Notre Dame Journal of Formal Logic
Volume 31, Number 2, Spring 1990

The Homogeneous Form of Logic Programs with Equality

WILLIAM DEMOPOULOS*

Abstract Let P be a Horn clause logic program. We suppose that P is *symmetric* in the sense that if C is a clause in P whose head is $s = t$, then there is a clause C^* in P which is like C except for having the head $t = s$. The *homogeneous form* of a clause $p(t_1, \ldots, t_n) \leftarrow B_1, \ldots, B_q$ is $p(x_1, \ldots, x_n) \leftarrow x_1 = t_1, \ldots, x_n = t_n, B_1, \ldots, B_q$. The *homogeneous form* P' of P is the set of homogeneous forms of clauses of P. Let T be a set of axioms asserting the reflexivity, symmetry, transitivity, and congruence (with respect to the predicates of P) of $=$. Then $P \cup T$ is *goal equivalent* to $P' \cup \{x = x\}$; i.e., for any goal G, $P \cup T \cup \{G\}$ is unsatisfiable iff $P' \cup \{x = x\} \cup \{G\}$ is unsatisfiable. The main interest of the paper lies in its construction of the Herbrand model M and in the proof that M is the *minimal* Herbrand model of both $P \cup T$ and $P' \cup \{x = x\}$.

1 Introduction In their analysis of Prolog II van Emden and Lloyd [1] introduce the notion of the homogeneous form of a Horn clause logic program.

Definition The *homogeneous form* of a clause $p(t_1, \ldots, t_n) \leftarrow B_1, \ldots, B_q$ is

$$p(x_1, \ldots, x_n) \leftarrow x_1 = t_1, \ldots, x_n = t_n, B_1, \ldots, B_q$$

where $x_1, \ldots, x_n$ are distinct variables not appearing in the original clause.

Definition Let P be a program. The *homogeneous form* P' of P is the set of homogeneous forms of the clauses in P.

*I wish to thank E. W. Elcock, Edward P. Stabler, and Peter Hoddinott for introducing me to the theory of logic programming. I wish to thank A. Abdallah for suggesting Section 4. I am especially indebted to Kwok Hung Chan for carefully reading earlier drafts and providing critical comments and helpful suggestions too numerous to detail. Research support by the Social Sciences and Humanities Research Council and by the IBM Corporation is gratefully acknowledged.

Received August 31, 1987; revised March 23, 1989

van Emden and Lloyd use the homogeneous form to provide a simple and elegant characterization of the relation of Prolog II to standard Prolog. Their results may be summarized as follows. Let E be the equality theory

(1) $\forall x\, x = x$
(2) $\forall x \forall y\, x = y \rightarrow y = x$
(3) $\forall x \forall y \forall z\, x = y \wedge y = z \rightarrow x = z$
(4) $\forall x_1 \cdots \forall x_n \forall y_1 \cdots \forall y_n (x_1 = y_1) \wedge \cdots \wedge (x_n = y_n) \rightarrow f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n)$, for all function symbols f
(5) $\exists x_1 \cdots \exists x_n \exists y_1 \cdots \exists y_k (x_1 = t_1) \wedge \cdots \wedge (x_n = t_n)$, where the x_i's are distinct variables, the t_i's are terms, and $\{x_1, \ldots, x_n, y_1, \ldots, y_k\}$ is the set of all variables in the formula.

Let P be a logic program not containing the equality predicate, let P' be its homogeneous form, and let G be a goal (again, not containing =). Then van Emden and Lloyd's first proposition states that $P \cup \{G\}$ is unsatisfiable iff $P' \cup \{x = x\} \cup \{G\}$ is unsatisfiable. And their second proposition tells us that Prolog II solves G only if $P' \cup E \cup \{G\}$ is unsatisfiable. Thus Prolog II is sound with respect to the first-order theory $P' \cup E \cup \{G\}$, while standard Prolog is sound (and complete) with respect to the first-order theory $P' \cup \{x = x\} \cup \{G\}$.

In general, logic programming systems which differ from standard Prolog only in their unification theory can be usefully analyzed within van Emden and Lloyd's framework: one first applies the homogeneous form transformation to the logic program in order to obtain a first-order theory; this theory is then combined with an equality theory, such that the original programming system is sound (and possibly complete) with respect to the combination, homogeneous form plus equality theory. The equality theory specifies the conditions under which two terms are unifiable.

Our interest in the homogeneous form of a logic program begins with the observation that the homogeneous form of a definite clause is itself a definite clause. It is thus possible to view the homogeneous form of a set of clauses as simply another logic program. We have found that when we lift the restriction that P and G be equation-free, the homogeneous form can be shown to have a theoretical interest which is independent of its role as a tool for comparison. This interest, together with our principal results, will now be briefly summarized.

Unlike van Emden and Lloyd we allow for the possibility that the head of the original clause may be an *equation*, i.e., an atomic formula whose predicate symbol is =.

