Completeness Results for Circumscription

Donald Perlis[1] and *Jack Minker*[1,2]

MYDepartment of Computer Science[1] MXUniversity of Maryland Institute[2]
of Maryland Advanced Computer Studies

College Park, MD 20742

Abstract

We investigate the model theory of the notion of circumscription, and find completeness theorems that provide a partial converse to a result of McCarthy. We show that the circumscriptive theorems are precisely the truths of the minimal models, in the case of various classes of theories, and for various versions of circumscription. We also present an example of commonsense reasoning in which first-order circumscription does not achieve the intuitive and desired minimization.

1. Introduction

McCarthy [1980] defines predicate circumscription and provides the soundness half of a model theory for this notion but not the completeness half. In McCarthy [1984a] formula circumscription is defined, generalizing predicate circumscription in two distinct ways (variable predicate interpretations and second-order wffs), and Etherington [1984] has provided a model theory and proven soundness there as well. Here we establish the completeness half of the model theory for predicate and variable circumscription in various broad classes of theories.

The paper is organized as follows: In section 2 the idea of circumscription is reviewed, and notations are fixed. In section 3 the completeness problem is discussed, and counterexamples are given. In section 4 we prove a simple but fundamental lemma (Lemma 0) and then go on to establish completeness results in certain broad classes of theories, namely, ones we call P-finite, explicitly-P-defining, and disjunctively-P-defining. Finally in section 5 we turn to theories not falling into any of these categories, and show that variable (first-order) circumscription fails to be complete in a commonsense setting. Throughout we will rely on certain key examples largely borrowed from (or variations on those of) others.

2. The Idea of Circumscription

We review briefly the idea of circumscription. Given a predicate symbol P and a formula A[P] containing P, the circumscription of P by A[P] can be thought of as saying that the P-things consist of certain ones as needed to satisfy A[P] and no more, in the sense that any set of P-things x given by a wff Zx such that A[Z] holds, already includes *all* P-things. This is expressed by means of a schema or set of wffs, which we symbolize as A[P]/P, as follows:

$$A[P]/P = \{[A[Z] \,\&\, (x)(Z(x) \rightarrow P(x))] \;\rightarrow\; (y)(P(y) \rightarrow Z(y)) \mid Z \text{ is a wff}\}$$

A key example, a variation on one emphasized by McCarthy [1980], is the following: let A[P] be a≠b &. P(a) v P(b). Let Z_1(x) be x=a and Z_2(x) be x=b. Then from P(a) v P(b) we get that either Z_1 or Z_2 will serve for circumscription, i.e., either Z_1(x) → P(x) and hence P(x) → Z_1(x), or Z_2(x) → P(x) and hence P(x) → Z_2(x). Thus either a is the only P-thing, or b is; indeed, ¬P(a) v ¬P(b) will then be provable from A[P] + A[P]/P.

McCarthy [1984a] generalized his original notion of (predicate) circumscription to allow specified predicates other than P to vary as well as P; this decisively extends the range of applicability of circumscription. In the new formulation, called *formula circumscription*, the schema is replaced by a single second-order formula, which further increases the power, as will be indicated later. In most of what follows we will however examine and compare the specific innovation of varying predicates (what we shall call *variable circumscription*) with ordinary predicate circumscription. For this purpose we retain a schema, denoted A[P_1,...,P_n]/E, in the following form:

$$\{A[Z_1,...,Z_n] \,\&\, (x)(E[Z_1,...,Z_n] \rightarrow E) \rightarrow (x)(\, E \rightarrow E(Z_1,...,Z_n)) \mid \text{wffs } Z_1,...,Z_n\}$$

where E = E[P_1,...,P_n] is a formula in which P_1,...,P_n may appear, and E[Z_1,...,Z_n] is obtained from E by substituting Z_i for each P_i. Here the intuitive idea is to minimize (the extension of) the formula E, by allowing variations in (the extensions of) P_1,...,P_n. (In the second-order (formula) version, the "variable" wffs Z_i become instead "predicate variables", and the entire schema expression is universally quantified over these "second-order" variables to turn it into a single second-order formula. See van Dalen [1983] for a concise introduction to second-order logic and its relation to first-order logic.)

