 5 and 6 are rather straightforward (and were also used by Burris & Werner (1979) in the analysis of existentially closed members of discriminator varieties). The striking insight of McKenzie is STEP 7 which ensures that we still have a discriminator variety, the simple algebras in $\mathrm{Red}_V(\Sigma, \sigma) \cup \mathrm{Ax}(V[\mathcal{F}])$ are just expansions of the simple algebras from V, and t is still a discriminator term for the simple algebras. Then, for the nontrivial simple algebras, $\mathrm{Red}_V(\Sigma, \sigma)$ is an encoding of $\Sigma \cup \{\neg\sigma\}$. Thus, in view of STEP 1, the set of statements $\Sigma \cup \{\neg\sigma\}$ has a model of size κ iff $\mathrm{Ax}(V[\mathcal{F}]) \cup \mathrm{Red}_V(\Sigma, \sigma)$ has a simple model of size κ, for $\kappa > 1$. ■

A simple example will be used to illustrate the basic steps. Let Σ be the following set of first-order sentences:

$$a \not\approx b$$
$$\forall x\,(x \approx a \vee x \approx b)$$
$$\forall x \forall y \forall z\,([r(x,y) \wedge r(x,z)] \to y \approx z)$$
$$\forall x \forall y \forall z\,([r(x,z) \wedge r(y,z)] \to x \approx y)$$
$$\forall x \exists y\, r(x,y),$$

and let σ be the sentence

$$\forall y \exists x\, r(x,y)$$

STEP 1.
From the first two sentences we see that the only possible models of Σ are two-element models, so we could choose any of the discriminator varieties from our examples which has a two-element algebra in it. One can *always* use the Pure Discriminator variety when carrying out the McKenzie reduction, so let us use it for this example.

STEP 2.
The sentences of Σ are in prenex form, and $\neg\sigma$ in prenex form is $\exists y \forall x\, \neg r(x,y)$.

STEP 3.
Skolemizing the sentences from STEP 2 (and then *omit writing the universal quantifiers*, as is customary) gives:

$$a \not\approx b$$
$$x \approx a \vee x \approx b$$
$$[r(x,y) \wedge r(x,z)] \to y \approx z$$
$$[r(x,z) \wedge r(y,z)] \to x \approx y$$
$$r(x, g(x))$$
$$\neg r(x,c).$$

STEP 4.
We replace the binary relation r by a binary function f to eliminate the relation symbols:

$$a \not\approx b$$
$$x \approx a \vee x \approx b$$
$$[f(x,y) \approx x \wedge f(x,z) \approx x] \to y \approx z$$
$$[f(x,z) \approx x \wedge f(y,z) \approx y] \to x \approx y$$
$$f(x, g(x)) \approx x$$
$$f(x,c) \not\approx x.$$

STEP 5.
Now we eliminate the binary propositional connectives:

$$a \not\approx b$$
$$s(x,a,x,b) \approx x$$
$$s(t(f(x,y),x,f(x,z)), t(x,f(x,y),x), y, z) \approx z$$
$$s(t(f(x,z),x,f(y,z)), t(x,f(x,z),y), x, y) \approx y$$
$$f(x, g(x)) \approx x$$
$$f(x,c) \not\approx x.$$

STEP 6.
This step eliminates negation and gives $\mathrm{Red}_V(\Sigma, \sigma)$:

$$
\begin{aligned}
&s(a,b,u,v) \approx v\\
&s(x,a,x,b) \approx x\\
&s(t(f(x,y),x,f(x,z)),t(x,f(x,y),x),y,z) \approx z\\
&s(t(f(x,z),x,f(y,z)),t(x,f(x,z),y),x,y) \approx y\\
&f(x,g(x)) \approx x\\
&s(f(x,c),x,u,v) \approx v.
\end{aligned}
$$

STEP 7.
$\mathcal{F} = \{f, g\}$, so our final set of equations is $\mathrm{Ax}(V[\mathcal{F}]) \cup \mathrm{Red}_V(\Sigma, \sigma)$:

