

---

## ANNOTATIONS ON THE CONSISTENCY OF THE CLOSED WORLD ASSUMPTION

---

**GERHARD JÄGER**

---

- ▷ The treatment of negation and negative information in a logic programming environment has turned out to be a major problem. We introduce a relativized version of Reiter's closed world assumption and study it from a logical point of view. In particular, we look at the questions of consistency and conservative extension. ◁
- 

### 0. INTRODUCTION

The treatment of negation and negative information in a logic programming environment has turned out to be a major problem. A starting point for all discussions is the programming language PROLOG, which is fairly efficient and realizes some ideas of logic programming. However, the expressive power of (pure) PROLOG is very limited, and it is not possible to treat negation naturally, as would be desirable for such an elementary concept.

A typical PROLOG program consists of a finite set of so called definite Horn clauses, i.e. formulas of the form

$$A_1 \& A_2 \& \dots \& A_k \rightarrow B$$

where  $k$  is larger than or equal to 0, and  $B, A_1, \dots, A_k$  are atomic formulas. Definite Horn formulas state what is true provided that something else is true; they do not state what is false.

It is clear that every PROLOG program  $\pi$  has a model and that  $\pi \not\models \neg B$  for any atomic  $B$ . If one chooses the declarative semantics and identifies the meaning of a program  $\pi$  with the set of its logical consequences,

$$\text{meaning}(\pi) := \{ A : \pi \vdash A \},$$

then  $\pi$  does not reflect negative information.

---

*Address correspondence to* Gerhard Jäger, Institut für Informatik und angewandte Mathematik, Universität Bern, CH-3012 Bern, Switzerland.

Received October 1987; accepted June 1988.

THE JOURNAL OF LOGIC PROGRAMMING

©Elsevier Science Publishing Co., Inc., 1990  
655 Avenue of the Americas, New York, NY 10010

0743-1066/90/\$3.50

This deficiency in the ability to express negation can be healed by adding new (meta) concepts to the plain deductive approach to logic programming. Clark's negation as failure rule and his completion of theories [5], Reiter's closed world assumption [18], and McCarthy's circumscription [12] are probably the most familiar ones, but there are more, and the interested reader should consult Shepherdson's excellent survey article [23] on negation in logic programming for further information.

Reiter's closed world assumption (CWA) is the mathematically most rigid form for introducing negation. It is motivated by applications in data base theory and proceeds on the assumption that a data base  $T$  contains all positive information and that any positive ground literal which is not implied by  $T$  is false. Although based on a very elementary idea, the use of the closed world assumption is not unproblematic at all, since it often leads to inconsistencies.

In order to gain more flexibility, we introduce a relativized version of Reiter's original notion. Given a first order theory  $T$  and a sequence  $\underline{P} = P_1, \dots, P_n$  of (unary) relation symbols in the language of  $T$ , we write  $CWA_{\underline{P}}(T)$  for the theory  $T$  plus

$$\bigcup_{i=1}^n \{ \neg P_i(a) : a \text{ a ground term \& } T \not\vdash P_i(a) \}$$

as additional axioms. The idea behind this approach is the following: (1)  $T$  is a formal representation of our (present) knowledge. (2) It can be split into stable facts SF which do not refer to the relations  $\underline{P}$ , and data DB about  $\underline{P}$ :  $T = \text{SF} + \text{DB}$ . (3) The meaning of the constants, function symbols, and relation symbols different from  $\underline{P}$  is reflected by SF; the meaning of the relation symbols  $\underline{P}$  is given by SF, DB, and the closed world assumption with respect to  $\underline{P}$ .

In this paper we study logical and mathematical aspects of the relativized closed world assumption. It is our aim to provide directions for the legitimate use of this important concept and to gain a better understanding of its meaning. In particular, we develop criteria for  $CWA_{\underline{P}}(T)$  to be consistent and a conservative extension of  $T$  for suitable classes of formulas.

Section 1 describes the general scenario and presents some basic definitions. Section 2 is devoted to a semantic approach to the closed world assumption. We introduce the notions of primary model (Definition 2.5) and weak categoricity (Definition 2.11) modulo  $\underline{P}$ , and use them to prove some consistency results for the CWA. In addition, we turn to the domain closure property, a version of completeness, and a conservative extension statement. We conclude, in Section 3, by discussing the closed world assumption in connection with the important classes of inductive and  $\Sigma$  inductive data bases. We study their intended models and apply results of the previous section to gain consistency and conservative extension statements for the CWA.

We hope that the results of this paper contribute to pointing out the usefulness and the limitations of the closed world assumption.

There is related work which should be mentioned: Chapter 3 of Lloyd's textbook on logic programming [11] is dedicated to negation. Reiter [18] introduces the closed world assumption and provides a lot of motivation for the use of this concept in data base theory. Three papers [21–23] of Shepherdson's address the treatment of negation in logic programming in general and establish important results concerning

the connections between the closed world assumption and alternative approaches. Lifschitz [10] treats the relationship between circumscription and the closed world assumption. Makowsky [13] is interested in the closed world assumption in the context of his generic (term) models. Minker [14] and Henschen and Park [8] formulate a so-called generalized closed world assumption which is able to take care of some forms of indefinite knowledge. In all these articles the question of the consistency of the closed world assumption plays a more or less important role. In this paper we focus on this essential logical property and try to provide some sound theoretical foundations for the safe use of the closed world assumption in logic programming, logical data bases, information processing, and the like.

### 1. GENERAL SCENARIO

We will deal with a countable first order language  $L$  with equality and an arbitrary number of function and relation symbols. The terms  $a, b, c, a_1, b_1, c_1, \dots$  and formulas  $A, B, C, A_1, B_1, C_1, \dots$  of  $L$  are defined as usual; formulas and terms without free variables are called *ground*. We write  $\underline{a}$  for a finite string  $a_1, \dots, a_n$  of  $L$  terms and use the notation  $A[\underline{x}]$  to indicate that all free variables of  $A$  come from the list  $\underline{x}$ ;  $A(\underline{x})$  may contain other free variables besides  $\underline{x}$ . An  $L$  theory is a (possibly infinite) collection of ground  $L$  formulas. By  $T \vdash A$  we express that the formula  $A$  can be deduced from the theory  $T$  by the usual axioms and rules of predicate logic with equality. An  $L$  theory  $T$  is *inconsistent* if every  $L$  formula is deducible from  $T$ ; otherwise  $T$  is *consistent*.

The collection of *Horn clauses* consists of all  $L$  formulas of the form

- (i)  $A$ ,
- (ii)  $\neg B_1 \vee \dots \vee \neg B_n \vee A$ ,
- (iii)  $\neg B_1 \vee \dots \vee \neg B_n$

where  $A$  and  $B_1, \dots, B_n$  are atomic formulas; Horn clauses of the form (i) and (ii) are called *definite* Horn clauses. If  $A$  is a formula, then the universal closure of  $A$  is the formula obtained by adding a universal quantifier for every variable having a free occurrence in  $A$ . A *logic program* is a finite collection of universal closures of definite Horn clauses.

In addition,  $L$  is supposed to contain a sequence  $\underline{P} = P_1, \dots, P_n$  of relation symbols, which we assume to be unary in order to keep the notation as simple as possible; the extension of our results to relation symbols  $\underline{P}$  of arbitrary arities is straightforward.  $L_0$  is the sublanguage of  $L$  without occurrences of the relation symbols  $\underline{P}$ ; QF denotes the class of quantifier-free  $L$  formulas.

Reiter's *closed world assumption* (CWA) [18] will be used in this paper in the following form: For every  $L$  theory  $T$  we define the set

$$\text{NEG}_{\underline{P}}(T) := \bigcup_{i=1}^n \{ \neg P_i(a) : a \text{ ground} \ \& \ T \not\vdash P_i(a) \}$$

of all ground formulas  $\neg P_i(a)$  which cannot be proved in  $T$ . Then we put

$$\text{CWA}_{\underline{P}}(T) := T + \text{NEG}_{\underline{P}}(T).$$

Similar forms of a relativized closed world assumption have been introduced in

Genesereth and Nilsson [6] and Jäger [9]. Working with the closed world assumption means working in two different levels. In the first level one has the theory  $T$  and checks whether certain atomic formulas  $P_i(a)$  are provable or not. In the second level  $T$  is extended by negations of some nonprovable atoms and then the usual derivation procedure is initiated. Formally these two levels are reflected by the fact that provability with the closed world assumption is  $\Sigma_2^0$ , and not  $\Sigma_1^0$  (recursively enumerable) as usual.