As McCarthy [1984b] has observed, it is the presence of the "variables" Z_1,...,Z_n that gives variable circumscription its power, and not the fact that E may be a formula. Indeed, forming an extension-by-definitions of A[P] by adding the new axiom (x)(P_0x↔Ex) where P_0 is a new predicate letter, one can simply circumscribe P_0 with P_0,...,P_n as variable predicates in the extension of A[P]. That is, we can just as well take E to be a single predicate letter P_0, since any formula that we may wish to minimize can be made equivalent to such a P_0 by means of an appropriate axiom included in A[P] itself. Thus we will employ this version of circumscription. In the sequel then, E is the predicate letter P_0, and P stands for P_0,P_1,...,P_n, i.e., E plays the role of P_0 above, unless context dictates otherwise. Then the schema A[P]/P is as above except that the parameters P_0,...,P_n appear rather than simply P_1,...,P_n, and the wffs Z_0,...,Z_n as well, again where P_0 is E[P_0,...,P_n] and Z_0 substitutes for E[Z_0,...,Z_n]. To be precise, A[P]/P will be the set of wffs

$$\{[A[Z_0,...,Z_n] \,\&\, (x)(Z_0x \rightarrow P_0x)] \;\rightarrow\; (y)(P_0y \rightarrow Z_0y) \;\mid\; Z_0,...,Z_n \text{ are wffs}\}$$

We abbreviate the theory obtained from A[P] by adjoining the set A[P]/P as new axioms, with the notation A[P]* whenever the P can be understood from context. I.e., A[P]* = A[P] + A[P]/P.

As an example using variable circumscription, we present the following "Life and Death" problem: Let A[D,L] be the axiom

$$(x)(Dx \leftrightarrow \neg Lx) \,\&\, La \,\&\, Db \,\&\, Kc \,\&\, (a \neq b \,\&\, a \neq c \,\&\, b \neq c)$$

which we intend to have the interpretation that dead things (D) are those that are not living (L), and a is living, b is dead, and c is a kangaroo (K). The circumscription of D then corresponds to the notion that as few things as possible are to be considered dead. However, if we were to use mere predicate circumscription, i.e., A[D]* rather than A[D,L]*, then D could not be "squeezed" down by means of an appropriate Z predicate since L, being unchanged, would force D to be its unchanging complement. Thus A[D]* would not have either Dc or Lc as theorems. On the other hand, A[D,L]* does have ¬Dc, and hence Lc, as a theorem. This can be seen by circumscribing with the two predicates x=b (for Z_0) and x≠b (for Z_1).

3. The Completeness Problem

In McCarthy [1980] the concept of minimal model was discussed in the context of predicate circumscription. Etherington [1984] has re-defined minimal model in a manner appropriate to McCarthy´s new (formula) version of circumscription, which we present in slightly modified form for variable circumscription as follows. Let M and N be models of A[P] = A[P_0,P_1,...,P_n] with the same domains and the same interpretations of all constant, function, and predicate symbols except possibly P_0,P_1,...,P_n. We say M *P-reduces* N if the extension of P_0 in M is a proper subset of that in N. Then N is a *P-minimal model* of A[P_0,...,P_n] if N is a model of A[P_0,...,P_n] and no model M of A[P_0,...,P_n] P-reduces N. (We should mention here that by ´model´ we mean ´normal model´, i.e., a model in which equality is interpreted as identity. This incidentally shows the pointlessness of choosing P_0 to be the equality predicate, for then two distinct elements necessarily cannot be identical and so all (normal) models are minimal for equality.)

As an example, consider again McCarthy´s axiom A[P]: a≠b &. Pa v Pb. Here P_0 is just P. It is easily seen that the P-minimal models are precisely ones of the form {Pa a=a b=b c_1 = c_1 c_2= c_2 ...} or {Pb a=a b=b c_1 = c_1 c_2= c_2...} where the number of c_i´s may be none or any other cardinality. In particular, M_1 = {Pa} and M_2 = {Pb} are two such models. But {Pa Pb}, although it is a model of A[P], is not P-minimal.

DEFINITION: A |P-- B means A* |-- B (i.e., A + A[P]/P |-- B), and A |P= B means B holds in any P-minimal model of A. Thus |P-- expresses circumscriptive provability, and |P= expresses the intended circumscriptive consequences. The extent to which these match will now be examined.

