# Circumscription: Completeness reviewed

Manfred Jaeger
Max-Planck-Institut für Informatik, Saarbrücken
Correspondence address:
Manfred Jaeger, Max-Planck-Institut für Informatik
Im Stadtwald, 6600 Saarbruecken, FRG

September 7, 2001

**Abstract**

In this paper we demonstrate that some results on the completeness of P-defining theories published earlier are incorrect. We point out that by restricting the original propositions to well-founded theories results somewhat weaker than the origninal ones can be retained. We also we present a theorem that provides some insight into the relation between completeness and reducibility and helps to identify the theories whose minimal models can be adequately handled with circumscription.

# 1 Introduction

The concept of P-minimal models of first-order sentences was introduced as a formalization of nonmonotonic reasoning in [7]. There are several approaches to capturing P-minimality syntactically. Usually either a second-order formula axiomatizing minimality, or a first-order sentence schema, yielding an essentially weaker description of minimality, are used.

While in the first case the question arises whether the second-order circumscription formula is equivalent to some first-order formula, one has to contend with the problem of completeness in the second case. Informally speaking, a first-order sentence T is said to be complete if the sentences that can be deduced from the circumscription schema of T are precisely those that are true in all minimal models of T. Hence it is the class of complete sentences for which the circumscription schema is a fully appropriate means of determining everything one might wish to know about their minimal models.

Unfortunately, completeness seems to be a rather scarce quality among first-order formulas. Davis ([3]) gave a first example of a theory that fails to be complete. Perlis and Minker ([8]) studied the problem of completeness in some detail, identifying various classes of theories for which completeness holds. For some of these classes, however, the results stated by Perlis and Minker are incorrect.

After briefly reviewing the fundamental definitions in section 2, we will demonstrate this by giving a counterexample in section 3. Section 4 presents a new theorem providing some insight into the relation between completeness and reducibility.

# 2 Fundamentals

We will be concerned with that version of circumscription that has been called variable circumscription in [8]. The following is a brief summary of the fundamental definitions.

Let T(P,$\mathbf{Q}$) be a first-order sentence in a language containing the predicate symbols P and $Q_1, \ldots, Q_n$ (this list will be abbreviated by $\mathbf{Q}$).

Let M and N be models of T(P,$\mathbf{Q}$). We say that N P;$\mathbf{Q}$-*reduces* M (writing $N \prec_{P;\mathbf{Q}} M$) if the domains of the two models are the same, the extensions of symbols other than P and $Q_1, \ldots, Q_n$ are the same in M and N, and N[P] is a proper subset of M[P]. No restriction is imposed upon the extensions of the $Q_i$. They may be varied freely in order to minimize the extension of P.

A model of T(P,$\mathbf{Q}$) that cannot be P;$\mathbf{Q}$-reduced by any other model of T(P,$\mathbf{Q}$) is called a P;$\mathbf{Q}$-*minimal* model of T(P,$\mathbf{Q}$).

The class of P;$\mathbf{Q}$-minimal models of T(P,$\mathbf{Q}$) can be axiomatized using the second-order circumscription formula:

$$Circ_{\mathrm{II}}(\mathrm{T} : \mathrm{P};\mathbf{Q}) := \mathrm{T}(\mathrm{P},\mathbf{Q}) \ \wedge \forall \mathrm{p}, \mathbf{q}(\mathrm{T}(\mathrm{p}, \mathbf{q}) \wedge \mathrm{p} \subset \mathrm{P} \rightarrow \mathrm{P} \subset \mathrm{p}).$$