*Example 1.1.* Let  $S_5$  be a theory which formalizes that its universe consists of exactly five elements  $a_1, \dots, a_5$ . Then define for  $1 \leq m \leq 5$

$$T[m] := S_5 + \{P(a_i) : 1 \leq i \leq m\}.$$

With the help of the closed world assumption we can easily show that

$$\text{CWA}_P(T[m]) \vdash \forall x (P(x) \leftrightarrow x = a_1 \vee \dots \vee x = a_m).$$

REMARK 1.2. The closed world assumption is a nonmonotonic concept in the sense that  $T \subset T'$  does not imply that the theorems of  $\text{CWA}_P(T)$  are theorems of  $\text{CWA}_P(T')$  as well. To give an example, consider the theories  $T[2]$  and  $T[3]$  above. Then we have:

- (i)  $T[2] \subset T[3]$ ;
- (ii)  $\text{CWA}_P(T[2]) \vdash \neg P(a_3)$ ;
- (iii)  $\text{CWA}_P(T[3]) \not\vdash \neg P(a_3)$ .

The closed world assumption is a very handy and well-motivated concept as long as elementary assertions about  $P$  are considered. Then the use of the CWA causes no problems, and the meaning of  $\text{CWA}_P(T)$  is perfectly clear. However, as soon as more complex situations are taken into account one has to be careful. Consider the following example.

*Example 1.3.* Let  $T$  be the theory  $S_5 + \{P(a_1) \vee P(a_2)\}$ . Then neither  $P(a_1)$  nor  $P(a_2)$  is provable in  $T$  such that  $\text{CWA}_P(T) \vdash \neg P(a_1) \ \& \ \neg P(a_2)$ . This yields the inconsistency of  $\text{CWA}_P(T)$ , since  $\text{CWA}_P(T)$  proves  $P(a_1) \vee P(a_2)$  as well.

This example tells us that  $\text{CWA}_P(T)$  may be inconsistent even if  $T$  is consistent. But in order to use the closed world assumption properly one has to be sure that  $\text{CWA}_P(T)$  is consistent. A complete characterization of those theories  $T$  which give rise to a consistent  $\text{CWA}_P(T)$  has not been found yet, and it seems very unlikely that there exists a natural one. However, there are some partial results which cover most of the relevant cases.

## 2. SEMANTIC APPROACH TO THE CONSISTENCY OF THE CWA

In this section we discuss the consistency of the closed world assumption from a semantic point of view by using modifications of well-understood concepts of traditional model theory. To start with, we repeat some standard notions and introduce variations thereof.

An  $L$  structure is a pair  $\mathbf{M} = (M, I)$  consisting of a nonempty set  $M$  and a mapping  $I$  which assigns a function  $I(f): M^k \rightarrow M$  to every  $k$ -ary function symbol

$f$  and a set  $I(R) \subset M^k$  to every  $k$ -ary relation symbol  $R$  of  $L$ . In the following we often write  $f^{\mathbf{M}}$  and  $R^{\mathbf{M}}$  instead of  $I(f)$  and  $I(R)$ , and denote the universe of  $\mathbf{M}$  by  $|\mathbf{M}|$ . The  $L$  structure  $\mathbf{M}$  is called *countable* if the cardinality of  $|\mathbf{M}|$  is less than or equal to  $\omega$ .

The validity of ground formula  $A$  in the structure  $\mathbf{M}$  is defined as usual and denoted by  $\mathbf{M} \models A$ .  $\mathbf{M}$  is a *model* of the theory  $T$  if  $\mathbf{M} \models A$  holds for all  $A$  in  $T$ . Finally we write  $T \models A$  if the ground formula  $A$  is valid in all models of the theory  $T$ . Gödel's completeness theorem then implies that  $T \vdash A$  if and only if  $T \models A$ .

Let  $\mathbf{M}$  and  $\mathbf{N}$  be two  $L$  structures. We call  $\mathbf{M}$  *equal to  $\mathbf{N}$  modulo  $\underline{P}$* , in symbols  $\mathbf{M} = \mathbf{N}$  modulo  $\underline{P}$ , if  $f^{\mathbf{M}} = f^{\mathbf{N}}$  for all function symbols  $f$  and  $R^{\mathbf{M}} = R^{\mathbf{N}}$  for all relation symbols  $R$  which do not occur in  $\underline{P}$ . An *isomorphism from  $\mathbf{M}$  to  $\mathbf{N}$  modulo  $\underline{P}$*  is a bijective mapping  $H$  from  $|\mathbf{M}|$  to  $|\mathbf{N}|$  which satisfies

- (i)  $f^{\mathbf{N}}(H(m_1), \dots, H(m_k)) = H(f^{\mathbf{M}}(m_1, \dots, m_k))$ ,
- (ii)  $\langle H(m_1), \dots, H(m_k) \rangle \in R^{\mathbf{N}} \Leftrightarrow \langle m_1, \dots, m_k \rangle \in R^{\mathbf{M}}$

for all  $m_1, \dots, m_k \in |\mathbf{M}|$ , all function symbols  $f$ , and all relation symbols  $R$  which do not occur in  $\underline{P}$ .  $\mathbf{M}$  and  $\mathbf{N}$  are *isomorphic modulo  $\underline{P}$*  if there exists an isomorphism from  $\mathbf{M}$  to  $\mathbf{N}$  modulo  $\underline{P}$ ; in this case we write  $\mathbf{M} \equiv \mathbf{N}$  modulo  $\underline{P}$ .  $\mathbf{M}$  is a *substructure of  $\mathbf{N}$  modulo  $\underline{P}$*  if  $|\mathbf{M}|$  is a subset of  $|\mathbf{N}|$  and (i) and (ii) are satisfied for  $H$  replaced by the identity mapping restricted to  $|\mathbf{M}|$ . If we choose  $\underline{P}$  to be the empty sequence, we obtain the usual definitions of isomorphism and substructure.

*Lemma 2.1.* *Let  $\mathbf{M}$  and  $\mathbf{N}$  be  $L$  structures and  $H$  an isomorphism from  $\mathbf{M}$  to  $\mathbf{N}$ . Then we have for all  $L$  formulas  $A[x_1, \dots, x_k]$  and  $m_1, \dots, m_k \in |\mathbf{M}|$*

$$\mathbf{M} \models A[m_1, \dots, m_k] \Leftrightarrow \mathbf{N} \models A[H(m_1), \dots, H(m_k)].$$

*Lemma 2.2.* *Let  $\mathbf{M}$  and  $\mathbf{N}$  be isomorphic  $L$  structures modulo  $\underline{P}$ . Then there exists an  $L$  structure  $\mathbf{M}'$  which is equal to  $\mathbf{M}$  modulo  $\underline{P}$  and isomorphic to  $\mathbf{N}$ .*

PROOF. Let  $H$  be an isomorphism from  $\mathbf{M}$  to  $\mathbf{N}$  modulo  $\underline{P}$ . Then we define  $\mathbf{M}'$  by

- (i)  $|\mathbf{M}'| := |\mathbf{M}|$ ,
- (ii)  $f^{\mathbf{M}'} := f^{\mathbf{M}}$ ,
- (iii)  $R^{\mathbf{M}'} := R^{\mathbf{M}}$

for all function symbols  $f$  and relation symbols  $R$  different from  $\underline{P} = P_1, \dots, P_n$ . In addition, we define for all  $P_i, \dots, P_n$

- (iv)  $P_i^{\mathbf{M}'} := \{m \in |\mathbf{M}'| : H(m) \in P_i^{\mathbf{N}}\}$ .

It is obvious that  $\mathbf{M}'$  is equal to  $\mathbf{M}$  modulo  $\underline{P}$  and isomorphic to  $\mathbf{N}$ .  $\square$

*Definition 2.3 (Inductive definition of  $\Sigma$  and  $\Pi$  formulas of  $L$ ).*

1. Every QF formula is a  $\Sigma$  and  $\Pi$  formula.
2. If  $A$  is a  $\Sigma$  formula [ $\Pi$  formula], then  $\neg A$  is a  $\Pi$  formula [ $\Sigma$  formula].
3. If  $A$  and  $B$  are  $\Sigma$  formulas [ $\Pi$  formulas], then  $A \ \& \ B$  and  $A \ \vee \ B$  are  $\Sigma$  formulas [ $\Pi$  formulas].