$\mathrm{Ax}(V[\mathcal{F}])$ consists of:

$$
\begin{aligned}
t(x,x,y) &\approx y && (31)\\
t(x,y,x) &\approx x && (32)\\
t(x,y,y) &\approx x && (33)\\
t(x,t(x,y,z),y) &\approx y && (34)\\
t(u,v,t(x,y,z)) &\approx t(u,v,t(t(u,v,x),t(u,v,y),t(u,v,z))) && (35)
\end{aligned}
$$

(compatibility axioms for f,g:)

$$
\begin{aligned}
t(u,v,f(x,y)) &\approx t(u,v,f(t(u,v,x),t(u,v,y))) && (36)\\
t(u,v,g(x)) &\approx t(u,v,g(t(u,v,x))) && (37)
\end{aligned}
$$

$\mathrm{Red}_V(\Sigma, \sigma)$ consists of:

$$
\begin{aligned}
s(a,b,u,v) &\approx v && (38)\\
s(x,a,x,b) &\approx x && (39)\\
s(t(f(x,y),x,f(x,z)),t(x,f(x,y),x),y,z) &\approx z && (40)\\
s(t(f(x,z),x,f(y,z)),t(x,f(x,z),y),x,y) &\approx y && (41)\\
f(x,g(x)) &\approx x && (42)\\
s(f(x,c),x,u,v) &\approx v. && (43)
\end{aligned}
$$

Σ has no 1-element models (as $a \not\approx b$), so the claim is now that $\Sigma \vdash \sigma$ holds iff the equations 31–43 lead to the equation $x \approx y$, which is indeed the case. We do not recommend that you try this by hand unless you have a good understanding of how the encoding into discriminator varieties works. However we will give a trivial example that can easily be done by hand, and which draws on the above steps.

Let Σ be empty, and let σ be $\forall x\,(r(x) \vee \neg r(x))$. Then, again using the Pure Discriminator variety, and carrying through the steps we arrive at equations (31)–(35), the compatibility equation (37), and the encoding of $\neg\sigma$:

$$
t(x,x,y) \approx y \qquad (44)
$$

$$t(x,y,x) \approx x \tag{45}$$
$$t(x,y,y) \approx x \tag{46}$$
$$t(x,t(x,y,z),y) \approx y \tag{47}$$
$$t(u,v,t(x,y,z)) \approx t(u,v,t(t(u,v,x),t(u,v,y),t(u,v,z))) \tag{48}$$
$$t(u,v,g(x)) \approx t(u,v,g(t(u,v,x))) \tag{49}$$
$$s(s(g(c),c,g(c),c),c,u,v) \approx v. \tag{50}$$

To show the original σ is valid we need to show that it holds on one-element structures, which it does, and that the equation $x \approx y$ follows from the above seven equations, which we now demonstrate:

$$\begin{aligned} s(g(c),c,g(c),c) &\approx t(t(g(c),c,g(c)),t(g(c),c,c),c) && \text{by (3)} \\ &\approx t(g(c),g(c),c) && \text{by (45) and (46)} \\ &\approx c && \text{by (44).} \end{aligned}$$

Thus from (50) and the previous result we have

$$s(c,c,u,v) \approx v. \tag{51}$$

Now

$$\begin{aligned} s(c,c,u,v) &\approx t(t(c,c,u),t(c,c,v),v) && \text{by (3)} \\ &\approx t(u,v,v) && \text{by (44)} \\ &\approx u && \text{by (46).} \end{aligned}$$

Combining this with (51) gives $u \approx v$, or equivalently, the desired $x \approx y$ since we are dealing with universally quantified equations. ∎

Returning to our general investigations, if the terms $t(x,y,z)$ or $s(x,y,u,v)$ have repeat occurrences of variables then there is a likelihood of an explosion in the size of the reduction of a quantifier-free formula to an equation. One can avoid this by simply adding a new ternary function symbol t and a new 4-ary function symbol s to the language, and adding two equations which say they are equal to the appropriate terms.