Here p and $\mathrm{q}_1, \ldots, \mathrm{q}_n$ are second-order predicate variables of the same arity as P and $\mathrm{Q}_1, \ldots, \mathrm{Q}_n$ respectively. As a first order approximation of this second-order formula we define a sentence schema of the form

$$\mathrm{T}(\Phi, \mathbf{\Psi}) \wedge \ \Phi \subset \mathrm{P} \rightarrow \mathrm{P} \subset \Phi,$$

where $\Phi$ and $\Psi_1, \ldots, \Psi_n$ are formulas in the language of T, with a number of free variables at least as great as the arities of P and $\mathrm{Q}_1, \ldots, \mathrm{Q}_n$ respectively. The formula $\mathrm{T}(\Phi, \mathbf{\Psi})$ is obtained from T(P,$\mathbf{Q}$) by replacing every expression $\mathrm{Pt}_1 \ldots \mathrm{t}_k$ and $\mathrm{Q}_i\mathrm{s}_1 \ldots \mathrm{s}_l$ by $\Phi(\mathrm{t}_1, \ldots, \mathrm{t}_k)$ and $\Psi_i(\mathrm{s}_1, \ldots, \mathrm{s}_l)$. Here $\Phi(\mathrm{t}_1, \ldots, \mathrm{t}_k)$ is the result of substituting $\mathrm{t}_1, \ldots, \mathrm{t}_k$ for the first $k$ free variables of $\Phi$. Any remaining free variables of $\Phi$ are universally quantified. Similarly for $\Psi_i(\mathrm{s}_1, \ldots, \mathrm{s}_l)$.

The union of T(P,$\mathbf{Q}$) and all instances of this schema is denoted by *Circ*(T : P;$\mathbf{Q}$).

*Circ*(T : P;$\mathbf{Q}$) expresses the minimality of models of T(P,$\mathbf{Q}$) in regard to definable subsets. More precisely: A model M of T(P,$\mathbf{Q}$) is a model of *Circ*(T : P;$\mathbf{Q}$) iff for all models N that P;$\mathbf{Q}$-reduce M one of the extensions $\mathrm{N}[\mathrm{P}], \mathrm{N}[\mathrm{Q}_1], \ldots, \mathrm{N}[\mathrm{Q}_n]$ is not definable in M.

It is noteworthy, that *Circ*(T : P;$\mathbf{Q}$) does not necessarily yield the best possible first-order description of P;$\mathbf{Q}$-minimal models of T(P,$\mathbf{Q}$). There are theories T(P,$\mathbf{Q}$) for which a suitable first-order sentence fully characterizes the class of their P;$\mathbf{Q}$-minimal models, but *Circ*(T : P;$\mathbf{Q}$)does not.

We are now in a position to define the crucial notions of minimal entailment and circumscriptive inference (cf.[7]): We say that T(P,$\mathbf{Q}$) *minimally entails* a first order formula $\phi$ (writing $\mathrm{T}(\mathrm{P},\mathbf{Q}) \models_{\mathrm{P};\mathbf{Q}} \phi$) if $\phi$ holds in all P;$\mathbf{Q}$-minimal models of T(P,$\mathbf{Q}$). $\phi$ is said to be *circumscriptively inferable* from T(P,$\mathbf{Q}$) (writing $\mathrm{T}(\mathrm{P},\mathbf{Q}) \vdash_{\mathrm{P};\mathbf{Q}} \phi$) if $Circ(\mathrm{T} : \mathrm{P};\mathbf{Q}) \vdash \phi$.

As $Circ(\mathrm{T}:\mathrm{P};\mathbf{Q})$ is intended to be a syntactic representation of P;**Q**-minimality, the question to address now is: To what extend do the two relations $\models_{\mathrm{P};\mathbf{Q}}$ and $\vdash_{\mathrm{P};\mathbf{Q}}$ coincide?

As McCarthy has shown ([7]), $\mathrm{T}\vdash_{\mathrm{P};\mathbf{Q}} \phi$ always implies $\mathrm{T}\models_{\mathrm{P};\mathbf{Q}} \phi$, but the converse does not hold ([3]). Theories for which $\mathrm{T}\models_{\mathrm{P};\mathbf{Q}} \phi$ also implies $\mathrm{T}\vdash_{\mathrm{P};\mathbf{Q}} \phi$ will be called P;**Q**-*complete*.