We restrict ourselves to what we term *symmetric programs*. If P is a symmetric program, then an equation $s = t$ occurs in the head of a program clause C of P only if there is a program clause C^* in P which is like C except that it has the head $t = s$. Aside from this requirement, we impose no restriction on P beyond the usual one that it be a finite set of definite Horn clauses. Thus P may contain equations, as well as other literals, and these may occur in any order in the program clauses of P. The restriction to symmetric programs is a natural one in the case of programs with equality. It is also clear that any definite clause logic program can always be extended to a symmetric program.

We are interested in the relation the homogeneous form of a symmetric program bears to the result of combining the original program with a particular equality theory. Specifically, we are interested in the question: To what extent does transforming a symmetric program P into its homogeneous form P' allow us to approximate the deductive power of $P \cup T$, where T is the equality theory

(1) $\forall x\, x = x$
(2) $\forall x \forall y\, x = y \rightarrow y = x$
(3) $\forall x \forall y \forall z\, x = y \wedge y = z \rightarrow x = z$
(4) $\forall x_1 \cdots \forall x_n \forall y_1 \cdots \forall y_n (x_1 = y_1) \wedge \cdots \wedge (x_n = y_n) \wedge p(x_1, \ldots, x_n) \rightarrow p(y_1, \ldots, y_n)$,

and (4) is an axiom schema with instances determined by the predicates occurring in P?

Letting G be any goal (thus G may contain equations), our principal result implies that if $P \cup T \cup \{G\}$ is unsatisfiable so is $P' \cup \{x = x\} \cup \{G\}$. (The proof of the converse is trivial.) The homogeneous form thus allows us to recover the effect of Axioms (3) and (4) of T without introducing transitivity and predicate substitutivity axiomatically. From the point of view of logic programming, the deductive power of $P' \cup \{x = x\}$ is, therefore, practically speaking equivalent to the deductive power of $P \cup T$, although they are not, of course, logically equivalent.

The main interest of the paper lies in the construction of the Herbrand model M in Section 2.1, and in the subsequent proof of its principal result, viz, that M is in fact the minimal Herbrand model of both $P \cup T$ and $P' \cup \{x = x\}$. We believe that our construction of M facilitates an especially transparent analysis of the Herbrand model theory of programs in homogeneous form. In Section 4 we show how M may be obtained as the least fixpoint of a program operator whose definition is motivated by the model-theoretic construction of Section 2.1. Section 3 relates our construction of M to the standard concept of transitive closure.

Before concluding this introduction let us note why an answer to our original question follows from the identity of the minimal Herbrand models of $P \cup T$ and $P' \cup \{x = x\}$. The following proposition is implicit in Lloyd [2], whom we follow throughout for terminology and notation.

Proposition 0 *Let P be any logic program and let M_P be its minimal Herbrand model. Then $P \cup \{G\}$ holds in some Herbrand model K iff $P \cup \{G\}$ holds in M_P.*

Proof: The $\Leftarrow$-direction is trivial. For the $\Rightarrow$-direction it clearly suffices to prove that $M_P \models G$. Let G be $\leftarrow A_1, \ldots, A_n$. Since $K \models G$, it follows that for every ground substitution θ there is an i, $i = 1, \ldots, n$, such that $K \models \leftarrow A_i\theta$. Hence $A_i\theta \notin K$. But then $A_i\theta \notin M_P$ since $M_P \subseteq K$. Hence for every θ, $M_P \models G\theta$, and therefore $M_P \models G$.

The following proposition is proved by van Emden and Lloyd in the course of the proof of their first proposition. The restriction which they impose on P and G (that they be equation-free) is not appealed to in the course of their proof.

Proposition 1 (van Emden and Lloyd) *Let P be a program, P' the homogeneous form of P. Then $P' \cup \{x = x\}$ logically implies P.*

2 Model-theoretic properties of the homogeneous form

2.1 General symmetric programs

Definition Let P be a program. Then P is *symmetric* if whenever a clause $s = t \leftarrow B_1, \ldots, B_q$ is in P so is $t = s \leftarrow B_1, \ldots, B_q$.

Throughout this section P is an arbitrary but fixed symmetric program, and P' is its homogeneous form. T is the equality theory defined in Section 1. To keep the notation simple, we consider only $=$ and the single unary predicate symbol p. (We also drop unnecessary parentheses and write pt for $p(t)$.) The generalization to predicates of arbitrary arity is, in all cases, obvious. We let H be the minimal Herbrand model of P; U_L is the universe of H. Arbitrary elements of U_L are denoted by the letters a, b, c, d, with or without subscripts. Notice, therefore, that a, b, c, and d are *not* necessarily constants but may be arbitrary terms. The letters s and t denote arbitrary terms which may or may not belong to U_L; the context will always make clear which alternative is intended.