Next we look at some further reductions one can carry out — they are likely only of theoretical significance. If Σ is finite and $\mathrm{Ax}(V)$ is finite then one could make reductions as follows:
Let $\mathrm{Ax}(V[\mathcal{F}]) \cup \mathrm{Red}(\Sigma,\sigma) = \{p_i \approx q_i : 1 \le i \le n\}$. Let $M(x,y,z) = t(x,t(x,y,z),z)$.

STEP 8: Let y be a variable not appearing in the above list of n equations $p_i \approx q_i$. Then let $s_i = t(p_i,q_i,y) \qquad (1 \le i \le n)$.

STEP 9: Let $y_1,\ldots,y_{n-1}$ be new variables, and let $s'_1 = s_1$, $s'_{i+1} = M(s'_i,s_{i+1},y_i)$ $(1 \le i \le n-1)$.

STEP 10: Let $p = M(M(x,y,y),u,M(M(x,y,y),s'_{n-1},z))$, where z,u are new variables.

THEOREM 5.3. $\Sigma \models \sigma$ *holds iff*

(a) *one-element models of* Σ *are models of* σ, *and*

(b) $\forall(p \approx y) \models \forall x \forall y\,(x \approx y)$,

where $\forall(p \approx y)$ means all variables in $p \approx y$ are universally quantified.

PROOF. For the justification of the steps 8–10 see McKenzie (1975) and Padmanabhan (1977). ■

6. Reducing Mathematics to Equational Logic

In 1925 von Neumann was able to give a *finite* first-order axiomatization of set theory by introducing classes as well as sets. This was refined later by Bernays, and then Gödel used a slight modification of Bernays axioms in his 1940 study of the continuum hypothesis. Set theory is sufficiently powerful to encode all of mathematics, so if we combine set theory with Gödel's completeness theorem (for the first-order logic), proved in 1930, we see that we have a first-order system in which every mathematical fact can be routinely encoded; and such a fact can be proved with our current mathematical tools iff its encoded form can be proved in this set theory using the rules of first-order logic. Below we list the finitely many axioms of this set theory.

VON NEUMAN-BERNAYS-GÖDEL AXIOMS FOR SET THEORY

This is a list of first-order axioms for set theory with a single binary predicate symbol $\in$. The basic idea is to work with *classes* (which can be very big), and *sets* (which are those classes which stand in the relation $\in$ to some class). In the following all variables in lower case are to be relativized to sets, e.g., replace $\forall x$ by $\forall x(\mathrm{set}(x) \rightarrow$, and replace $\exists x$ by $\exists x(\mathrm{set}(x) \wedge$.

DEFINITIONS

$\mathrm{set}(X)$	$=:$	$\exists Y\, X \in Y$
$\{x, y\}$	$=:$	the unique z in Axiom 2 below
$\{x\}$	$=:$	$\{x, x\}$
$\langle x, y\rangle$	$=:$	$\{\{x\}, \{x, y\}\}$
$\langle x, y, z\rangle$	$=:$	$\langle x, \langle y, z\rangle\rangle$
$\emptyset(x)$	$=:$	$\forall u \neg u \in x$
$\mathrm{Un}(x)$	$=:$	$\forall u \forall v \forall w\, [\langle v, u\rangle \in x \wedge \langle w, u\rangle \in x \rightarrow v \approx w]$
$\mathrm{sep}(x, y)$	$=:$	$\forall u \neg(u \in x \wedge u \in y)$
$x \subseteq y$	$=:$	$\forall z\,(z \in x \rightarrow z \in y)$
$x \subset y$	$=:$	$x \subseteq y \wedge x \not\approx y$

AXIOMS

1. $\forall u\,(u \in X \leftrightarrow u \in Y) \rightarrow X \approx Y$
2. $\forall x \forall y \exists z\,(u \in z \leftrightarrow u \approx x \vee u \approx y)$
3. $\exists E \forall x \forall y\,(\langle x, y\rangle \in E \leftrightarrow x \in y)$
4. $\forall A \forall B \exists C \forall u\,(u \in C \leftrightarrow u \in A \wedge u \in B)$
5. $\forall A \exists B \forall u\,(u \in B \leftrightarrow \neg u \in A)$
6. $\forall A \exists B \forall x\,(x \in B \leftrightarrow \exists y(\langle y, x\rangle \in A))$