4. If  $A$  is a  $\Sigma$  formula [ $\Pi$  formula] and  $B$  a  $\Pi$  formula [ $\Sigma$  formula], then  $A \rightarrow B$  is a  $\Pi$  formula [ $\Sigma$  formula].
5. If  $A(x)$  is a  $\Sigma$  formula, then  $\exists x A(x)$  is a  $\Sigma$  formula.
6. If  $A(x)$  is a  $\Pi$  formula, then  $\forall x A(x)$  is a  $\Pi$  formula.

*Lemma 2.4.* Let  $\mathbf{M}$  be a substructure of  $\mathbf{N}$ . Then we have for all QF formulas  $A[\underline{x}]$ ,  $\Sigma$  formulas  $B[\underline{x}]$ ,  $\Pi$  formulas  $C[\underline{x}]$ , and  $\underline{m} \in |\mathbf{M}|$

- (a)  $\mathbf{M} \models A[\underline{m}] \Leftrightarrow \mathbf{N} \models A[\underline{m}]$ ;
- (b)  $\mathbf{M} \models B[\underline{m}] \Rightarrow \mathbf{N} \models B[\underline{m}]$ ;
- (c)  $\mathbf{N} \models C[\underline{m}] \Rightarrow \mathbf{M} \models C[\underline{m}]$ .

The proof of this lemma is standard and can be found in any textbook on mathematical logic or model theory, e.g. in [4,24]. The following definition of primary model is a first step in direction of the consistency of the CWA. The notion of primary model introduced here resembles the notions of prime model (cf. [4]) and initial model (cf. [7]) but is not equivalent to either of these. In the case of prime models substructure is replaced by elementary substructure, and in the case of initial models uniqueness of the substructure is required.

*Definition 2.5.* Let  $T$  be an  $L$  theory, and  $\mathbf{M}$  an  $L$  structure.  $\mathbf{M}$  is a *primary model of  $T$  modulo  $\underline{P}$*  if  $\mathbf{M}$  is a model of  $T$  and every model of  $T$  has a substructure which is isomorphic to  $\mathbf{M}$  modulo  $\underline{P}$ .  $\mathbf{M}$  is a *primary model of  $T$*  if  $\mathbf{M}$  is a primary model of  $T$  modulo the empty sequence.

*Example 2.6.*

1. Peano arithmetic PA is a theory which has a primary model, namely the standard structure of the natural numbers.
2. Let  $a$  be a ground term and  $R$  a unary relation symbol. Then the theory  $\{\neg R(a) \vee R(a)\}$  does not possess a primary model.

*Theorem 2.7.* If the  $L$  structure  $\mathbf{M}$  is a primary model of the  $L$  theory  $T$ , then  $\mathbf{M}$  is a model of  $\text{CWA}_{\underline{P}}(T)$  as well.

**PROOF.** Let  $\mathbf{M}$  be a primary model of  $T$ , and assume that  $T \not\models P_i(a)$  for a ground term  $a$  and a relation symbol  $P_i$  from  $\underline{P}$ . In view of Gödel's completeness result there exists a model  $\mathbf{N}$  of  $T$  such that  $\mathbf{N} \models \neg P_i(a)$ . Since  $\mathbf{M}$  is a primary model of  $T$ , there exists a substructure  $\mathbf{N}'$  of  $\mathbf{N}$  which is isomorphic to  $\mathbf{M}$ . By Lemma 2.4 we conclude  $\mathbf{N}' \models \neg P_i(a)$ , and therefore  $\mathbf{M} \models \neg P_i(a)$  by Lemma 2.1. Hence we have shown that  $\mathbf{M} \models \text{NEG}_{\underline{P}}(T)$ , i.e.  $\mathbf{M} \models \text{CWA}_{\underline{P}}(T)$ .  $\square$

The applicability of this theorem is very restricted. Example 2.6 above shows that even collections of (definite) Horn clauses do not necessarily possess primary models. Therefore Theorem 2.7 does not validate the use of the closed world assumption, for example, in the context of logic programming.

In order to obtain deeper results and as motivation for the following, it is helpful to recall an important idea of logic programming: Every logic program  $\pi$  has the intersection property, i.e., if  $\mathbf{M}$  and  $\mathbf{N}$  are Herbrand models of  $\pi$ , then their intersection is a model of  $\pi$  as well; hence  $\pi$  possesses a least Herbrand model (the intersection of all of its Herbrand models), and the meaning of  $\pi$  can be identified with what is valid in this least Herbrand model.

Carried over to the closed world assumption with respect to  $\underline{P}$ , this means that we have to consider the intersection of  $L$  structures in the interpretations of  $\underline{P}$ . For two  $L$  structures  $\mathbf{M}$  and  $\mathbf{N}$  which are equal modulo  $\underline{P}$ , we introduce a new  $L$  structure  $\mathbf{M} \cap \mathbf{N}$  by

- (i)  $|\mathbf{M} \cap \mathbf{N}| := |\mathbf{M}| = |\mathbf{N}|$ ,
- (ii)  $f^{\mathbf{M} \cap \mathbf{N}} := f^{\mathbf{M}} = f^{\mathbf{N}}$ ,
- (iii)  $R^{\mathbf{M} \cap \mathbf{N}} := R^{\mathbf{M}} \cap R^{\mathbf{N}}$

for all function symbols  $f$  and relation symbols  $R$  of  $L$ . The new structure  $\mathbf{M} \cap \mathbf{N}$  differs from  $\mathbf{M}$  and  $\mathbf{N}$  only in the interpretation of the relation symbols  $\underline{P}$ .

*Definition 2.8.* The  $L$  theory  $T$  has the  $\underline{P}$  intersection property if for all models  $\mathbf{M}$  and  $\mathbf{N}$  of  $T$

$$\mathbf{M} = \mathbf{N} \text{ modulo } \underline{P} \Rightarrow \mathbf{M} \cap \mathbf{N} \models T.$$

Similar concepts, but for completely different reasons, have been introduced into model theory by Rabin [16,17] and Robinson [20]; Makowsky [13] studies the intersection property in connection with generic (term) models and Horn logic.

Typical theories with the  $\underline{P}$  intersection property will follow later. Now we consider examples which show that the  $\underline{P}$  intersection property of an  $L$  theory  $T$  does not guarantee the consistency of  $\text{CWA}_{\underline{P}}(T)$ . These examples depend on the existence of an  $L_0$ -incomplete theory, but it is obvious that  $L_0$ -incomplete theories exist for nearly all languages  $L$  and  $L_0$ .

*Definition 2.9.* An  $L$  theory  $T$  is called  $L_0$ -complete if  $T$  is consistent and  $T \vdash A$  or  $T \vdash \neg A$  for every ground formula  $A$  of the sublanguage  $L_0$  of  $L$ .  $T$  is  $L_0$ -incomplete if there exists a ground  $L_0$  formula  $A$  such that  $T$  proves neither  $A$  nor  $\neg A$ .

*Example 2.10.* Let SF be an incomplete  $L_0$  theory, and suppose that  $A$  is a ground  $L_0$  formula such that SF proves neither  $A$  nor  $\neg A$ .

1. For two ground terms  $a_1$  and  $a_2$  which are provably different we define

$$B(P) \quad := \Leftrightarrow \quad \forall x [(A \ \& \ x = a_1) \vee (\neg A \ \& \ x = a_2) \rightarrow P(x)],$$

$$T_1 := \text{SF} + \{B(P)\}.$$

$T_1$  has the  $\underline{P}$  intersection property. It is also clear that  $T$  proves  $P(a_1) \vee P(a_2)$  but does not prove  $P(a_i)$  for  $i = 1, 2$ . Using the closed world assumption, we conclude that  $P(a_1) \vee P(a_2)$  and  $\neg P(a_1) \ \& \ \neg P(a_2)$  are theorems of  $\text{CWA}_{\underline{P}}(T_1)$ . Hence  $\text{CWA}_{\underline{P}}(T_1)$  is inconsistent.

2. If only a single ground term  $a$  is available, we define

$$C_1(P_1) \quad :\Leftrightarrow \quad \forall x [A \ \& \ x = a \rightarrow P_1(x)],$$

$$C_2(P_2) \quad :\Leftrightarrow \quad \forall x [\neg A \ \& \ x = a \rightarrow P_2(x)],$$

$$T_2 := \text{SF} + \{C_1(P_1), C_2(P_2)\}.$$

Following the pattern of the previous argument, we conclude that  $T_2$  has the  $P_1, P_2$  intersection property and  $\text{CWA}_{P_1, P_2}(T_2)$  is inconsistent.

Our next goal is to find criteria for the  $\underline{P}$ -free part of  $T$  which, together with the  $\underline{P}$  intersection property, ensures the consistency of  $\text{CWA}_{\underline{P}}(T)$ . To achieve this, we need some more terminology.