We begin by stating a result, variants of which have been given in Davis [1980] (for what is often called ´domain´ circumscription), in McCarthy [1980] (for predicate circumscription), in Minker and Perlis [1984] (for ´protected´ circumscription), and extended by Etherington [1984] to formula circumscription.

SOUNDNESS THEOREM: For any formula B

A[P] |P-- B implies A[P] |P= B

where P is a vector of predicate symbols $P_0, P_1, ..., P_n$

Again, the example above will illustrate this. Since A[P] |P-- ¬Pa v ¬Pb as we saw earlier, then it follows that ¬Pa v ¬Pb holds in the models M_1 and M_2. Of course, we also see directly that this is the case.

Unfortunately in general the converse does not hold, as shown by Davis [1980]. Let A[N] be Peano arithmetic (with the postulates N(0), (x)(N(x)→N(x+1)), etc.) Then the N-minimal models contain N-extensions isomorphic to the natural numbers, so that the formulas B relativized to N that are true in these models are precisely those which are true in arithmetic. But no recursive first-order theory, including one of the form A[N]* = A[N] + A[N]/N, has as its theorems precisely those sentences true of the natural numbers, nor even its N-relativized theorems. Moreover, no recursive *second-order* theory has as its N-relativized theorems precisely those sentences true of the natural numbers, so this is a counterexample to completeness for formula circumscription as well as for variable circumscription. Later we shall present a more commonsense example where completeness fails for variable, but not formula, circumscription (in the process illustrating the greater power of formula circumscription).

Nevertheless, we shall show that certain partial converses do hold, which have rather broad application. First we fix some terminology. We say a theory A[P] is *P-complete* if A[P] |P= B implies A[P] |P-- B for all B, i.e., if the converse to the Soundness Theorem holds for A[P]. (Note that the full converse to Soundness, which as we have just observed is false, is simply the assertion that every theory is P-complete for every vector P of predicate symbols.)

4. Completeness Results

Although as we have noted above, A[P] |P-- B iff A[P]* |-- B (by definition), a similar result does not hold for the double-turnstile: in general it is not the case that A[P] |P= B implies A[P]* |= B, as the example of Davis shows, even though the converse does hold. This is in essence the difficulty we address now.

The following result is fundamental to the rest of our treatment.

LEMMA 0: If every model of A[P]* is P-minimal, then A[P] is P-complete.

Proof: We must show that A[P] |P= B implies A[P] |P-- B. So assume A[P] |P= B. Now if M is a model of A[P]* then M is P-minimal by hypothesis, and consequently M |= B. But then by completeness of first-order logic, we have A[P]* |-- B, i.e., A[P] |P-- B.

As we saw above, Davis's counterexample shows precisely that models of A[P]* are not necessarily P-minimal, contrary to one´s expectations. Indeed, the distinction seems to be that A[P]/P expresses minimization merely with respect to definable subsets of the domains of models, whereas the notion of P-minimal model refers to arbitrary subsets for extensions of predicates. [We note parenthetically that second-order (formula) circumscription is only in a slightly better situation here, for it too is prey to Davis's example, and therefore also fails to express minimization satisfactorily in general, at least for purposes of *deduction* of truths in minimal models, even though the *intended* interpretation of second-order (formula) circumscription is that of these models.] Lemma 0 then has the consequence that if we restrict attention only to models of A[P]*, a kind of completeness holds for all theories A[P], which can be formulated as follows:

DEFINITION: A P-circumscriptive model of A[P] is a model of A[P]* = A[P] + A[P]/P.

Note that A[P] |P-- B iff B holds in all P-circumscriptive models of A[P]. However, although this is in fact correct, it is not particularly useful in itself since in general it is not obvious which models of A[P] are P-circumscriptive, whereas it often is quite an easy matter to characterize the P-minimal models. Moreover, the underlying motivation for circumscription corresponds more closely to the idea of a P-minimal model than to that of a P-circumscriptive model. We then must face the fact that not all P-circumscriptive models are P-minimal and that P-completeness in general fails. However, Lemma 0 also holds out hope of special cases in which P-completeness may hold. We seek then conditions under which all models of A[P]* (i.e., all P-circumscriptive models of A[P]) will be P-minimal.

4.1 P-finite theories

We begin by considering theories A[$P_0, ..., P_n$] where extensions of the P_i are finite in all models.