# 3 Completeness

In this section we will discuss some of the results on completeness presented in [8]. It will be shown that the results on explicitly-P;**Q**-defining theories presented in this paper need to be revised.

We begin with restating lemma 4.1 of [8], a basic criterion for completeness.

**Lemma 3.1** If every model of $Circ(\mathrm{T}:\mathrm{P};\mathbf{Q})$ is P;**Q**-minimal, then $\mathrm{T}(\mathrm{P},\mathbf{Q})$ is P;**Q**-complete.

In what follows, Perlis and Minker claim that the prerequisite of this lemma is fulfilled by a class of theories we are now to investigate.

**Definition 3.2** A theory T is called *explicitly*-P;**Q**-*defining* if there exists a formula $\Phi$ in the language of T without the predicate letters $\mathrm{P}, \mathrm{Q}_1, \ldots, \mathrm{Q}_n$ such that $\mathrm{T} \models \forall \mathbf{x}(\mathrm{P}\mathbf{x} \leftrightarrow \Phi(\mathbf{x}))$.

Obviously, for any explicitly-P;**Q**-defining theory T every model of T will be P;**Q**-minimal, because no minimization whatsoever of P with $\mathrm{Q}_1, \ldots, \mathrm{Q}_n$ as variable predicates can be achieved, the extension of P being completely determined by the extensions of symbols other than $\mathrm{P}, \mathrm{Q}_1, \ldots, \mathrm{Q}_n$. It is more interesting therefore, to look at $Circ(\mathrm{T}:\mathrm{P};\mathbf{Q})$ rather than T. Doing so, Perlis and Minker propose the following: If $Circ(\mathrm{T}:\mathrm{P};\mathbf{Q})$ is explicitly-P;**Q**-defining, then every model of $Circ(\mathrm{T}:\mathrm{P};\mathbf{Q})$ is a P;**Q**-minimal model of T.

We will now present an example that proves this proposition to be false. The idea is to define a theory T as a disjunction of two theories $\mathrm{T}_1$ and $\mathrm{T}_2$ so that the following holds:

a) $\mathrm{T}_2$ is explicitly-P;**Q**-defining.

b) For any model M of $\mathrm{T} := \mathrm{T}_1 \vee \mathrm{T}_2$: If $\mathrm{M} \models Circ(\mathrm{T} : \mathrm{P};\mathbf{Q})$, then $\mathrm{M} \models \mathrm{T}_2$.

c) There exists a model $\mathrm{M}_0$ of $Circ(\mathrm{T} : \mathrm{P};\mathbf{Q})$ that can be P;**Q**-reduced by a model $\mathrm{N}_0$ of $\mathrm{T}_1$.

a) and b) then show $Circ(\mathrm{T} : \mathrm{P};\mathbf{Q})$ to be explicitly-P;**Q**-defining, while c) establishes the existence of a model of $Circ(\mathrm{T} : \mathrm{P};\mathbf{Q})$ that is not P;**Q**-minimal.

For the definition of T we use two binary predicate symbols P and R. Minimization will be with respect to P, while R will remain fixed, i.e. there are no variable predicates. $\mathrm{T}_1$ is an adaptation of an example first used in [3]. It consists of the following axioms:

$$\forall x(\exists^{\leq 1} y\mathrm{P}xy \wedge \exists^{\leq 1} y\mathrm{P}yx) \tag{1}$$
$$\forall x(\exists y\mathrm{P}yx \rightarrow \exists y\mathrm{P}xy) \tag{2}$$
$$\exists^{=1} x(\exists y\mathrm{P}xy \wedge \forall y\neg\mathrm{P}yx) \tag{3}$$

The quantifier $\exists^{\leq 1}$ is an abbreviation for 'there exists at most one' which clearly is expressible in first-order logic.