*Definition 2.11.* The  $L$  theory  $T$  is called *weakly categorical modulo  $\underline{P}$*  if  $T$  has a countable model and any two countable models of  $T$  are isomorphic modulo  $\underline{P}$ .

This definition means that the class of all countable models of a weakly categorical theory  $T$  has, up to isomorphism modulo  $\underline{P}$ , exactly one element. It does not say whether there are uncountable models of  $T$  and how many. The concept “weakly categorical” and the familiar “ $\omega$ -categorical” (cf. e.g. [4]) are related but not identical.

REMARK 2.12.

1. If  $T$  is weakly categorical modulo  $P$ , then  $T$  has a primary model modulo  $\underline{P}$ .
2. If  $T$  is weakly categorical modulo  $\underline{P}$ , then  $T$  is  $L_0$ -complete.

Uncountable structures do not really matter in computer science, and so we ignore them for a moment. Then one can think of a weakly categorical theory modulo  $\underline{P}$  as a theory which provides enough information to pin down its universe and the meaning of all function and relation symbols different from  $\underline{P}$ .

Weak categoricity is a very strong assumption if one deals with theories which have an infinitary domain. Typical (mathematical) examples are atomless Boolean algebras, the four complete theories of dense simple order, the theory of infinite Abelian groups with all elements of order  $p$  ( $p$  prime), and the theory of an equivalence relation with infinitely many equivalence classes and each class infinite.

For applications, weak categoricity is more important in connection with finite domains. Then there are many examples of weakly categorical theories modulo  $\underline{P}$ , and this notion is even equivalent to  $L_0$  completeness if we add Reiter’s domain closure property [19] that every object of the universe of  $T$  can be represented by a ground term available in  $L$ . Domain closure plays a role in finite data base theory, where it is very natural to assume that every object has a name.

*Definition 2.13.* The  $L$  theory  $T$  has the *domain closure property* (DCP) if there exist finitely many ground terms  $a_1, a_2, \dots, a_k$  such that  $T$  proves

$$\forall x (x = a_1 \vee \dots \vee x = a_k).$$



*Lemma 2.14.* *Let  $T$  be an  $L$  theory which satisfies the domain closure property. Then  $T$  is weakly categorical modulo  $\underline{P}$  if and only if  $T$  is  $L_0$ -complete.*

**PROOF.** It is obvious that weak categoricity modulo  $\underline{P}$  implies  $L_0$  completeness (cf. Remark 2.12 above). To prove the converse direction, assume that  $\mathbf{M}$  and  $\mathbf{N}$  are countable models of  $T$  and

$$T \vdash \forall x (x = a_1 \vee \dots \vee x = a_k) \tag{1}$$

for some ground  $L$  terms  $a_1, \dots, a_k$ . We have to show that  $\mathbf{M}$  is isomorphic to  $\mathbf{N}$  modulo  $\underline{P}$ . By (1)

$$[m] := \{ a \in \{ a_1, \dots, a_k \} : a^{\mathbf{M}} = m \}$$

is nonempty for every  $m \in |\mathbf{M}|$ . We choose a representative  $b_m$  of each set  $[m]$  and define a mapping  $H$  from  $|\mathbf{M}|$  to  $|\mathbf{N}|$  by

$$H(m) := b_m^{\mathbf{N}}$$

for all  $m \in |\mathbf{M}|$ . This mapping is independent of the choice of the representatives of the  $[m]$ , since  $T$  is  $L_0$ -complete.

It is also clear that  $H$  is 1-1 and onto: If  $H(m_1) = H(m_2)$ , there are  $b_1 \in [m_1]$  and  $b_2 \in [m_2]$  such that  $b_1^{\mathbf{N}} = b_2^{\mathbf{N}}$ . By the  $L_0$  completeness of  $T$ , this implies  $T \vdash b_1 = b_2$ , i.e.  $m_1 = b_1^{\mathbf{M}} = b_2^{\mathbf{M}} = m_2$ . Moreover, if  $n \in |\mathbf{N}|$ , we use (1) to find a term  $b \in \{ a_1, \dots, a_k \}$  such that  $n = b^{\mathbf{N}}$ . Thus  $n = H(b^{\mathbf{M}})$ .

Now let  $b_1 \in [m_1], \dots, b_r \in [m_r]$ . Then we have for all  $r$ -ary function symbols  $f$  and relation symbols  $R$  different from  $\underline{P}$  and all  $b \in [f^{\mathbf{M}}(m_1, \dots, m_r)]$

$$\begin{aligned} \langle m_1, \dots, m_r \rangle \in R^{\mathbf{M}} &\Leftrightarrow T \vdash R(b_1, \dots, b_r) \\ &\Leftrightarrow \langle H(m_1), \dots, H(m_r) \rangle \in R^{\mathbf{N}}, \end{aligned} \tag{2}$$

$$T \vdash b = f(b_1, \dots, b_r), \tag{3}$$

$$H(f^{\mathbf{M}}(m_1, \dots, m_r)) = b^{\mathbf{N}} = f^{\mathbf{N}}(b_1^{\mathbf{N}}, \dots, b_r^{\mathbf{N}}) = f^{\mathbf{N}}(H(m_1), \dots, H(m_r)). \tag{4}$$

This proves that  $T$  is weakly categorical modulo  $\underline{P}$ .  $\square$

Now we return to the closed world assumption. The next results give sufficient conditions for the consistency of  $CWA_{\underline{P}}(T)$ . Actually, we can do better and prove that, under suitable assumptions,  $CWA_{\underline{P}}(T)$  is a conservative extension of  $T$  for all  $\underline{P}$  positive ground formulas of  $L$ .

*Definition 2.15 (Inductive definition of  $\underline{P}$ -positive and  $\underline{P}$ -negative  $L$  formulas).*

1. All formulas of  $L_0$  are in  $\text{Pos}(\underline{P})$  and  $\text{Neg}(\underline{P})$ .
2. If  $A$  is in  $\text{Pos}(\underline{P})$  [ $\text{Neg}(\underline{P})$ ], then  $\neg A$  is in  $\text{Neg}(\underline{P})$  [ $\text{Pos}(\underline{P})$ ].
3. If  $A$  and  $B$  are in  $\text{Pos}(\underline{P})$  [ $\text{Neg}(\underline{P})$ ], then  $A \ \& \ B$  and  $A \ \vee \ B$  are in  $\text{Pos}(\underline{P})$  [ $\text{Neg}(\underline{P})$ ].
4. If  $A$  is in  $\text{Pos}(\underline{P})$  [ $\text{Neg}(\underline{P})$ ] and  $B$  in  $\text{Neg}(\underline{P})$  [ $\text{Neg}(\underline{P})$ ], then  $A \rightarrow B$  is in  $\text{Neg}(\underline{P})$  [ $\text{Pos}(\underline{P})$ ].
5. If  $A(x)$  is in  $\text{Pos}(\underline{P})$  [ $\text{Neg}(\underline{P})$ ], then  $\forall x A(x)$  and  $\exists x A(x)$  are in  $\text{Pos}(\underline{P})$  [ $\text{Neg}(\underline{P})$ ].

A formula is called  $\underline{P}$ -positive or positive in  $\underline{P}$  [ $\underline{P}$ -negative or negative in  $\underline{P}$ ] if it belongs to the class  $\text{Pos}(\underline{P})$  [ $\text{Neg}(\underline{P})$ ].

The significance of  $\underline{P}$ -positive and  $\underline{P}$ -negative formulas for the closed world assumption comes partly from some model-theoretic properties of these classes. The following notion of  $\underline{P}$  extension can be regarded as the semantic counterpart to the syntactic notion of  $\underline{P}$  positivity. Lemma 2.17 below makes these connections more precise.

*Definition 2.16.* Suppose that  $\mathbf{M}$  and  $\mathbf{N}$  are  $L$  structures,  $T$  is an  $L$  theory, and  $\underline{P} = P_1, \dots, P_n$ .

1.  $\mathbf{N}$  is said to be a  $\underline{P}$  extension of  $\mathbf{M}$ , in symbols  $\mathbf{M} \leq_{\underline{P}} \mathbf{N}$ , if the following conditions are satisfied:
  - (i)  $\mathbf{M} = \mathbf{N}$  modulo  $\underline{P}$ ,
  - (ii)  $P_i^{\mathbf{M}} \subset P_i^{\mathbf{N}}$  for  $1 \leq i \leq n$ .
2.  $\mathbf{N}$  is a proper  $\underline{P}$  extension of  $\mathbf{M}$  if  $\mathbf{N}$  is a  $\underline{P}$  extension of  $\mathbf{M}$  and  $\mathbf{M} \neq \mathbf{N}$ ; then we write  $\mathbf{M} <_{\underline{P}} \mathbf{N}$ .
3.  $\mathbf{M}$  is a  $\underline{P}$ -minimal model of  $T$  if  $\mathbf{M}$  is a model of  $T$  and there is no model  $\mathbf{N}$  of  $T$  such that  $\mathbf{N} <_{\underline{P}} \mathbf{M}$ .