7. $\forall A \exists B \forall x \forall y\, (\langle y, x\rangle \in B \leftrightarrow x \in A)$
8. $\forall A \exists B \forall x \forall y\, (\langle x, y\rangle \in B \leftrightarrow \langle y, x\rangle \in A)$
9. $\forall A \exists B \forall x \forall y \forall z\, (\langle x, y, z\rangle \in B \leftrightarrow \langle y, z, x\rangle \in A)$
10. $\forall A \exists B \forall x \forall y \forall z\, (\langle x, y, z\rangle \in B \leftrightarrow \langle x, z, y\rangle \in A)$
11. $\exists A\, [\neg \emptyset(A) \wedge \forall x\, (x \in A \rightarrow \exists y\, (y \in A \wedge x \subset y))]$
12. $\forall x \exists y \forall u \forall v\, (u \in v \wedge v \in x \rightarrow u \in y)$
13. $\forall x \exists y \forall u\, (u \subseteq x \rightarrow u \in y)$
14. $\forall x \forall A\, [\mathrm{Un}(A) \rightarrow \exists y \forall u\, (u \in y \leftrightarrow \exists v\, (v \in x \wedge \langle u, v\rangle \in A))]$
15. $\neg \emptyset(A) \rightarrow \exists u\, (u \in A \wedge \mathrm{sep}(u, A))$
16. $\exists A\, [\mathrm{Un}(A) \wedge \forall x (\neg \emptyset(x) \rightarrow \exists y\, (y \in x \wedge \langle y, x\rangle \in A))]$.

If we let Σ be the above set of axioms then we can take any mathematical assertion and encode it as a sentence σ in the language of set theory, and then use the results of the previous section to reduce the assertion $\Sigma \models \sigma$ to an equivalent assertion about a finite set of equations implying $x \approx y$. Thus we have our final reduction of mathematics to equational logic.

Alternate Reductions of First-Order Logic

In Tarski & Givant (1987) one has a reduction of first-order Zermelo-Fraenkel set theory to traditional (one-sorted) equational logic by using a sophisticated encoding into the equational logic of *relation algebras*. The reduction presented here seems to be rather more transparent.

Dershowitz & Hsiang (1983) showed how to work with clauses as *two-sorted terms*, one sort referring to the original first-order predicates and operations, the other to Boolean operations (including Boolean ring operations). The thrust of their work was to introduce term rewrite rules for these two-sorted terms for refutational theorem proving — the most popular versions use the Boolean rewrite rules of Hsiang (1985); see, e.g., the complete system in Bachmair & Dershowitz (1987). However admitting equality ($\approx$) created problems. Hsiang (1985) proposed a rewrite system, but it is not known to be complete when working with equality. However Hsiang (1987) presents a complete set of rewrite rules for refutational theorem proving using two-sorted terms with equality present.

Paul (1985), following Dershowitz & Hsiang, gave a Birkhoff-style completeness theorem for an equational version of the two-sorted terms without equality; and succeeded in doing the same for restricted cases with equality.

The McKenzie reduction of first-order logic with equality to one-sorted equational logic has, in full generality, a completeness theorem (Theorem 5.2) based on Birkhoff's completeness theorem. However we have not yet determined the efficiency with which this result can be used, nor do we have any meaningful comparisons with the alternatives above.

Concluding Remarks

First a few comments on the foundations of mathematics. In 1935 and 1944 Birkhoff proved two fundamental results in universal algebra: (1) the completeness of the five rules of inference, and (2) every algebra is isomorphic to a subdirect product of subdirectly irreducible algebras. These basic results of Birkhoff are the tools needed to show that the first-order notion of $\Sigma \models \sigma$ has an equivalent syntactic version in equational logic (as in section 5). From this one can easily derive the compactness theorem of first-order logic. In this sense universal algebra could be taken as a foundation for first-order logic.