Let $[x = x]$ denote the set of all ground instances of $x = x$. We define a double sequence of Herbrand models, all with the same Herbrand universe U_L, as follows:

$$M_0 = H \cup [x = x]$$

$$N_0 = H$$

$$\begin{aligned} M_{n+1} = M_n \cup N_n \cup \{a = b \mid\ & \text{there are } s \text{ are } t \text{ such that} \\ & a = s,\ b = t \in M_n \cup N_n, \text{ and } s = t \in N_n\} \\ \cup \{pa \mid\ & \text{there is a } t \text{ such that} \\ & a = t \in M_n \cup N_n, \text{ and } pt \in N_n\} \end{aligned}$$

$$\begin{aligned} N_{n+1} = N_n \cup \{a = b \mid\ & \text{there is a clause } C \text{ in } P \text{ and a substitution } \theta \text{ such that} \\ & C \text{ is } s = t \leftarrow B_1, \ldots, B_q, \\ & B_1\theta, \ldots, B_q\theta \in M_{n+1}, \\ & (s = t)\theta \text{ is } a = b\} \\ \cup \{pa \mid\ & \text{there is a clause } C \text{ in } P \text{ and a substitution } \theta \text{ such that} \\ & C \text{ is } pt \leftarrow B_1, \ldots, B_q, \\ & B_1\theta, \ldots, B_q\theta \in M_{n+1}, \\ & (pt)\theta \text{ is } pa\}. \end{aligned}$$

We call M_{n+1} a *rectangle completion* of N_n. The following graph of the principal clause in the definition of M_{n+1} gives the intuition underlying this terminology:

$$\begin{tikzcd} a \arrow[r, "M_n/N_n"] & s \arrow[d, "N_n"] \\ b \arrow[r, "M_n/N_n"] & t \end{tikzcd}$$

Γ

The graph Γ is read left-to-right and top-down. Thus Γ represents the following: the equations $a = s$, $b = t$ belong either to M_n or to N_n, and $s = t$ belongs to N_n. M_{n+1} *completes the rectangle* by including the equation $a = b$.

The double sequence of Herbrand models is pictured below. Boxes ($\square$) signify rectangle completions and double arrows ($\Rightarrow$) signify closures under program clauses.

$$\begin{tikzcd}
M_0 \arrow[r, "\square"] & M_1 \arrow[d, "\Downarrow"] & \cdots & M_n \arrow[r, "\square"] \arrow[d, "\Downarrow"'] & M_{n+1} \arrow[d, "\Downarrow"] & \cdots \\
N_0 \arrow[r, "\Rightarrow"] \arrow[ur, "\square"] & N_1 & \cdots & N_n \arrow[r, "\Rightarrow"] \arrow[ur, "\square"] & N_{n+1} & \cdots
\end{tikzcd}$$

Definition $\quad M = \bigcup_{n<\omega} M_n.$

Definition $\quad N = \bigcup_{n<\omega} N_n.$

Remark We observe that $N \subseteq M$, although the converse inclusion need not hold.

In the following theorems M and N are the Herbrand models just defined (relative to the symmetric program P).

Theorem 0 *M is a model of $P' \cup \{x = x\}$.*

Proof: M is a model of $\{x = x\}$ since M and M_0 have the same universe U_L, M_0 is a model of $\{x = x\}$, and $M_0 \subseteq M$.

To show that M is a model of P', let C' in P' be $x = y \leftarrow x = s$, $y = t$, $B_1, \ldots, B_q$. Let θ be a substitution such that $x\theta$ is a and $y\theta$ is b. Suppose further that $B_1\theta, \ldots, B_q\theta, (x = s)\theta, (y = t)\theta \in M$. Then there is an n such that these are all in M_n. Since $B_1\theta, \ldots, B_q\theta \in M_n$ and $C\theta$ is a ground instance of a clause in P, it follows that $(s = t)\theta \in N_n$. Thus we have

$$\begin{tikzcd}
a \arrow[r, "M_n"] & s\theta \arrow[d, "N_n"] \\
b \arrow[r, "M_n"] & t\theta
\end{tikzcd}$$

Hence by the definition of M_{n+1}, $a = b \in M_{n+1}$; i.e. $(x = y)\theta \in M_{n+1}$.

If C' is $px \leftarrow x = t$, $B_1, \ldots, B_q$, then C is $pt \leftarrow B_1, \ldots, B_q$. Let θ be such that $(x = t)\theta$, $B_1\theta, \ldots, B_q\theta \in M$ (and thus, in some M_n). Since $C\theta$ is a ground instance of a clause in P, $(pt)\theta \in N_n$. Hence $(px)\theta \in M_{n+1}$.

Theorem 1 *M is the minimal Herbrand model of $P' \cup \{x = x\}$.*

Proof: The proof is by induction: We first show that any Herbrand model K of $P' \cup \{x = x\}$ contains M_0. This is the *base case*. Next we show that if K contains

M_n, it contains M_{n+1}, providing the *inductive step.* It follows that K contains M. Since K is arbitrary, it follows that M is minimal.

Base case. Let K be a model of $P' \cup \{x = x\}$. Then by Proposition 1, K is a model of P. Thus K is a model of $P \cup \{x = x\}$. Thus $K \supseteq [x = x]$, and since H is the minimal model of P, $K \supseteq H$. Hence $K \supseteq H \cup [x = x] = M_0$.