*Lemma 2.17.* Let  $\mathbf{M}$  and  $\mathbf{N}$  be  $L$  structures,  $\mathbf{M} \leq_{\underline{P}} \mathbf{N}$ . If  $A[x]$  is a  $\underline{P}$ -positive formula,  $B[x]$  a  $\underline{P}$ -negative formula, and  $C[x]$  a formula of  $L_0$ , then we have for all  $\underline{m} \in |\mathbf{M}|$

- (a)  $\mathbf{M} \models A[\underline{m}] \Rightarrow \mathbf{N} \models A[\underline{m}]$ ,
- (b)  $\mathbf{N} \models B[\underline{m}] \Rightarrow \mathbf{M} \models B[\underline{m}]$ ,
- (c)  $\mathbf{M} \models C[\underline{m}] \Leftrightarrow \mathbf{N} \models C[\underline{m}]$ .

The proof of this lemma is trivial and proceeds by induction on  $A$  and on  $B$ . We will make use of these results in the proof of the following main theorem of this section and later on in connection with inductive data bases.

*Definition 2.18.* We assume that  $S_1$  and  $S_2$  are  $L$  theories and  $K$  is a class of ground  $L$  formulas.

1.  $S_2$  is an extension of  $S_1$  if every theorem of  $S_1$  is a theorem of  $S_2$ .
2.  $S_2$  is a conservative extension of  $S_1$  for  $K$  if  $S_2$  is an extension of  $S_1$  and  $S_2 \vdash A$  implies  $S_1 \vdash A$  for all  $A$  in  $K$ .

*Theorem 2.19.* Let  $T$  be an  $L$  theory such that

- (A1)  $T$  is weakly categorical modul  $\underline{P}$ ,
- (A2)  $T$  has the  $\underline{P}$  intersection property.

Then  $\text{CWA}_{\underline{P}}(T)$  is a conservative extension of  $T$  for all  $\underline{P}$ -positive ground formulas.

PROOF.  $\text{CWA}_{\underline{P}}(T)$  is clearly an extension of  $T$ . To prove conservativeness for  $\underline{P}$ -positive ground formulas, assume that

$$\text{CWA}_{\underline{P}}(T) \vdash A, \tag{1}$$

$$T \not\vdash A \tag{2}$$

for some  $\underline{P}$ -positive ground formula  $A$  of  $L$ . We will show that this leads to a contradiction. From (1) we conclude that there are finite sets  $J_1, \dots, J_k$  of ground terms with the properties

$$T \cup \bigcup_{i=1}^k \{ \neg P_i(a) : a \in J_i \} \vdash A, \tag{3}$$

$$T \not\vdash P_i(a) \tag{4}$$

for all  $1 \leq i \leq k$  and  $a \in J_i$ . Because of (2) and the Löwenheim-Skolem theorem we can choose a countable  $L$  structure  $\mathbf{M}$  which satisfies

$$\mathbf{M} \models T, \tag{5}$$

$$\mathbf{M} \models \neg A. \tag{6}$$

Again by the Löwenheim-Skolem theorem, now using (4), there exist countable  $L$  structures  $\mathbf{N}_{ia}$  for all  $1 \leq i \leq k$  and  $a \in J_i$  such that

$$\mathbf{N}_{ia} \models T, \tag{7}$$

$$\mathbf{N}_{ia} \models \neg P_i(a). \tag{8}$$

It follows from assumption (A1), (5), and (7) that all  $\mathbf{N}_{ia}$  are isomorphic to  $\mathbf{M}$  modulo  $\underline{P}$ . By Lemma 2.2, to each  $\mathbf{N}_{ia}$  there corresponds an  $L$  structure  $\mathbf{M}_{ia}$  such that

$$\mathbf{M}_{ia} = \mathbf{M} \text{ modulo } \underline{P}, \tag{9}$$

$$\mathbf{M}_{ia} \text{ is isomorphic to } \mathbf{N}_{ia}. \tag{10}$$

Hence (7), (8), (10), and Lemma 2.1 imply

$$\mathbf{M}_{ia} \models T, \tag{11}$$

$$\mathbf{M}_{ia} \models \neg P_i(a) \tag{12}$$

for all  $1 \leq i \leq k$  and  $a \in J_i$ . Now we define the  $L$  structure

$$\mathbf{M}^* := \mathbf{M} \cap \bigcap_{i=1}^k \bigcap_{a \in J_i} \mathbf{M}_{ia},$$

which is a model of  $T$  by assumption (A2) and (9):

$$\mathbf{M}^* \models T. \tag{13}$$

Note that  $\mathbf{M}$  and each  $\mathbf{M}_{ia}$  is a  $\underline{P}$  extension of  $\mathbf{M}^*$ . Thus we see by (6), (12), and Lemma 2.17 that

$$\mathbf{M}^* \models \neg A, \tag{14}$$

$$\mathbf{M}^* \models \neg P_i(a) \tag{15}$$

for all  $1 \leq i \leq k$  and  $A \in J_i$ . Finally, (3), (13), and (15) imply  $\mathbf{M}^* \models A$ , which contradicts (14). This completes the proof of Theorem 2.19.  $\square$

*Corollary 2.20.* *If the  $L$  theory  $T$  is weakly categorical modulo  $\underline{P}$  and has the  $\underline{P}$  intersection property, then  $\text{CWA}_{\underline{P}}(T)$  is consistent.*

**PROOF.** By the previous theorem,  $T$  is consistent if and only if  $\text{CWA}_{\underline{P}}(T)$  is consistent; on the other hand,  $T$  is consistent, since it is weakly categorical modulo  $\underline{P}$ .  $\square$

**REMARK 2.21 (OPEN QUESTIONS).**

1. The concept of weak categoricity modulo  $\underline{P}$  has been developed with Theorem 2.19 in mind. Although it turns out to be very natural and useful, at least in the context of finite data bases, it is sometimes rather complicated to check whether a theory is weakly categorical modulo  $\underline{P}$  or not. Therefore it would be desirable to find an alternative approach which serves the same purpose.
2. The definition of weakly categorical modulo  $\underline{P}$  is based on the notion of isomorphism. Is it possible to obtain similar results if we work with elementary equivalences or elementary embeddings instead?

### 3. INDUCTIVE DATA BASES

In this section we introduce the notions of inductive data base and  $\Sigma$  inductive data base. Using results of the previous section, we then study these concepts in connection with the closed world assumption, where we concentrate again on questions concerning conservative extension and (relative) consistency. The term “inductive data base” is motivated by the fact that the intended models of inductive data bases are constructed by inductive definition; see Lemma 3.6 for details.

If  $T$  is an  $L$  theory, we write  $T_-(\underline{P})$  for the set  $T \cap L_0$  and  $T_+(\underline{P})$  for the complement of  $T_-(\underline{P})$  in  $T$ . Hence  $T = T_-(\underline{P}) + T_+(\underline{P})$ , where  $T_-(\underline{P})$  contains the information which does not refer to  $\underline{P}$ , and  $T_+(\underline{P})$  collects all data about  $\underline{P}$ .

*Definition 3.1.* A nonempty and finite collection  $T$  of  $L$  sentences is called an *inductive data base for  $\underline{P} = P_1, \dots, P_n$*  if each element of  $T_+(\underline{P})$  is of the form

$$(ID.1) \quad \forall x (A[\underline{P}, x] \rightarrow P_i(x)) \text{ or}$$

$$(ID.2) \quad B(\underline{P})$$

where  $A[\underline{P}, x]$  belongs to  $\text{Pos}(\underline{P})$  and  $B(\underline{P})$  to  $\text{Neg}(\underline{P})$ . We call an inductive data base for  $\underline{P}$  *definite* if it does not contain sentences of the form (ID.2). An inductive data base for  $\underline{P}$  is said to be  $\Sigma$ -inductive if the formulas  $A[\underline{P}, x]$  in (ID.1) are  $\Sigma$  formulas and the formulas  $B(\underline{P})$  in (ID.2) are  $\Pi$  formulas.