There is of course an interesting parallel with the Löwenheim-Skolem-Herbrand analysis of first-order logic. Given a first-order sentence σ they showed how to generate a list of *ground formulas* $\varphi_1, \ldots$ such that $\models \varphi$ holds iff some φ_i is not satisfiable (see, e.g., Mostowski (1965)). The reduction we have described in this section would take a sentence φ and produce a *finite* set of *equations* $\varepsilon_1, \ldots, \varepsilon_n$ such that $\models \varphi$ holds iff $\varepsilon_1, \ldots, \varepsilon_n \vdash x \approx y$, where we take the $\vdash$ to be derivation in equational logic. By appropriately introducing constants 0,1 into the discriminator variety V one can show that $\varepsilon_1, \ldots, \varepsilon_n \vdash x \approx y$ iff $\varepsilon_1, \ldots, \varepsilon_n \vdash 0 \approx 1$ iff some ground instances of $\varepsilon_1, \ldots, \varepsilon_n$ suffice to derive $0 \approx 1$. This version would seem to be more in the spirit of their work.

From a theoretical point of view one could even handle all of mathematics as a finite *string rewrite system.* To see this just note that the consequences of von Neumann-Bernays-Gödel set theory form a recursively enumerable set of statements. Thus there is a Turing machine which, given a sentence in the language of this set theory, halts iff that sentence is a theorem. A Turing machine can be easily converted to a finite string rewrite system (and hence, if one wishes, to a finite unary term rewrite system). Thus some finite string rewrite system encompasses all of mathematics. However one would naturally design such a Turing machine around a sweep out search for a proof, and consequently the string rewrite system would be computationally disastrous. (With McKenzie's reduction to equations we have not yet committed ourselves to a strategy for finding an equational proof!)

Since we see that even the simplest formal systems are adequate to capture the whole of mathematics, we return to the fundamental question: which ones can be profitably exploited? The study of equations is of fundamental importance to mathematics, and as we have seen, a better understanding of how to carry out derivations in equational logic could have strong implications for automated theorem proving in all areas of mathematics.

7. Appendix: The Completeness Theorem

We will give the detailed proof of Theorem 5.2, based on McKenzie (1975). First we note that the following equations are consequences of the axioms $\mathrm{Ax}(V)$:

$$t(x,x,y) \approx y \tag{52}$$

$$t(x,y,x) \approx x \tag{53}$$

$$t(x,y,y) \approx x \tag{54}$$

$$t(x,t(x,y,z),y) \approx y. \tag{55}$$

To see this just verify that the above equations hold on the simple algebras in V, using (1), and then use the fact that the simple algebras of V generate V, so the equations must hold on V as well.

Also we have the compatibility equations

$$t(u,v,f(x_1,\ldots,x_n)) = t(u,v,f(t(u,v,x_1),\ldots,t(u,v,x_n))) \tag{56}$$

for *all* the function symbols occurring in $\mathrm{Ax}(V[\mathcal{F}]) \cup \mathrm{Red}_V(\Sigma,\sigma)$ — for a function symbol not in V we know (56) is one of our equations by STEP 7; and for a function symbol in V the equation (56) holds because it is a consequence of $\mathrm{Ax}(V)$ (to see this again use (1) to show that it holds on the simple models of V, and hence on all models of V).

CLAIM 7.1. For $\mathbf{A} \models \mathrm{Ax}(V[\mathcal{F}]) \cup \mathrm{Red}_V(\Sigma,\sigma)$ and for $a, b \in \mathrm{A}$,

$$\Theta_{\mathbf{A}}(a,b) = \{\langle c,d\rangle \in \mathrm{A}\times\mathrm{A} : t(a,b,c) = t(a,b,d)\}, \tag{57}$$

that is, the principal congruence $\Theta_{\mathbf{A}}(a,b)$ of $\mathbf{A}$ generated by the pair $\langle a,b\rangle$ is very simply described by an equation involving the term t. (This is precisely where we need the mysterious compatibility condition.)