LEMMA 1: If A[P] |P= B then B is true in all finite models of A[P]*, where B is an arbitrary formula.

Proof: We begin by showing that any *finite* model of A[P]* is also P-minimal. Let N be a finite model of

A[P]*, and suppose N is not P-minimal; then there is a model N´ which P-reduces N, i.e., all predicates other than $P_0,P_1,...,P_n$ have the same extension in N´ as in N, and yet the extension C of P_0(x) in N´ is a proper subset of its extension in N, so that for some b, N |= P_0(b) and N´ |= ¬P_0(b). Let $c_1,...,c_k$ be the elements of C, so that

$$C = \{c_1,...,c_k\} = \{x \text{ in dom(N): N´} \models P_0(x)\}$$

and thus $b \neq c_1,...,b \neq c_k$.

Let Z_0(x) be x=c_1 v ... v x=c_k, and more generally let Z_j be the disjunction of atoms x=c where c ranges over the extension of P_j in N´ . (Note that this is a correct use of a Z predicate, even though the constants c_i may not be in the language L; for we can start with the formula x=y_1 v ... v x=y_n, which is in L, and instantiate the constants c_i for the variables y_i once we pass to the model N.) Then Z_0(b) is false in both N and N´ (indeed $Z_0,...,Z_n$ each have the same extensions in N as in N´). Also, $P_i(x) \leftrightarrow Z_i(x)$ in N´ .

Since N is a model of A[P]*, then the circumscriptive schema holds in N, and so the results of circumscribing $P_0,...,P_n$ using $Z_0,...,Z_n$ must also hold in N. First, we show that A[Z] holds in N, where we write Z to stand for $Z_0,...,Z_n$. But A[P] holds in N´ , and so then must A[Z] since $P_i(x) \leftrightarrow Z_i(x)$ in N´ . Now since A[Z] has the same meaning in both N and N´ (no P_i´ s are left in this formula) then also A[Z] is true in N.

Next we see that $Z_0(x) \rightarrow P_0(x)$ in N. But this is trivial since we already have $P_0(c_1),...,P_0(c_k)$ in N, and $Z_0(x) \leftrightarrow$ x=c_1 v...v x=c_k. Now if we circumscribe, we get $P_0(x) \rightarrow Z_0(x)$ in N. As a consequence, $P_0(b) \rightarrow Z_0(b)$ in N. But P_0(b) is true in N, and Z_0(b) is false, which is a contradiction.

Now we can conclude that if A[P] |P= B then B holds in all finite models N of A[P]*, for we have just seen that any such model is minimal.

THEOREM 1: If A[P] has only finite models, then for all sentences B, A[P] |P= B iff A[P] |P-- B.

Proof: We have already noted the right-to-left order of entailment in the Soundness Theorem. Let us proceed to the converse. If A[P] |P= B, then B holds in all finite P-circumscriptive models of A[P] by Lemma 1. Since all models of A[P] are finite, then in fact B holds in all circumscriptive models of A[P], i.e., A[P]* |= B, so by the completeness theorem of first-order logic, A[P] |P-- B.

We remind the reader that *model* throughout refers always to a *normal* model, so that the existence of a theory with only finite models does not contradict the Lowenheim-Skolem Theorem. We also point out that for the same reason, the Compactness Theorem shows that any such theory will be one for which there are not models of arbitrarily large finite cardinality, i.e., there will be a maximum finite cardinality of its models.

This result is not as restrictive as it may sound. In many applications, it is perfectly appropriate to assume that the universe is finite and that there is a (possibly implicit) axiom to the effect that there are at most Max objects in the universe. (See Reiter [1980] for another view of this sort.) However, we can easily extend the result to certain infinite universes as well, namely ones in which the extensions of the P (i.e., of $P_0,...,P_n$) remain finite. This we amplify in several results which follow.

COROLLARY: If the extensions of P are finite in every P-circumscriptive model of A[P] then for all sentences B, A[P] |P= B iff A[P] |P-- B.

Proof: The same as for Theorem 1, noting that the proof of Lemma 1 requires only that the P and P_0 extensions be finite in each circumscriptive model of A[P].