*Lemma 3.2.* *Every inductive data base  $T$  for  $\underline{P}$  has the  $\underline{P}$  intersection property.*

**PROOF.** Let  $\mathbf{M}$  and  $\mathbf{N}$  be models of  $T$  which satisfy

$$\mathbf{M} = \mathbf{N} \text{ modulo } \underline{P}. \tag{1}$$

Then we have

$$\mathbf{M} \cap \mathbf{N} \leq_{\underline{P}} \mathbf{M}, \tag{2}$$

$$\mathbf{M} \cap \mathbf{N} \leq_{\underline{P}} \mathbf{N}, \tag{3}$$

so that Lemma 2.17 implies

$$\mathbf{M} \cap \mathbf{N} \models T_{-}(\underline{P}), \tag{4}$$

$$\mathbf{M} \cap \mathbf{N} \models B(\underline{P}) \tag{5}$$

for all sentences  $B(\underline{P})$  in  $T_{+}(\underline{P})$  which are  $\underline{P}$ -negative. Finally we choose a sentence  $\forall x(A[\underline{P}, x] \rightarrow P_i(x))$  of  $T_{+}(\underline{P})$  of the form (ID.1). In view of Lemma 2.17 we then obtain

$$\mathbf{M} \cap \mathbf{N} \models A[\underline{P}, m] \Rightarrow \mathbf{M} \models A[\underline{P}, m] \Rightarrow m \in P_i^{\mathbf{M}}, \tag{6a}$$

$$\mathbf{M} \cap \mathbf{N} \models A[\underline{P}, m] \Rightarrow \mathbf{N} \models A[\underline{P}, m] \Rightarrow m \in P_i^{\mathbf{N}}, \tag{6b}$$

for all  $m \in |\mathbf{M} \cap \mathbf{N}|$ . Hence it is proved that

$$\mathbf{M} \cap \mathbf{N} \models \forall x(A[\underline{P}, x] \rightarrow P_i(x)). \tag{7}$$

To sum up, (4), (5), and (7) imply that  $\mathbf{M} \cap \mathbf{N}$  is a model of  $T$ . Thus our lemma is proved.  $\square$

*Theorem 3.3. Let  $T$  be an  $L$  theory such that*

(B1)  *$T$  is weakly categorical modulo  $\underline{P}$ ,*

(B2)  *$T$  is an inductive data base for  $\underline{P}$ .*

*Then  $CWA_{\underline{P}}(T)$  is a conservative extension of  $T$  for all  $\underline{P}$ -positive ground formulas.*

This theorem is an immediate consequence of Theorem 2.19 and Lemma 3.2. As in Section 2, we obtain the following corollary.

*Corollary 3.4. If  $T$  is weakly categorical modulo  $\underline{P}$  and an inductive data base for  $\underline{P}$ , then  $CWA_{\underline{P}}(T)$  is consistent.*

As a digression, we will briefly address the question about the converse of Corollary 3.4: What can we say about an  $L_0$  theory SF, i.e. a theory which does not refer to  $\underline{P}$ , if we know that  $CWA_{\underline{P}}(\text{SF} + \text{DB})$  is consistent for all (definite) inductive data bases DB for  $\underline{P}$ ? First we consider two special cases:

1.  $L$  has no ground terms at all. In this case  $CWA_{\underline{P}}(\text{SF} + \text{DB}) = \text{SF} + \text{DB}$ .
2.  $L$  has exactly one ground term  $a$ , and DB is an inductive data base for one unary relation symbol  $P$ . If  $\text{SF} + \text{DB}$  proves  $P(a)$ , then  $CWA_{\underline{P}}(\text{SF} + \text{DB}) = \text{SF} + \text{DB}$ . Otherwise  $CWA_{\underline{P}}(\text{SF} + \text{DB})$  is consistent if and only if  $\text{SF} + \text{DB}$  is consistent.

The more complex situations have already been taken care of by Example 2.10 and Lemma 2.14:

*Theorem 3.5. Let SF be an  $L_0$  theory.*

- (a) *If there exist two ground terms  $a$  and  $b$  such that  $SF \vdash a \neq b$  and if  $CWA_P(SF + DB)$  is consistent for all definite data bases DB for  $P$ , then SF is  $L_0$ -complete.*
- (b) *If there exists a ground term  $a$  and if  $CWA_{P_1, P_2}(SF + DB)$  is consistent for all definite data bases DB for  $P_1, P_2$ , then SF is  $L_0$ -complete.*
- (c) *If we add the requirement that SF satisfies the domain closure property (DCP), then assertions (a) and (b) can be strengthened to SF being weakly categorical modulo  $P$  and modulo  $P_1, P_2$ , respectively.*

Now we turn to the explicit description of the intended model of an inductive data base for  $\underline{P}$ . The construction we use is very familiar from the theory of inductive definitions for the case of definite inductive data bases (cf. e.g. Moschovakis [15] or Barwise [2]) and is often called the “simultaneous induction lemma”. The extension to general inductive data bases is straightforward.

*Lemma and Definition 3.6. Let  $T$  be an inductive data base for  $\underline{P}$ , and  $\mathbf{M}$  a model of  $T$ . Then there exists an  $L$  structure  $IND_{\underline{P}}(\mathbf{M}, T)$  which has the following properties:*

- (a)  $IND_{\underline{P}}(\mathbf{M}, T) = \mathbf{M}$  modulo  $\underline{P}$ ,
- (b)  $IND_{\underline{P}}(\mathbf{M}, T) \models T$ ,
- (c) *if  $\mathbf{N}$  is an  $L$  structure such that*
  - (i)  $\mathbf{N} = \mathbf{M}$  modulo  $\underline{P}$ ,
  - (ii)  $\mathbf{N} \models T$ ,

*then  $\mathbf{N}$  is a  $\underline{P}$  extension of  $IND_{\underline{P}}(\mathbf{M}, T)$ .*

*We call  $IND_{\underline{P}}(\mathbf{M}, T)$  the inductive model of  $T$  over  $\mathbf{M}$ .*

**PROOF.** We assume that  $T$  is an inductive data base for  $\underline{P} = P_1, \dots, P_n$ , and  $\mathbf{M}$  an  $L$  structure such that  $\mathbf{M} \models T$ . In a first step we introduce sets  $I_1, \dots, I_n \subset |\mathbf{M}|$  which will be used as interpretations for the relation symbols  $P_1, \dots, P_n$  later. For every  $s = 1, \dots, n$  we choose  $K_s$  to be the (finite) collection of all formulas  $A[\underline{P}, x]$  such that  $\forall x(A[\underline{P}, x] \rightarrow P_s(x))$  belongs to  $T$ . By  $D$  we denote the set of all  $s \in \{1, \dots, n\}$  such that  $K_s \neq \emptyset$ . Finally we set

$$C_s[\underline{P}, x] := \bigvee_{A \in K_s} A[\underline{P}, x]$$

for all  $s \in D$ ; i.e.,  $C_s[\underline{P}, x]$  is the disjunction of the elements in  $K_s$ . Now use transfinite induction over the ordinals to define sets  $I_s^\alpha \subset |\mathbf{M}|$  for all  $s = 1, \dots, n$  and  $\alpha \in \text{On}$ :

$$I_s^{<\alpha} := \bigcup_{\xi < \alpha} I_s^\xi,$$

$$I_s^\alpha := \begin{cases} \{m \in |\mathbf{M}| : \mathbf{M} \models C_s[I_1^{<\alpha}, \dots, I_n^{<\alpha}, m]\} & \text{if } s \in D, \\ \emptyset & \text{if } s \notin D. \end{cases}$$

Since each formula  $C_s[\underline{P}, x]$  is  $\underline{P}$ -positive, we have  $I_s^\alpha \subset I_s^\beta$  for all  $\alpha < \beta$ . Hence there exists an ordinal  $\kappa$  of cardinality  $\leq \text{cardinality}(\mathbf{M})$  such that

$$I_s^\kappa = I_s^{<\kappa} = \bigcup_{\xi \in \text{On}} I_s^\xi. \quad (1)$$

Depending on  $\mathbf{M}$  and  $T$ , we now introduce an  $L$  structure  $\text{IND}_P(\mathbf{M}, T)$ , which we abbreviate as  $\text{IND}$ :

$$|\text{IND}| := |\mathbf{M}|, \quad (2)$$

$$P_s^{\text{IND}} := I_s^\kappa, \quad (3)$$

$$f^{\text{IND}} := f^{\mathbf{M}}, \quad (4)$$

$$R^{\text{IND}} := R^{\mathbf{M}} \quad (5)$$

for all function symbols  $f$  and relation symbols  $R$  which do not occur in  $\underline{P}$ . It is clear that  $\text{IND}$  is equal to  $\mathbf{M}$  modulo  $\underline{P}$ .