PROOF.

(i) The right-hand side of (57) is clearly an equivalence relation.

(ii) For $\langle c_i, d_i\rangle$, $1 \le i \le n$, in the right-hand side of (57) and for f any n-ary function symbol in our language we have

$$\begin{aligned} t(a,b,f(c_1,\ldots,c_n)) &= t(a,b,f(t(a,b,c_1),\ldots,t(a,b,c_n))) && \text{by (56)} \\ &= t(a,b,f(t(a,b,d_1),\ldots,t(a,b,d_n))) && \text{as } t(a,b,c_i) = t(a,b,\\ &= t(a,b,f(d_1,\ldots,d_n)) && \text{by (56).} \end{aligned}$$

Thus $\langle f(\vec{c}), f(\vec{d})\rangle$ is also in the right-hand side of (57).

Items (i) and (ii) show that the right-hand side of (57) is a congruence on $\mathbf{A}$.

(iii) Since $t(a,b,a) = a = t(a,b,b)$ by (53) and (54) it follows that $\langle a,b\rangle$ is in the right-hand side of (57). As $\Theta_{\mathbf{A}}(a,b)$ is the smallest congruence to which $\langle a,b\rangle$ belongs we must have $\Theta_{\mathbf{A}}(a,b) \subseteq \{\langle c,d\rangle \in \mathrm{A}\times\mathrm{A} : t(a,b,c) = t(a,b,d)\}$.

(iv) Now suppose $\langle c,d\rangle$ is in the right-hand side of (57). Then we have $t(a,b,c) = t(a,b,d)$. Let a',b',c',d' be the elements corresponding to a,b,c,d in the quotient algebra $\mathbf{A}/\Theta_{\mathbf{A}}(a,b)$. Then $a' = b'$, and in the quotient algebra we have $t(a',b',c') = t(a',b',d')$, so $t(a',a',c') = t(a',a',d')$. Now we can use (52) to conclude $c' = d'$, and thus $\langle c,d\rangle \in \Theta_{\mathbf{A}}(a,b)$. Consequently $\Theta_{\mathbf{A}}(a,b) \supseteq \{\langle c,d\rangle \in \mathrm{A}\times\mathrm{A} : t(a,b,c) = t(a,b,d)\}$, so the claim is proved. ∎

CLAIM 7.2. The term $t(x,y,z)$ defines a discriminator function on the subdirectly irreducible algebras in the variety defined by

$$\mathrm{Ax}(V[\mathcal{F}]) \cup \mathrm{Red}_V(\Sigma,\sigma). \tag{58}$$

This guarantees that (58) defines a discriminator variety and $t(x,y,z)$ is a discriminator term on the simples of this variety.

PROOF. Let $\mathbf{A}$ be a nontrivial subdirectly irreducible model of (58). Choose a, b such that $\Theta_{\mathbf{A}}(a, b)$ is the monolith of $\mathbf{A}$, i.e., the smallest non-identity congruence. For $c \in \mathbf{A}$, if $t(a, b, c) \neq a$ then $\Theta(a, b) \subseteq \Theta_{\mathbf{A}}(a, t(a, b, c))$ so $\langle a, b\rangle \in \Theta_{\mathbf{A}}(a, t(a, b, c))$, and thus by Claim 7.1 the pair $\langle a, b\rangle$ satisfy a simple equation, namely

$$t(a, t(a, b, c), a) = t(a, t(a, b, c), b). \tag{59}$$

From (53) we have

$$t(a, t(a, b, c), a) = a \tag{60}$$

and from (55) we have

$$t(a, t(a, b, c), b) = b \tag{61}$$

so combining (59),(60) and (61) we have $a = b$; but this is a contradiction. Thus

$$t(a, b, c) = a. \tag{62}$$

But then $t(a, b, c) = t(a, b, d) = a$ for all $c, d \in \mathrm{A}$, and then by Claim 7.1 we see that $\Theta_{\mathbf{A}}(a, b) = \mathrm{A} \times \mathrm{A}$. This says that the monolith is the largest congruence on $\mathbf{A}$, and hence $\mathbf{A}$ is indeed simple. But then for *any* $a \neq b$ in A we have $\Theta(a, b)$ is the monolith, and thus for any a, b, c with $a \neq b$ we must have (62) holding, i.e., $t(a, b, c) = a$. Now by (52) we see that if $a = b$ then $t(a, b, c) = c$. But this shows, by (1), that t indeed defines a ternary discriminator function on each subdirectly irreducible algebra satisfying (58). ■