Inductive step. (1) Suppose $a = b \in M_{n+1}$. We have the following cases:

(i) $a = b \in M_n$. Then $a = b \in K$ by the inductive hypothesis.

(ii) $a = b \in N_n$. Then by the definition of N_n, there is a clause C in P and a substitution θ such that C is $s = t \leftarrow B_1, \ldots, B_q$, $B_1\theta, \ldots, B_q\theta \in M_n$ and $(s = t)\theta$ is $a = b$. Since by the inductive hypothesis, $K \supseteq M_n$, $B_1\theta, \ldots, B_q\theta \in K$. By hypothesis K is a model of $P' \cup \{x = x\}$; thus by Proposition 1 it follows that K is a model of P. Hence $(s = t)\theta \in K$. But $(s = t)\theta$ is $a = b$.

(iii) There are s, t such that $a = s \in M_n/N_n$, $s = t \in N_n$, $b = t \in M_n/N_n$ (i.e. $a = s$, $b = t$ are in M_n or N_n). Since $s = t \in N_n$, there is a clause C in P and a θ such that $C\theta$ is $s = t \leftarrow B_1\theta, \ldots, B_q\theta$. It follows that C' in P' has the instance

$$x = y \leftarrow x = s,\ y = t,\ B_1\theta, \ldots, B_q\theta. \tag{$\dagger$}$$

Since K is a model of P', K is a model of ($\dagger$). If $a = s$, $b = t \in M_n$ then $a = s$, $b = t \in K$, by the inductive hypothesis. If one (or both) of $a = s$, $b = t$ belongs to N_n, then by the argument for Case (ii) that equation belongs to K. Thus in any of the four possible cases, $a = s$, $b = t \in K$. Hence $a = b \in K$.

(2) Next we observe that the argument for a literal $pa \in M_{n+1}$ is exactly similar. We include a slightly abbreviated version for completeness.

Suppose $pa \in M_{n+1}$. There are three cases:

(i) $pa \in M_n$. Then by the inductive hypothesis $pa \in K$.

(ii) $pa \in N_n$. Then by the definition of N_n there is a clause C and a substitution θ such that $C\theta$ is $pa \leftarrow B_1\theta, \ldots, B_q\theta$ and $B_1\theta, \ldots, B_q\theta \in M_n$. Hence $pa \in K$.

(iii) There is a t such that $a = t \in M_n$ and $pt \in N_n$. Hence there is a clause C and a substitution θ such that $pt \leftarrow B_1\theta, \ldots, B_q\theta$. Thus C' in P' has the instance

$$px \leftarrow x = t,\ B_1\theta, \ldots, B_q\theta. \tag{$\ddagger$}$$

Since K is a model of P', K is a model of ($\ddagger$). Since $a = t \in M_n$, $a = t \in K$ by the inductive hypothesis. Thus $pa \in K$.

Theorem 2 *M is a model of $P \cup T$.*

Proof: We have already shown that M is a model of $P \cup \{x = x\}$. There are three cases remaining.

(1) $\forall x \forall y\ x = y \rightarrow y = x$. We show that M_n is a model of (1) for all n.

Base case. $M_0 \vDash (1)$. Let $a = b \in M_0$. Then there is a clause C in P and a substitution θ such that $C\theta$ is $a = b \leftarrow B_1\theta, \ldots, B_q\theta$, with $B_1\theta, \ldots, B_q\theta \in M_0$. Since P is symmetric, there is a clause C^* in P such that $C^*\theta$ is $b = a \leftarrow B_1\theta, \ldots, B_q\theta$. Hence $b = a \in M_0$.

Inductive step. If $M_n \models (1)$, then $M_{n+1} \models (1)$. Let $a = b \in M_{n+1}$.

(i) $a = b \in M_n$. Then $b = a \in M_n$ by the inductive hypothesis.

(ii) $a = b \in N_n$. Then there is a clause C in P such that for some θ, $C\theta$ is $a = b \leftarrow B_1\theta, \ldots, B_q\theta$ and $B_1\theta, \ldots, B_q\theta \in M_n$. Hence, since P is symmetric we also have that $b = a \leftarrow B_1\theta, \ldots, B_q\theta$ is an instance of a clause in P. It now follows by the definition of N_n that $b = a \in N_n$.

(iii) There are s, t such that $a = s \in M_n/N_n$, $s = t \in N_n$, $b = t \in M_n/N_n$. Hence by the argument for Case (ii) we have $t = s \in N_n$. So by the definition of M_{n+1}, $b = a \in M_{n+1}$.

(2) $\forall x \forall y \forall z\, x = y \wedge y = z \rightarrow x = z$. Notice that (2) follows from the following proposition:

$$\text{For all } n, \text{ if } a = b,\ b = c \in M_n \text{ then } a = c \in M. \qquad (\dagger)$$

We will prove (†) by induction on n. Before proceeding with the proof notice that if one of $a = b$ or $b = c$ is a syntactic identity, the result follows immediately. Thus we may assume without loss of generality that neither equation is a syntactic identity.