Now we prove (c). By induction on  $\alpha$  one can easily show that  $I_s^\alpha \subset P_s^{\mathbf{N}}$  for each  $L$  structure  $\mathbf{N}$  with the properties (i) and (ii). Hence  $P_s^{\text{IND}} \subset P_s^{\mathbf{N}}$  for these  $\mathbf{N}$ , and therefore

$$\text{IND} \leq_P \mathbf{N}. \quad (6)$$

In particular,  $\mathbf{M}$  is a  $\underline{P}$  extension of  $\text{IND}$ .

It remains to show (b). From Lemma 2.17 and (6), with  $\mathbf{N}$  replaced by  $\mathbf{M}$ , we can conclude that all  $\underline{P}$  negative sentences of  $T$  are valid in  $\text{IND}$ , i.e.

$$\mathbf{M} \models T \cap \text{Neg}(\underline{P}). \quad (7)$$

From (1) and the definition of  $P_s^{\text{IND}}$  we obtain that

$$\text{IND} \models \forall x (A[\underline{P}, x] \rightarrow P_s(x)) \quad (8)$$

for all sentences  $\forall x (A[\underline{P}, x] \rightarrow P_s(x))$  in  $T$  which have the form (ID.1). By (7) and (8)  $\text{IND}$  is a model of  $T$ .  $\square$

*Lemma 3.7. If  $T$  is an inductive data base for  $\underline{P}$ , and  $\mathbf{M}, \mathbf{N}$  are models of  $T$ , then*

$$\mathbf{M} \equiv \mathbf{N} \text{ modulo } \underline{P} \quad \Rightarrow \quad \text{IND}_P(\mathbf{M}, T) \equiv \text{IND}_P(\mathbf{N}, T).$$

**PROOF.** Let  $H$  be an isomorphism from  $\mathbf{M}$  to  $\mathbf{N}$  modulo  $\underline{P}$ . Over  $\mathbf{M}$  and  $\mathbf{N}$  we define the sets  $I_s^\alpha$  as in the previous lemma and denote them by  $I_s^\alpha(\mathbf{M})$  and  $I_s^\alpha(\mathbf{N})$ , respectively. Then we have for all  $\alpha \in \text{On}$  and  $m \in |\mathbf{M}|$

$$H(m) \in I_s^\alpha(\mathbf{N}) \quad \Leftrightarrow \quad m \in I_s^\alpha(\mathbf{M}).$$

This implies the assertion.  $\square$

*Theorem 3.8. Assume that*

(C1)  $T$  is an inductive data base for  $\underline{P}$ ,

(C2)  $T$  is weakly categorical modulo  $\underline{P}$ ,

(C3)  $\mathbf{M}$  is a countable model of  $T$ .

*Then  $\text{IND}_P(\mathbf{M}, T)$  is a  $\underline{P}$ -minimal model of  $\text{CWA}_P(T)$ .*

**PROOF.** Lemma 3.6 implies that

$$\text{IND}_{\underline{P}}(\mathbf{M}, T) \models T. \quad (1)$$

Now let  $a$  be a ground term such that

$$T \not\models P_i(a). \quad (2)$$

for some  $i = 1, \dots, n$ . Then there exists a countable  $L$  structure  $\mathbf{N}$  which has the properties

$$\mathbf{N} \models T, \quad (3)$$

$$\mathbf{N} \models \neg P_i(a). \quad (4)$$

By assumption (C2) we conclude that

$$\mathbf{M} \equiv \mathbf{N} \text{ modulo } \underline{P} \quad (5)$$

and hence, by the previous lemma,

$$\text{IND}_{\underline{P}}(\mathbf{M}, T) \equiv \text{IND}_{\underline{P}}(\mathbf{N}, T). \quad (6)$$

Lemma 3.6 and (3) also tell us that  $\mathbf{N}$  is an  $\underline{P}$  extension of  $\text{IND}_{\underline{P}}(\mathbf{N}, T)$ . Thus

$$\text{IND}_{\underline{P}}(\mathbf{N}, T) \models \neg P_i(a) \quad (7)$$

according to Lemma 2.17 and (4), whereas

$$\text{IND}_{\underline{P}}(\mathbf{M}, T) \models \neg P_i(a) \quad (8)$$

according to Lemma 2.1, (6), and (7). This, together with (1), proves that  $\text{IND}_{\underline{P}}(\mathbf{M}, T)$  is a model  $\text{CWA}_{\underline{P}}(T)$ ; the  $\underline{P}$  minimality follows from Lemma 3.6.  $\square$

In order to get rid of the strong requirement of weak categoricity modulo  $\underline{P}$ , we now pay special attention to  $\Sigma$ -inductive data bases. We will later see that in this case we get along with the weaker assumption of the existence of a primary model.

*Lemma 3.9. Assume that  $T$  is a  $\Sigma$ -inductive data base for  $\underline{P}$ , and  $\mathbf{M}$  a primary model of  $T$  modulo  $\underline{P}$ . Then  $\text{IND}_{\underline{P}}(\mathbf{M}, T)$  is a model of*

$$T + \{ \neg A : A \text{ is a } \underline{P}\text{-positive ground } \Sigma \text{ formula \& } T \not\models A \}.$$

*In particular,  $\text{IND}_{\underline{P}}(\mathbf{M}, T)$  is a model of  $\text{CWA}_{\underline{P}}(T)$ .*

**PROOF.** According to Lemma 3.6,  $\text{IND}_{\underline{P}}(\mathbf{M}, T)$  is a model of  $T$ . Now assume that  $T \not\models A$  for some  $\underline{P}$ -positive ground  $\Sigma$  formula  $A$ . Then there exists a countable  $L$  structure  $\mathbf{N}$  with the properties

$$\mathbf{N} \models T, \quad (1)$$

$$\mathbf{N} \models \neg A. \quad (2)$$

Since  $\mathbf{M}$  is a primary model of  $T$  modulo  $\underline{P}$ ,  $\mathbf{N}$  has a substructure  $\mathbf{N}'$  which is isomorphic to  $\mathbf{M}$  modulo  $\underline{P}$ :

$$\mathbf{N}' \equiv \mathbf{M} \text{ modulo } \underline{P}. \quad (3)$$

Hence  $\mathbf{N}' \models T_-(\underline{P})$ . But  $T_+(\underline{P}) \cup \{ \neg A \}$  is a set of  $\Pi$  formulas, and  $\mathbf{N}'$  is a model of this set by Lemma 2.4. Thus we have

$$\mathbf{N}' \models T \cup \{ \neg A \}. \quad (4)$$



Now consider the structure  $\text{IND}_{\underline{P}}(\mathbf{N}', T)$ . From Lemma 3.6 we see that  $\mathbf{N}'$  is a  $\underline{P}$  extension of  $\text{IND}_{\underline{P}}(\mathbf{N}, T)$ . Because of (4) and the  $\underline{P}$  negativity of  $\neg A$  we conclude first that

$$\text{IND}_{\underline{P}}(\mathbf{N}', T) \models \neg A \tag{5}$$

and use (3) and Lemma 3.7 to show that  $\text{IND}_{\underline{P}}(\mathbf{M}, T)$  is a model of  $\neg A$ . Thus our assertion is proved. It is a trivial consequence that  $\text{IND}_{\underline{P}}(\mathbf{M}, T)$  is a model of  $\text{CWA}_{\underline{P}}(T)$ .  $\square$

*Theorem 3.10. Assume that*

- (D1) *T has a primary model modulo  $\underline{P}$ ,*
- (D2) *T is a  $\Sigma$ -inductive data base for  $\underline{P}$ .*

*Then  $\text{CWA}_{\underline{P}}(T)$  is a conservative extension of T for all  $\underline{P}$ -positive ground  $\Sigma$  formulas.*

**PROOF.** First of all,  $\text{CWA}_{\underline{P}}(T)$  is an extension of  $T$  by definition. Now suppose that  $A$  is a  $\underline{P}$ -positive ground  $\Sigma$  formula which is provable in  $\text{CWA}_{\underline{P}}(T)$ . Hence, by Lemma 3.9,  $\text{IND}_{\underline{P}}(\mathbf{M}, T)$  is a model of  $A$ . And, again by Lemma 3.9, this implies that  $A$  is provable in  $T$ .  $\square$

*Corollary 3.11. If T has a primary model modulo  $\underline{P}$  and is a  $\Sigma$ -inductive data base for  $\underline{P}$ , then  $\text{CWA}_{\underline{P}}(T)$  is consistent.*

It is an easy exercise to show that the previous results can be extended to a sequence  $\underline{P} = P_1, \dots, P_n$  of relation symbols of arbitrary (finite) arities. One can either redo all arguments for this more general case or assume that each  $T$  contains sufficient coding possibilities to reduce an  $n$ -ary relation to a unary one. Anyway, using these generalizations, Corollary 3.11 has a nice consequence for logic programs.