Now we can use Lemma 3.5 to see that our axioms (58) can be decoded on the simple models of these axioms to give the assertion $\Sigma \cup \{\neg\sigma\}$. If $\Sigma \cup \{\neg\sigma\}$ is not satisfiable then there are no nontrivial simple models of our axioms, and hence no nontrivial models. But then by Birkhoff's completeness theorem we can derive $x \approx y$.

If $\Sigma \cup \{\neg\sigma\}$ is satisfiable in a nontrivial model $\mathbf{A}$ then choose any simple algebra $\mathbf{S}$ in V of the same size (this is where our hypothesis about the generalized spectrum of V comes in) and simply put $\mathbf{A}$ and $\mathbf{S}$ on a common domain to get a structure $\mathbf{C}$. Now replace the relations r of $\mathbf{C}$ by appropriate functions f_r and one has a model $\mathbf{D}$ of our axioms (this is easy to check since t satisfies (1) on $\mathbf{D}$). Consequently we have a nontrivial model of our axioms (58), and therefore we cannot derive $x \approx y$. ■

The author is grateful to the referees for pointing out numerous places where the text in the original version could be improved. This research has been supported by NSERC Grant No. A7256.

REFERENCES

Ackermann, W. (1954). *Solvable Cases of the Decision Problem.* North Holland

Albert, M.H., Lawrence, J. (1991). Solving equations in nilpotent groups. *Preprint.*

Arens, R.F., Kaplansky, I. (1948). Topological representations of algebras. *Trans. Amer. Math. Soc.* **63**, 457–481.

Bernays, P. (1937/41), A system of axiomatic set theory. Part I., *J. Symbolic Logic* **2**, 65–77; Part II., *J. Symbolic Logic* **6**, 1–17.

Birkhoff, G. (1935). On the structure of abstract algebras. *Proc. Camb. Phil. Soc.* **31**, 433–454.

Birkhoff, G. (1944). Subdirect unions in universal algebra. *Bull. Amer. Math. Soc.* **50**, 764–768.

Bloom, S.L., Tindell, R. (1983) Varieties of "if-then-else". *Siam J. Computing* **12**, 677-707.

Bulman-Fleming, S., Werner, H. (1977). Equational compactness in quasi-primal varieties. *Algebra Universalis* **7**, 33–46.

Büttner, W., Simonis, H. (1987). Embedding Boolean expressions into logic programming. *J. Symbolic Computation* **4**, 191 -207.

Burris, S. (1984), Boolean constructions, *in Universal Algebra and Lattice Theory*, Springer Lect. Notes in Math. **1004**, 67–90.

Burris, S., McKenzie, R., Valeriote, M., Decidable discriminator varieties. To appear in *J. Symbolic Logic.*

Burris, S., Sankappanavar, H.P. (1981). *A Course in Universal Algebra.* Graduate Texts in Math. **78**, Springer Verlag.

Burris, S., Werner, H. (1979). Sheaf constructions and their elementary properties. *Trans. Amer. Math. Soc.* **248**, 269–309.

Dauns, J., Hofmann, K.H. (1966). The representation of biregular rings by sheaves. *Math. Z.* **91**, 103–123.

Dershowitz, N. (1989) Completion and its applications. In *Resolution of Equations in Algebraic Structures, Vol. 2*, Academic Press.

Foster, A.L. (1953). Generalized " Boolean" theory of universal algebras. Part I. Subdirect sums and normal representation theorem. *Math. Z.* **58**, 306–336. Part II. Identities and subdirect sums in functionally complete algebras. *Math. Z.* **59** 191–199.