Base case. $n = 0$. Then $a = b$, $b = c \in H = N_0$. Since $c = c \in M_0$, we have $a = b$, $c = c \in M_0$ and $b = c \in n_0$. Hence $a = c \in M_1 \subseteq M$.

Inductive step. Suppose (†) holds for n and let $a = b$, $b = c \in M_{n+1}$. We have the following cases:

(i) Both $a = b$, $b = c \in M_n$. Then (†) follows by the inductive hypothesis.

(ii) At least one of $a = b$, $b = c$ is in N_n. Then the argument is essentially identical to the base case. E.g., if $b = c \in N_n$ then $a = b$, $c = c \in M_{n+1}$; then $a = c \in M_{n+2} \subseteq M$.

(iii) There are s, t such that $a = s$, $b = t \in M_n/N_n$, $s = t \in N_n$, and $b = c \in M_n$. Or there are s, t such that $b = s$, $c = t \in M_n/N_n$, $s = t \in N_n$, and $a = b \in M_n$. Consider the first alternative. There are two subcases.

(iiia) $b = t \in M_n$. Then by part (1) (symmetry) and the inductive hypothesis we have $c = t \in M$, and hence $c = t \in M_l$ for some l. Thus we have $a = s \in M_n/N_n$, $c = t \in M_l$, and $s = t \in N_n$. Let j be the greater of l, n. Then $a = c \in M_{j+1} \subseteq M$.

(iiib) $b = t \in N_n$. Then by symmetry and the fact that $M_0 \subseteq M_n$ we have $c = b$, $t = t \in M_n$. Since $b = t \in N_n$, $c = t \in M_n/N_n$, $a = s \in M_n/N_n$, and $s = t \in N_n$. Thus $a = c \in M_{n+2} \subseteq M$.

(iv) For both $a = b$, $b = c$ there are a_1 and b_1, and b_2 and c_1 such that $a = a_1$, $b = b_1 \in M_n/N_n$, $a_1 = b_1 \in N_n$, $b = b_2$, $c = c_1 \in M_n/N_n$, and $b_2 = c_1 \in N_n$. If both $b_2 = b$ and $b = b_1$ belong to M_n then $b_2 = b_1 \in M$ by the inductive hypothesis (compare Case (i)). If $b = b_2$ or $b = b_1$ belongs to N_n, then $b_2 = b_1$, by the argument for Case (ii). So in either case $b_2 = b_1 \in M$, and therefore $b_2 = b_1 \in M_l$, for some l. By symmetry $b_1 = b_2 \in M_l$. By hypothesis $b_2 = c_1 \in N_n$ and $c = c_1 \in M_n/N_n$. Thus $b_1 = c \in M_{j+1}$, where j is the greater of l, n. By symmetry, $c = b_1 \in M_{j+1}$. By hypothesis $a = a_1 \in M_n/N_n$ and $a_1 = b_1 \in N_n$. Hence $a = c \in M_{j+2} \subseteq M$.

(3) $\forall x \forall y\, x = y \wedge px \rightarrow py$. Let $a = b$, $pa \in M$. There are two cases to consider, according to whether $pa \in N_n$ for some n or $pa \notin N_n$ for any n.

(i) $pa \in N_n$ and $a = b \in M_k$ for some n and k. Without loss of generality we may suppose that $k > n$. Then $pa \in N_k$ and $a = b \in M_k$ so that $pb \in M_{k+1} \subseteq M$.

(ii) $pa \in M_n$ for some $n > 0$ and $pa \notin N_n$ for any n. Let n be the smallest index such that $pa \in M_n$. Then by the definition of M_n there is a t such that $a = t \in M_{n-1} \cup N_{n-1}$ and $pt \in N_{n-1}$. By hypothesis $a = b \in M$ and therefore $a = b \in M_q$ for some q. Without loss of generality we may suppose that $q > n$. Then $a = t$, $a = b \in M_q$. By part (1) $b = a \in M_q$. Thus by part (2) $b = t \in M_{q+1}$. Since $pt \in N_{n-1}$, $pt \in N_{q+1}$. Thus $pb \in M_{q+2} \subseteq M$.

Theorem 3 *M is the minimal Herbrand Model of $P \cup T$.*

Proof: Our strategy follows the proof of Theorem 1. Let K be a model of $P \cup T$.

Base case. $M_0 = H \cup [x = x] \subseteq K$, since K is a model of $T \supseteq \{x = x\}$ and H is the minimal model of P.

Inductive step. If $K \supseteq M_n$, then $K \supseteq M_{n+1}$.

(1) Let $a = b \in M_{n+1}$.

(i) $a = b \in M_n$. Then $a = b \in K$ by the inductive hypothesis.

(ii) $a = b \in N_n$. Then $a = b \in K$ by an argument identical to our earlier proof of Theorem 1 (Case (ii) of the inductive step).