DEFINITION: A theory is P – finite[1] if it has the axioms (or theorems)

$$(Ey_1...y_{k_i})(x)(P_i(x) \rightarrow x=y_1 \text{ v ... v } x=y_{k_i})$$

for some k_i for each i=0,...,n.

THEOREM 2: If A[P]* is P-finite, then A[P] |P= B iff A[P] |P-- B.

Proof: By the above definitions, if A[P]* is P-finite, then the extensions of $P_0,P_1,...,P_n$ are finite in every P-circumscriptive model, so by the previous corollary we are done.

This then provides a soundness and completeness result for ´ finite´ variable circumscription. The applicability of this result should be broad. Except in the case of deliberate reference to infinite structures such as the integers, it is usually entirely within the scope of the intended domains that the entities enjoying properties P_i be finite in number. Since A[P]* is P-finite whenever A[P] is, then theories that explicitly assert that the number of P-things is bounded are P-complete. Furthermore, theories such as McCarthy's example a≠b &. Pa v Pb have P-finite circumscriptive closure, i.e., {a≠b &. Pa v Pb}* is P-finite, so the theory a≠b &. Pa v Pb is P-complete.

Reiter [1982] addresses the problem of selecting appropriate formulas Z, in his work relating circumscription and the closed world assumption using ideas in Clark [1978]. Here we have provided some insight into the manner in which certain fundamental Z´ s operate. McCarthy [1980] uses two examples to illustrate circumscription and chooses judiciously a disjunction in one and multiple cases for Z in the other; we have used exactly these situations to show that all circumscriptive theorems are characterizable this way in the finitary case.

4.2 Explicitly-P-defining theories

Consider the following example suggested by Ray Reiter. Let the axiom A[P] be

$$Q(b) \,\&\, (x)(Q(x) \rightarrow P(x)).$$

Let N be {Q(b) P(b) P(a)}. Then N clearly is not P-minimal, for N´ = {Q(b) P(b)} P-reduces N. N also is not circumscriptive, as is seen by considering Z(x): x=b. The schema A[P]/P then becomes false in N, for A[Z] and Z(x)→P(x) are true there, while P(x)→Z(x) is not. Note that all models of this sentence have the form

$$\{Q(b)\ P(b)\ Q(c_1)\ P(c_1)\ Q(c_2)\ P(c_2)\ \ldots\ P(d_1)\ P(d_2)\ \ldots\}$$

where there may be an infinite number of c´ s and d´ s. Any minimal model clearly has this form where there are no d´ s, and any such model is minimal.

Pursuing this example further, we note that using Z(x) to be Q(x) allows a quick (circumscriptive) derivation that P(x) → Q(x). Yet our Lemma 1 above employs only equalities and disjunctions for Z, suggesting that P(x) → Q(x) may not be obtainable by the method of the proof. This in fact is the case.

However, Reiter´ s example A[P] is nonetheless P-complete, as the following simple argument shows: Let M be a P-circumscriptive model of A[P]. Then in M the extension of P must contain that of Q. However, letting Z be Q we easily verify that A[Z] and Zx→Px, so that also Px→Qx must hold in M since it satisfies A[P]*. But then M is P-minimal since no model can P-reduce M; Q↔P will have to hold in any model of A[P]* and this means the extension of P cannot be less than it already is in M. So all models of A[P]* are P-minimal, and by Lemma 0 then A[P] is P-complete.

So this remains as a problem to address, since clearly Lemma 1 does not apply to Reiter´ s example. From the previous section it is clear we could introduce a suitable P-finiteness condition which would yield completeness, but the point here is that this is not necessary. Following Doyle [1984] we suggest that the following is appropriate.

> DEFINITION: A theory T will be called explicitly-P-defining if it has a theorem of the form (x)(P_0x↔Wx) where W does not involve the predicate letters $P_0,P_1,\ldots,P_n$.

Then for any explicitly-P-defining theory A[P]*, every P-circumscriptive model of A[P] will be P-minimal. This result is trivial to prove, for in any P-circumscriptive model M the extension of P_0 will have to be that of W, and since W cannot have a smaller extension in a model that P-reduces M, neither can P_0, hence there are no models that P-reduce M and so M is P-minimal. Lemma 0 then shows that A[P] will be P-complete.

We codify this in the following theorem.

> THEOREM 3: If A[P]* is explicitly-P-defining then A[P] is P-complete.