Garey, M.R., Johnson, D.S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W.H. Freeman, Co.

Gödel, K. (1940). *The Consistency of the Axiom of Choice and the Generalized Continuum Hypothesis with the Axioms of Set Theory.* Princeton Univ. Press.

Henkin, L., Monk, D., Tarski. A. (1975/85). *Cylindric algebras. Part I; Part II.* North Holland.

Herbrand, J. (1930). *Recherches sur la théorie de la démonstration.* Travaux de la Société des Sciences et des Lettres Varsovie. Cl. III, **33**, 128 pp.

Herstein, I.N. (1975). *Topics in Algebra.* 2nd Edition. John Wiley & Sons.

Hsiang, J., Dershowitz, N. (1983). Rewrite methods for clausal and non-clausal theorem proving. *Proceedings 10th ICALP*, Springer-Verlag Lecture Notes in Computer Science, **154**, 331–346.

Hsiang, J. (1985). Refutational theorem proving using term rewriting systems. *Artificial Intelligence*, **25**, 255–300.

Hsiang, J. (1987). Rewrite method for theorem proving in first-order theory with equality. *J. Symbolic Computation* **3**, 133–151.

Huntington, E.V. (1904). Sets of independent postulates for the algebra of logic. *Trans. Amer. Math. Soc.* **5**, 288–309.

Jacobson. N. (1945). Structure theory for algebraic algebras of bounded degree. *Ann. of Math.* **46**, 695 -707.

McCoy, N.H., Montgomery, D. (1937). A representation of generalized Boolean rings. *Duke Math. J.* **3**, 455–459.

Lawrence, J (1991a) A note on the unification of group equations. *Preprint.*

Lawrence, J (1991b) Unification in nilpotent and solvable varieties of groups. *Preprint.*

McKenzie, R. (1975). On spectra, and the negative solution of the decision problem for identities having a finite non-trivial model. *J. Symbolic Logic* **40**, 186–196.

McKenzie, R., McNulty, G., Taylor, W. *Algebras, Lattices, Varieties. Vol. I.* Wadsworth & Brooks/Cole, 1987.

McNulty, G. (1989). An equational logic sampler. In *Rewriting Techniques and Applications, RTA-89 Proceedings*, ed. N. Dershowitz, Lecture Notes in Computer Science, **355**, 234–262.

Mekler, A.H., Nelson, E.M. (1987). Equational bases for if-then-else. *Siam J. Computing* **16**, 465–485.

Mostowski, A. (1965). *Thirty Years of Foundational Studies.* Acta Philosophica Fennica **17**.

Nipkow, T. (1990). Unification in primal algebras, their powers and their varieties. *J. Assoc. Comp. Mach.* **37**, 742–776.

Padmanabhan, R. (1977). Equational theory of algebras with a majority polynomial. *Algebra Universalis* **7**, 273–275.

Paul, E. (1985). Equational methods in first-order predicate calculus. *J. Symbolic Computation* **1**, 7–29.

Robinson, J.A. (1965). A machine oriented logic based on the resolution principle. *J. Assoc. Comp. Mach.* **12**, 23–41.

Rosenbloom, P.C. (1942). Post algebras I. Postulates and general theory. *Amer. J. Math.* **64**, 167–188.

Schröder, E. (1890–1902). *Algebra der Logik. Vol. I–III.* Chelsea Reproductions.

Siekmann, J. (1989). Unification theory. *J. Symbolic Computation* **7**, 207–274.

Tarski, A., Givant, S. (1987). *A Formalism of Set Theory without Variables.* Amer. Math. Soc. Colloq. Publications **41**.

Taylor, W. (1979). Equational Logic. *Houston J. of Math* **5**.

von Neumann, J. (1925). Eine Axiomatisierung der Mengenlehre. *J. Reine Angew. Math.* **154**, 219–240.

Werner, H. (1978). *Discriminator Algebras.* Studien zur Algebra und ihre Anwendungen, Band 6, Akademie-Verlag, Berlin.