(iii) There are s, t such that $a = s \in M_n/N_n$, $s = t \in N_n$, and $b = t \in M_n/N_n$. As in the proof of Theorem 1 (see Case (iii) of the inductive step), we may infer that $a = s \in K$, $s = t \in K$, and $b = t \in K$. Thus we have

$$\begin{tikzcd} a \arrow[r, dash, "K"] \arrow[d, dash, dashed, "K"'] & s \arrow[d, dash, "K"] \arrow[dl, dash, dashed, "K"] \\ b \arrow[r, dash, "K"'] & t \end{tikzcd}$$

where the broken lines represent equations in K which are inferred by the transitivity of the equality relation of K.

(2) *Let* $pa \in M_{n+1}$. Again there are three cases to consider:

(i) $pa \in M_n$. Then $pa \in K$ by the inductive hypothesis.

(ii) $pa \in N_n$. Then there is a clause C in P and a substitution θ such that

$$pa \leftarrow B_1\theta, \ldots, B_q\theta \tag{†}$$

and

$$B_1\theta, \ldots, B_q\theta \in M_n.$$

Thus by the inductive hypothesis $B_1\theta, \ldots, B_q\theta \in K$. Since, by the inductive hypothesis, K is a model of (†), $pa \in K$.

(iii) There is a t such that $a = t \in M_n$ and $pt \in N_n$. By Case (ii), $pt \in K$. And by the inductive hypothesis, $a = t \in K$. Thus $pa \in K$, since by hypothesis predicate substitutivity holds in K.

2.2 Unit clause symmetric programs

Definition A *unit clause symmetric program* is a symmetric program all of whose clauses are of the form $s = t \leftarrow$ or $pa \leftarrow$.

Let P be a unit clause symmetric program, and let P' be its homogeneous form. Let T be the equality theory defined in Section 1. The minimal Herbrand model of $P' \cup \{x = x\}$ is identical to the minimal Herbrand model of $P \cup T$ of Section 2.1. We note that the minimal model takes an especially simple form in this case.

Let H be the minimal Herbrand model of P. We define a sequence of Herbrand models all with the same universe U_L:

$$M_0 = H \cup [x = x]$$

$$\begin{aligned} M_{n+1} = M_n &\cup \{a = b \mid \text{ there are } s, t \text{ such that } \\ &\qquad s = t \in H \text{ and } a = s,\ b = t \in M_n\} \\ &\cup \{pa \mid \text{ there is a } t \text{ such that } \\ &\qquad pt \in H \text{ and } a = t \in M_n\}. \end{aligned}$$

The corresponding diagram is

$$M_0 \xrightarrow{\square} M_1 \cdots M_n \xrightarrow{\square} M_{n+1} \cdots$$

Definition $M = \bigcup_{n<\omega} M_n$.

Thus, in the case of unit clause symmetric programs, the operation of rectangle completion suffices to generate the chain of Herbrand models.

3 Rectangle completion and transitive closure The concept of rectangle completion is closely related to the more standard notion of the transitive closure of a set with respect to a relation R. In this section we first explain this connection in an abstract setting. We conclude by relating our discussion to the minimal Herbrand model of Subsection 2.2.

Let $\mathfrak{M} = \langle M, R \rangle$ be a structure with universe M and binary relation R on M.

Definition $\mathfrak{M}$ is *transitively closed with respect to* R if whenever aRb, bRc hold, so does aRc.

The definition is expressed by a standard diagram:

$$\begin{tikzcd} a \arrow[r, dash] \arrow[dr, dash, dashed] & b \arrow[d, dash] \\ & c \end{tikzcd}$$

Transitive closure

where the broken line represents the relation inferred.

Definition $\mathfrak{M}$ is *$\Box$-complete with respect to R* if whenever aRb, bRc, dRc hold, so does aRd.

The associated diagram is

$$\begin{tikzcd} a \arrow[r, dash] \arrow[d, dash, dashed] & b \arrow[d, dash] \\ d \arrow[r, dash] & c \end{tikzcd}$$

$\Box$-completion

where, as before, the broken line represents the relation inferred. (We drop the explicit reference to R in what follows, and take it to be understood.) The relation between these two concepts is given by the following theorem.

Theorem 4 *Let $\mathfrak{M} = \langle M, R\rangle$ be as above and let M be the* domain *of R, i.e. for every $a \in M$ there is a $b \in M$ such that aRb. Then $\mathfrak{M}$ is $\Box$-complete and R is reflexive iff $\mathfrak{M}$ is transitively closed and R is symmetric.*

Proof: ($\Rightarrow$) Let $\mathfrak{M}$ be $\Box$-complete and R reflexive. We show first that R is symmetric. Suppose aRb holds. Then since R is reflexive bRb holds, i.e.

$$\begin{tikzcd} b \arrow[r, dash] & b \arrow[d, dash] \\ a \arrow[r, dash] & b \end{tikzcd}$$

and since $\mathfrak{M}$ is $\Box$-complete, bRa holds.

To show that R is transitive, let aRb, bRc hold in $\mathfrak{M}$. Since R is reflexive bRb holds, and since R is symmetric cRb also holds. Thus we have aRb, bRb, cRb. Hence by $\Box$-completion, aRc.