*Corollary 3.12. If T is a logic program (or a consistent set of universal closures of Horn clauses) and  $\underline{P} = P_1, \dots, P_n$  the list of all relation symbols which occur in T, then  $\text{CWA}_{\underline{P}}(T)$  is consistent; i.e., the use of the closed world assumption is justified in connection with logic programs (and consistent sets of universal closures of Horn clauses).*

**PROOF.** Let  $T^*$  be the theory

$$T + \{ a \neq b : a, b \text{ ground} \ \& \ T \not\models a = b \}.$$

By a proof-theoretic argument one can show that

$$T \vdash A \iff T^* \vdash A$$

for all atomic sentences which are positive in  $\underline{P}$  and do not contain the equality symbol. This implies that

$$\text{CWA}_{\underline{P}}(T) \subset \text{CWA}_{\underline{P}}(T^*).$$

We first observe that  $T^*$  is a  $\Sigma$ -inductive data base for  $\underline{P}$ . In view of Corollary 3.11, it only remains to show that  $T^*$  has a primary model modulo  $\underline{P}$ . In order to prove

this, we first define a binary relation  $\sim$  on the ground terms of  $L$ ; we set

$$a \sim b \quad :\Leftrightarrow \quad T^* \vdash a = b.$$

Clearly  $\sim$  is an equivalence relation on the set of all ground terms. The equivalence class of a ground term  $a$  with respect to  $\sim$  is designated by  $[a]_{\sim}$ . We define an  $L$  structure  $\mathbf{M}$  by setting

- (i)  $|\mathbf{M}| := \{[a]_{\sim} : a \text{ is a ground term of } L\}$ ,
- (ii)  $R^{\mathbf{M}} := |\mathbf{M}|$ ,
- (iii)  $f^{\mathbf{M}}([a_1]_{\sim}, \dots, [a_k]_{\sim}) := [f(a_1, \dots, a_k)]_{\sim}$

for all relation symbols  $R$ , function symbols  $f$ , and ground terms  $a_1, \dots, a_k$ . Each function  $f^{\mathbf{M}}$  is well defined by (iii), since  $\sim$  is an equivalence relation. It is easy to check that  $\mathbf{M}$  is a primary model of  $T^*$  modulo  $\underline{P}$  provided that  $T$  is a logic program.

If  $T$  is a consistent set of universal closures of Horn clauses, then there exists a model  $\mathbf{N}$  of  $T^*$ . In this case the structure  $\mathbf{M}$  is defined as before, but with (ii) replaced by

$$(ii') \quad R^{\mathbf{M}} := \{ \langle [a_1]_{\sim}, \dots, [a_k]_{\sim} \rangle : a_1, \dots, a_k \text{ ground} \ \& \ \mathbf{N} \models R(a_1, \dots, a_k) \}.$$

The new  $\mathbf{M}$  is a model of  $T^*$ , and therefore also a primary model of  $T^*$  modulo  $\underline{P}$ , since  $\underline{P}$  is the list of all relation symbols which occur in  $T^*$ .  $\square$

Reiter's proof of a similar result in [18] made use of ad hoc methods depending on unit refutation. An extension and generalization of his theorem have long been sought and are now achieved here. It is gratifying that they take their place within a systematic framework.

#### 4. FINAL REMARKS

We end this paper with some comments concerning possible extensions of and generalizations of the notions and results presented so far.

1. Let  $\Gamma$  be a collection of ground  $L$  formulas, and  $T$  an  $L$  theory. The closed world of  $T$  with respect to  $\Gamma$  is then defined to be the theory

$$\text{CWA}_{\Gamma}(T) := T + \{ \neg A : A \in \Gamma \ \& \ T \not\models A \}.$$

This relativized closed world assumption is a natural extension of the version studied above. It would be interesting to analyze it for various collections of formulas  $\Gamma$  and to see whether similar results can be achieved.

2. Using techniques of theories of iterated inductive definitions (e.g. [3]), one can easily introduce the notion of iterated (or stratified) inductive data base. It seems obvious that the results of this paper are useful for these data bases as well and that there is a close connection to the stratified programs of Apt, Blair, and Walker [1].

## REFERENCES

1. Apt, K. R., Blair, H., and Walker, A., Towards a Theory of Declarative Knowledge, in: J. Minker (ed.), *Foundations of Deductive Databases and Logic Programming*, Morgan Kaufmann Publishers, Los Altos, 1988.
2. Barwise, J., *Admissible Sets and Structures*, Springer-Verlag, New York, 1975.
3. Buchholz, W., Feferman, S., Pohlers, W., and Sieg, W., *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies*, Lecture Notes in Math. 897, Springer-Verlag, New York, 1981.
4. Chang, C. C. and Keisler, H. J., *Model Theory*, North Holland, Amsterdam, 1973.
5. Clark, K., Negation as Failure, in: H. Gallaire and J. Minker (eds.), *Logic and Databases*, Plenum, New York, 1978.
6. Genesereth, M. R. and Nilsson, N. J., *Logical Foundations of Artificial Intelligence*, Morgan Kaufmann, Los Altos, Calif., 1987.
7. Goguen, J. A., Thatcher, J. W., Wagner, E. G., and Wright, J. B., Initial algebra semantics and continuous algebras, *J. Assoc. Comput. Mach.* 24 (1977).
8. Henschen, L. J. and Park, H. S., Compiling the GCWA in indefinite data bases, in: *Proceedings Workshop on Deductive Databases and Logic Programming*, Morgan Kaufmann, Los Altos, Calif., 1987.
9. Jäger, G., Some logical aspects of information processing, in: R. L. Keeney, R. H. Möhring, H. Otway, F. J. Radermacher, and M. M. Richter (eds.), *Proceedings of the International Conference on Multi-attribute Decision Making via O.R.-Based Expert Systems*, Passau, 1986.
10. Lifschitz, V., Closed World Data Bases and Circumscription, *Artificial Intelligence* 27 (1985).
11. Lloyd J. W., *Foundations of Logic Programming*, Springer-Verlag, New York, 1984.
12. McCarthy, C., Circumscription—a Form of Non-monotonic Reasoning, *Artificial Intelligence* 13 (1980).
13. Makowsky, J. A., Why Horn Formulas Matter in Computer Science, in: *Mathematical Foundations of Software Development, Proceedings of the International Joint Conference on Theory and Practice of Software Development (TAPSOFT)*, Lecture Notes in Comput. Sci. 185, Springer-Verlag, New York, 1985.
14. Minker, J., On Indefinite Data Bases and the Closed World Assumption, in: *Proceedings of the 6th Conference on Automated Deduction*, Lecture Notes in Comput. Sci. 138, Springer-Verlag, New York, 1982.
15. Moschovakis, Y. N., *Elementary Induction on Abstract Structures*, North Holland, Amsterdam, 1974.
16. Rabin, M., Characterization of Convex Systems of Axioms, Abstract 571-65, *Amer. Math. Soc.*, 1960.
17. Rabin, M., Classes of Models and Sets of Sentences with the Intersection Property, *Ann. Fac. Sci. Univ. Clermont*, 71 (1962).
18. Reiter, R., On Closed World Data Bases, in: H. Gallaire and J. Minker (eds.), *Logic and Databases*, Plenum, New York, 1978.
19. Reiter, R., Equality and Domain Closure in First-Order Data Bases, *J. Assoc. Comput. Mach.* 27 (1980).
20. Robinson, A., *Introduction to Model Theory and the Metamathematics of Algebra*, North Holland, Amsterdam, 1963.
21. Shepherdson, J. C., Negation as Failure: A Comparison of Clark's Completed Data Base and Reiter's Closed World Assumption, *J. Logic Programming* 1 (1984).
22. Shepherdson, J. C., Negation as Failure II, *J. Logic Programming* 3 (1985).
23. Shepherdson, J. C., Negation in Logic Programming, in: J. Minker (ed.), *Foundations of Deductive Databases and Logic Programming*, Morgan Kaufmann Publishers, Los Altos, 1988.
24. Shoenfield, J., *Mathematical Logic*, Addison-Wesley, Reading, Mass., 1967.