Now, many theories A[P] are such that A[P]* is explicitly-P-defining. For instance, for any explicitly-P-defining theory A[P], A[P]* is also explicitly-P-defining, so the theory (x)(Px↔Qx) is P-complete. Also in Reiter´ s example, {Qb & (x)(Qx→Px)}* is explicitly-P-defining. Moreover any theory consisting of (a conjunction of) wffs of the form (x)(Q_ix→P_0x) will allow proof of P_0x↔(Q_1x v...v Q_mx) by circumscription with Z as Q_1v...vQ_m and so its circumscriptive closure (its starred extension) will be explicitly-P-defining.

4.3 Disjunctively-P-defining theories

There are theories that are not explicitly-P-defining and yet are P-complete. An example is that of McCarthy [1980]: A[P] = {a≠b &. Pa v Pb}. For here there is no wff Q not containing P such that (x)(Px↔Qx) is a theorem of A[P]*. This is easily seen since there are distinct minimal models (indeed P-circumscriptive models) of A[P] with the same domains and interpretations of all predicate letters other than P, so if (x)(Px↔Qx) were true in these then Q would have to contain the letter P after all. Fortunately, this example is P-finite, for circumscription leads to the result that there is a unique P-entity as we have seen earlier. This example though does lead us to another point of view. For it intuitively defines enough of P to provide for P-completeness; it is simply that P is characterized in terms of a disjunction that defies explicit-P-definability. Again Doyle [1984] supplies us with a convenient definition.

DEFINITION: A[P] is disjunctively – P – defining[2] if it has a theorem of the form

$$(x)(P_i x \leftrightarrow W_{i1} x) \text{ v } ... \text{ v } (x)(P_i x \leftrightarrow W_{ik_i} x)$$

for each i=0,...,n where the W's do not involve $P_0,...,P_n$.

This is in analogy with the definition of P-finiteness, where a condition is placed on each P_i. However, in contrast to this, the definition of explicitly-P-defining places a condition only on P_0. This is curious since disjunctively-P-defining is an extension of the definition of explicitly-P-defining. Below we comment further on this matter.

It then is easily seen that McCarthy's example (Pa v Pb)* is disjunctively-P-defining, for it has the theorem (x)(Px↔x=a) v (x)(Px↔x=b). Now we have the following result:

THEOREM 4: If A[P]* is disjunctively-P-defining then A[P] is P-complete.

Proof: This is established in a manner only a little more involved than that for the analogous result for explicitly-P-defining theories. Namely, for any disjunctively-P-defining theory A[P]*, every P-circumscriptive model M will have to have extensions of the P_i's being the same as those of one of the corresponding Wk's, and so if another model M' with the same domain and interpretation of the predicates other than the P_i's has a smaller P_0 extension then this contradicts the schema A[P]/P.

However, just as with P_0-finiteness, we have not found the expected result for disjunctively-P_0-defining theories[3]. The point here is that in the disjunctively-P-defining case, as in the P-finite case, there are predicates Z_i available to substitute for the P_i's, namely, the W's that are guaranteed by the condition of disjunctively-P-defining, which may not exist for P_0-characterizing theories, again supporting the thesis that definability considerations are the crucial ones. See Doyle [1984] for further discussion of definability and circumscription.

5. Commonsense and Completeness

We now turn to an assessment of how lack of completeness may affect commonsense reasoning. We know from Davis that in general there is no solution to the completeness issue in the usual form, but his example came from the rather abstract setting of recursive function theory, and moreover was only an existence argument: there exists a number-theoretic formula that is true in all P-minimal models of arithmetic but is not circumscriptively provable, yet no such formula is exhibited. Of course, these formulas are the undecidable Godel sentences, and presumably far from commonsense reasoning.

David Kueker [1984] has found the following simpler illustration: Let I[P] be the theory Pa, Px↔Psx, a ≠sx, sx=sy→x=y. Then models of I[P] are of two types: those that satisfy the sentence (x).Px→[¬(Ey)x=sy → x=a] and those that do not. But any P-minimal model is isomorphic to the natural numbers N, and is of the former type. Kueker has shown that this sentence is not a theorem of I[P]* = I[P]+I[P]/P, which demonstrates that I[P] is not P-complete. This example is of particular interest, since it shows a sentence that intuitively one would want circumscription to prove, and even more because it can be presented in a fairly simple commonsense setting, indeed a blocks-world setting.