($\Leftarrow$) Let $\mathfrak{M}$ be transitively closed and let R be symmetric. To show that R is reflexive, let a be any element of M. Since, by hypothesis, M is the domain of R, there is a b such that aRb. Since R is symmetric bRa also holds. Then by transitive closure

$$\begin{tikzcd} a \arrow[r, dash] \arrow[dr, dash, dashed] & b \arrow[d, dash] \\ & a \end{tikzcd}$$

To show that $\mathfrak{M}$ is $\Box$-complete, notice that by symmetry and transitive closure we have the following inferences

$$\begin{tikzcd} a \arrow[r, dash] \arrow[d, dash, dashed] & b \arrow[d, dash] \\ d \arrow[r, dash] \arrow[ur, dash, dashed] & c \end{tikzcd}$$

Remark The requirement that M be the domain of R is required only in the proof of the $\Leftarrow$-direction; notice also that it is only used to establish that R is reflexive.

Let us now consider the model M defined in Section 2.2 above. For Herbrand models the relevant notions of $\Box$-completion and transitive closure are clearly the following:

Definition Let M be a Herbrand model over U_L, R a binary predicate of L, and $a,b,c,d \in U_L$. Then M is *$\Box$-complete* with respect to R if whenever aRb, $dRc \in M$ and $bRc \in H$, $aRd \in M$.

Definition Let M be a Herbrand model over U_L, R a binary predicate of L, and $a,b,c \in U_L$. Then M is *transitively closed with respect to R* if whenever aRb, $bRc \in M$, $aRc \in M$.

Clearly, if M is the minimal Herbrand model of $P' \cup \{x = x\}$, then M is $\Box$-complete with respect to $=$, and $=$ is reflexive. For a complete coincidence with the definition of the Herbrand model M, we introduce the concept of a *hereditary Herbrand model.*

Definition Let M be a Herbrand model over U_L, and for simplicity let the predicates of L be p and $=$. Then M is *hereditary* if whenever $a = b \in M$ and $pb \in H$, $pa \in M$.

It is obvious that the minimal Herbrand model of $P' \cup \{x = x\}$ is the smallest hereditary and $\Box$-complete Herbrand model of $P' \cup \{x = x\}$.

Notice that there is a subtlety in our notion of a $\Box$-complete Herbrand model that is absent from the abstract notion of $\Box$-completion. This arises from the requirement that one of the equations must come from H. This precludes the use of instances of $\{x = x\}$ unless these arise from program clauses. Since the proof of the quasi-equivalence of rectangle completeness and transitive closure depends on the unlimited availability of reflexive pairs, a general argument of this kind cannot be appealed to in the context of Herbrand models. This is why the proof of part (2) of Theorem 3 is as complicated as it is. The same consideration precludes a simple argument for the symmetry of the equality relation on M (part (1) of Theorem 3). That is, we cannot argue:

> Suppose $a = b \in M$. Then $a = b \in M_n$ for some n. Now $b = b \in N_n$ for all n, since $b = b \in N_0$ and $N_0 \subseteq N_n$ for all n. Thus we have
>
> $$\begin{tikzcd} b \arrow[r, "M_n"] & b \arrow[d, "N_n"] \\ a \arrow[r, "M_n"] & b \end{tikzcd}$$
>
> and therefore $b = a \in M_{n+1} \subseteq M$.

For there is nothing to guarantee the presence of $b = b$ in N_0.

4 A fixpoint operator for M In this section we will describe a fixpoint operator for M. Our aim is to recover the model M of Section 2.1 as the least fixpoint of an operator defined on the complete lattice of all subsets of the Herbrand base for the program $P \cup \{x = x\}$. Of course the fixpoint operators associated with the programs $P \cup T$ and $P' \cup \{x = x\}$ can be shown to have M as their least fixpoint. We believe the operator described here is of interest because of the way it reflects the construction of the minimal model carried out in Section 2.1 above. This section relies heavily on Lloyd [2], to whom the reader is referred for the relevant concepts and propositions.

Let τ_P be the standard fixpoint operator associated with the program $P \cup \{x = x\}$. $\tau_P : 2^B \to 2^B$ is defined on the powerset 2^B of the Herbrand base B by the condition:

$$\tau_P(I) = \{A \mid A \leftarrow A_1, \ldots, A_n \text{ is a ground instance of a clause in } P \cup \{x = x\} \text{ and } \{A_1, \ldots, A_n\} \subseteq I\}.$$

As is well known, the least fixpoint of τ_P, $\mathrm{lfp}(\tau_P)$, is the minimal Herbrand model of $P \cup \{x = x\}$.