The commonsense interpretation we have in mind is as follows: A king has decreed that all marble blocks in the kingdom are to be painted in certain ways, such that at any block situated in the exact center ´ a´ of the floor of the royal mint is to be purple, purple blocks may lie on or below only purple blocks (or on the floor), and--since purple is a royal and exclusive color--the number of purple blocks is to be minimized. Here the universe of discourse can be thought of as positions x that blocks could occupy, Px as saying that any block at x is to be purple, and sx as the position immediately above x. Then Pa guarantees that only a purple block may be at position a, and similarly for the other axioms. Commonsense would seem to immediately show that no block on the floor at postions other than a should be purple, i.e., that x≠a & ¬(Ey)(x=sy). → ¬Px.

Yet Kueker's result shows that if b is another such block position on the floor, formalized as

a ≠ b & ¬(Ey)b=sy

then ¬Pb is *not* provable by variable circumscription, although one would hope otherwise, since only blocks in the tower over a are intuitively forced to be purple[4]. Indeed, in P-minimal models of I[P], only these blocks will be purple, and ¬Pb therefore is true there. So lack of P-completeness is a hindrance to proving an intuitive and potentially useful result.

Part of the trouble is that we can no longer form the disjunctions that were so useful in the P-finite theories:

$$x=a \vee x=sa \vee x=ssa \vee \ldots$$

is not a (finite) wff. An alternative would be to specify that b is not among the blocks of the tower. This could be expressed as b≠a, b≠sa, b≠ssa, etc., which is an infinite set of formulas and therefore not suitable for circumscription. What we need is a way to refer to the entire tower at once, i.e., in a single (finite) formula. No first-order formula in the language given is available for this purpose. Indeed, I[P]* is not explicitly-P-defining (or disjunctively-P-defining).

We can, nonetheless, introduce a new predicate symbol ´tower´ into the original language that is intended to apply precisely to the positions at and above a (and in particular to no positions on the floor other than at a) as expressed by the following wffs T[P]:

$$\mathrm{tower}(a)$$

$$\mathrm{tower}(x) \leftrightarrow \mathrm{tower}(sx)$$

$$\mathrm{tower}(x) \;\&\; \neg(Ey)(x=sy). \rightarrow x=a.$$

Now let I´[P] be I[P]+T[P]. Taking Z(x) to be Px & tower(x), we easily obtain by ordinary circumscription in the theory I´[P] that

$$P(x) \leftrightarrow \mathrm{tower}(x)$$

and therefore a≠b & ¬(Ey)(x=sy). → ¬Pb. So we have restored explicitly-P-definingness, and therefore P-completeness, to the augmented version I´[P] of I[P], since the theory I´[P]* has just been shown to be explicitly-P-defining. We do not pretend that this is a satisfactory solution, since it was necessary to augment the theory in a non-obvious way; in effect we had to mentally circumscribe in order to see how to alter the theory. It does seem to illustrate, however, that definability once more is a key to a proper formulation of circumscription (in this case, definability of the minimal closure of the successor function s, which we then partially developed by means of ´tower´). We view our discussion as indicative of what may be required for a powerful application of circumscription: for instance, a second-order language as in formula circumscription, or, more generally, some appropriate version of set theory in which minimal closures of functions are obtainable.

6. Conclusions

We have given conditions under which P-completeness, i.e., a match between proof theory and model theory, will occur for variable circumscription. We have also shown that counterexamples persist even in the commonsense realm, but that it may still be possible to take a theory whose circumscriptive closure is not P-complete, and to transform it into a theory that is P-complete. The underlying idea in doing this was to augment the original language by suitable expressions to then allow the appropriate Z to be defined in a single formula. It is of interest that another transformation, namely Skolemization, which also introduces new expressions (terms) into a language, has been used by Etherington, Mercer, and Reiter [1984] to guarantee the existence of minimal models in certain theories.

It remains to determine a general method for extending or transforming a theory into one for which completeness may hold, and to determine the limitations of such a method. It also remains to characterize P-completeness of second-order (formula) circumscription, since here too failure of P-completeness in general means that the intended models (in this case the P-minimal principal second-order models) do not coincide with the models characterizing the (second-order) deduction relation.