Let $\tau_\Box : 2^B \to 2^B$ be defined as follows:

$$\tau_\Box(I) = \{A \mid A \text{ is } a = b \text{ and there are } s, t \text{ such that } a = s,\ s = t,\ b = t \in I \text{ or A is } pa \text{ and there is a } t \text{ such that } a = t,\ pt \in I\}.$$

$\tau_\Box$ is thus the standard fixpoint operator associated with the program $\Box$, where $\Box$ consists of the clauses:

$$x = y \leftarrow x = u,\ u = v,\ y = v$$

$$px \leftarrow x = y,\ py.$$

It follows by standard properties of fixpoint operators associated with definite clause programs that τ_P and $\tau_\Box$ are continuous and thus monotonic. Least fixpoints of τ_P and $\tau_\Box$ therefore exist and we have

$$\mathrm{lfp}(\tau_P) = \tau_P \uparrow \omega, \quad \mathrm{lfp}(\tau_\Box) = \tau_\Box \uparrow \omega$$

where the ordinal powers, $\tau_P \uparrow \alpha$, $\tau_\Box \uparrow \alpha$, are defined in the usual way.

We are now in a position to introduce a fixpoint operator for M.

Definition $\tau(I) = \tau_\Box(\tau_P(I))$.

Thus the ordinal powers of τ that are relevant to our discussion are given by:

$$\tau \uparrow 0 = \tau_\Box(\tau_P \uparrow 0) = \tau_\Box(\varnothing) = \varnothing$$

$$\tau \uparrow \alpha = \tau(\tau \uparrow \alpha - 1) = \tau_\Box(\tau_P(\tau \uparrow \alpha - 1)), \text{ if } \alpha \text{ is a successor ordinal}$$

$$\tau \uparrow \alpha = \mathrm{lub}\{\tau \uparrow \beta \mid \beta < \alpha\}, \text{ if } \alpha \text{ is a limit ordinal.}$$

Since the operator composition of two continuous operators is again continuous, τ is continuous. Hence $\mathrm{lfp}\tau = \tau \uparrow \omega$. Assuming that M is the least fixpoint of τ, it follows that beginning with the empty Herbrand interpretation $\varnothing$ we can recover M after ω applications of the operator τ. This is in contrast to our earlier construction of M which began with the more complicated structure $H = \tau_P \uparrow \omega$.

We shall now verify that the least fixpoint of τ is indeed M. We will require a simple preliminary lemma.

Lemma *Let T be the equality theory of Section 1. Then T logically implies $\Box$.*

Proof: Let K be a Herbrand model over U_L such that $K \models T$. $K \models \forall x \forall y \forall u \forall v\ x = u \wedge u = v \wedge v = y \rightarrow x = y$. For, let $a = s$, $s = t$, $b = t \in K$. Then by transitivity $a = t \in K$; by symmetry $t = b \in K$. Thus by transitivity $a = b \in K$. Next, notice that $K \models \forall x \forall y\ x = y \wedge py \rightarrow px$, since this formula belongs to the theory T.

Theorem 5 $\mathrm{lfp}(\tau) = M$.

Proof: (1) $\mathrm{lfp}(\tau) \subseteq M$. Since τ is continuous $\mathrm{lfp}(\tau) = \tau \uparrow \omega$. Our argument proceeds by induction.

Base case. $\tau \uparrow 0 = \varnothing \subseteq M$.

Inductive step. Suppose $\tau \uparrow n \subseteq M$. We must show that $\tau \uparrow n + 1 \subseteq M$. We have

$$\tau \uparrow n + 1 = \tau(\tau \uparrow n).$$

By the inductive hypothesis,

$$\tau \uparrow n \subseteq M.$$

Since τ_P is monotonic, and M is model of $P \cup \{x = x\}$.

$$\tau_P(\tau \uparrow n) \subseteq \tau_P(M) \subseteq M.$$

By the monotonicity of $\tau_\Box$

$$\tau_\Box(\tau_P(\tau \uparrow n)) \subseteq \tau_\Box(M).$$

And by the Lemma $M \models \Box$, so that $\tau_\Box(M) \subseteq M$. Thus

$$\tau_\Box(\tau_P(\tau \uparrow n)) \subseteq M,$$

establishing the first inclusion.

(2) $M \subseteq \mathrm{lfp}(\tau)$. It suffices to show that

$$\mathrm{lfp}(\tau) \models \forall x \forall y \forall z\ x = y \wedge y = z \rightarrow x = z,$$

since $\mathrm{lfp}(\tau)$ is a model of all other axioms of the equality theory T and M is the minimal Herbrand model of $P \cup T$. So suppose $a = b$, $b = c \in \mathrm{lfp}(\tau)$. Then for some $n \geq 1$, $a = b$, $b = c \in \tau \uparrow n$. Notice if $b = c \in \tau \uparrow n$, then $c = b \in \tau \uparrow n$; also $b = b \in \tau \uparrow 1 \subseteq \tau \uparrow n$. Hence $a = b \in \tau \uparrow n + 1$, thus proving the reverse inclusion.

REFERENCES

[1] van Emden, M. H. and J. W. Lloyd, "A logical reconstruction of Prolog II," *The Journal of Logic Programming*, vol. 1 (1984), pp. 143–150.

[2] Lloyd, J. W., *Foundations of Logic Programming*, Springer-Verlag, New York, 1984.

Department of Philosophy
The University of Western Ontario
London, Ontario, Canada N6A 3K7