Footnotes

1. A problem suggested by the above is: If A[P]* is P_0-finite is A[P] necessarily P-complete? Lifschitz [private communication] has observed that a variation on Davis's example shows this to be false.

2. Lifschitz [1984] has shown the significance of a subclass of theories vis a vis formula (second-order) circumscription: the separable theories. These are related to the disjunctively-P-defining theories, namely, as his Proposition 5 shows, any separable theory is also disjunctively-P-defining. That is, he shows that the second-order circumscriptive axiom is equivalent (proof-theoretically) to a single first-order formula in the same language. It follows that such theories are P-complete in our sense.

3. This suggests the following question: If A[P]* is P_0-characterizing then is A[P] necessarily P-complete? Lifschitz [private communication] has observed that again a variation on Davis's example shows that this is not the case.

4. An anonymous referee has pointed out to us that formula (second-order) circumscription, however, does have the sentence in question as a theorem, thus establishing clearly the greater power of that version. The formula (Q)[Qa & (y)(Qy→Qsy) .→ Qx], where (Q) is a universally quantified predicate variable, serves as a suitable Zx to replace Px in the formula circumscription.

Acknowledgements

Our work obviously depends crucially on that of John McCarthy. We have also benefitted greatly from discussions with David Etherington, David Kueker, Ray Reiter, and John Schlipf. Vladimir Lifschitz has been especially helpful and generous with criticism and suggestions that have resulted in major improvements in the final version of the paper. We also wish to thank two anonymous referees for very constructive and instructive comments.

This research was supported by the following organizations:

AFOSR (AFOSR-82-0303) for J. Minker and D. Perlis

ARO (DAAG29-85-K-0177) for J. Minker and D. Perlis

NSF (NSF MCS 79 19418) for J. Minker

U. of Md. (Summer Research Award) for D. Perlis.

Martin Marietta Corporation for D. Perlis

Clark, K. [1978] "Negation as Failure", In: Logic and Databases, Gallaire, H. and Minker, J., Eds.) Plenum Press, NY 1978, 293-322.

van Dalen, D. [1983] *Logic and Structure*. 2nd edition. Springer-Verlag.

Davis, M. [1980] "The Mathematics of Non-Monotonic Reasoning". Artificial Intelligence 13 (1980), 73-80.

Doyle, J. [1984] "Circumscription and Implicit Definability", Workshop on Nonmonotonic Reasoning, Mohonk, Oct. 17-19, 1984.

Etherington, D. personal communication, Week on Logic and Artificial Intelligence, Univ. of Maryland, Oct. 22-26, 1984.

Etherington, D., Mercer, R., and Reiter, R. [1984] "Drawing the Line Around Circumscription", Workshop on Nonmonotonic Reasoning, Mohonk, Oct. 17-19, 1984.

Kueker, D. "Another Failure of Completeness for Circumscription", photocopied notes, Week on Logic and Artificial Intelligence, Univ. of Maryland, Oct. 22-26, 1984.

Lifschitz, V. "Some Results on Circumscription", Workshop on Nonmonotonic Reasoning, Mohonk, Oct. 17-19, 1984.

McCarthy, J. [1980] "Circumscription--A Form of Non-Monotonic Reasoning". Artificial Intelligence 13 (1980), 27-39.

McCarthy, J. [1984a] "Applications of Circumscription to Formalizing Common Sense Knowledge" Workshop on Nonmonotonic Reasoning, Mohonk, Oct. 17-19, 1984.

McCarthy, J. [1984b] personal communication, Workshop on Nonmonotonic Reasoning, Mohonk, Oct. 17-19, 1984.

Minker, J. and Perlis, D. [1984] "Protected Circumscription". Preprint, Workshop on Nonmonotonic Reasoning, Mohonk, Oct. 17-19, 1984.

Reiter, R. [1978] "On Closed World Databases". In: Logic and Data Bases, (Gallaire, H. and Minker, J., eds.) Plenum, 1978, 55-76.

Reiter, R. [1980] "Equality and Domain Closure in First-Order Databases". JACM 27, 2 (April 1980), 235-249.

Reiter, R. [1982] "Circumscription Implies Predicate Completion (Sometimes)". Proceedings of AAAI-82, 418-420.