P can be interpreted as a representation of the successor function in the natural numbers. As is shown in [3] and [4], there are no P-minimal models of $\mathrm{T}_1$.

The theory $\mathrm{T}_2$ describes binary trees. To facilitate matters, we use the abbreviation '$x \in \mathrm{Field(P)}$' for the formula $\exists y(\mathrm{P}xy \vee \mathrm{P}yx)$. The first three axioms of $\mathrm{T}_2$ then are:

$$\forall x(x \in \mathrm{Field(P)} \rightarrow \exists^{\leq 2} y\mathrm{P}xy) \tag{1}$$
$$\forall x(x \in \mathrm{Field(P)} \rightarrow \exists^{\leq 1} y\mathrm{P}yx) \tag{2}$$
$$\exists^{=1} x(x \in \mathrm{Field(P)} \;\wedge\; \forall y\neg\mathrm{P}yx) \tag{3}$$

The quantifiers $\exists^{\leq 2}$ and $\exists^{=1}$ have the obvious meaning. Finally, we make $\mathrm{T}_2$ explicitly-P-defining by adding the axiom

$$\forall xy(\mathrm{P}xy \;\leftrightarrow\; \mathrm{R}xy). \tag{4}$$

It remains to show that $\mathrm{T} := \mathrm{T}_1 \vee \mathrm{T}_2$ has the properties b) and c).

b) is easy to see: Let M be a model of $\mathrm{T}_1$. In this case M[P] contains a sequence of the form $(\mathrm{a}_0, \mathrm{a}_1), (\mathrm{a}_1, \mathrm{a}_2), (\mathrm{a}_2, \mathrm{a}_3), \ldots$, where $\mathrm{a}_0$ is the unique element in the field of M[P] with no P-predecessor. By removing $(a_0, a_1)$ from M[P] we can P-reduce M. As the set $M[P]\backslash\{(a_0, a_1)\}$ is definable in M, we conclude that M is not a model of *Circ*(T : P).

To prove c) consider a model $M_0$ of $T_2$ with

$$M_0[P] = M_0[R] = \{(r, a_1), (a_1, a_2), (a_2, a_3), \ldots \ (r, b_1), (b_1, b_2), (b_2, b_3), \ldots\},$$

where all elements $r, a_1, b_1, a_2, b_2, \ldots$ are distinct.

$M_0$ is not a P-minimal model of T: We can choose any one element from $a_1, b_1, a_2, b_2, \ldots$, e.g. $a_i$, and define

$$N_0[P] := \{(a_i, a_{i+1}), (a_{i+1}, a_{i+2}), \ldots\},$$

thereby obtaining a model $N_0$ of $T_1$ that P-reduces $M_0$.

It is quite evident, and can be established by a somewhat involved technical proof, that none of these extensions of P in P-reductions of $M_0$ is definable in $M_0$, hence $M_0 \models$ *Circ*(T : P).

To complete the discussion of our example, we may observe that T is indeed incomplete: No model of $T_1$ is P-minimal, neither is any binary tree that contains an infinite branch which could be used to construct a model of $T_1$. Hence the P-minimal models of T are exactly the models of $T_2$ with finite extensions of P, particularly

$$T \models_P \exists x (x \in \mathrm{Field}(P) \ \wedge \ \neg \exists y P x y).$$

This sentence however, does not hold in $M_0$, and is therefore not circumscriptively inferable from T.

Recall from [6] and [4] that a theory T(P,**Q**) is called *well-founded with respect to* P;**Q** iff for every model M of T that is not P;**Q**-minimal itself, there exists a P;**Q**-minimal model N with $N \prec_{P;\mathbf{Q}} M$.

Confining the original proposition of Perlis and Minker to well-founded theories yields the following theorem. [1]

**Theorem 3.3** If T(P,**Q**) is well-founded with respect to P;**Q**, and *Circ*(T : P;**Q**) is explicitly-P;**Q**-defining, then every model of *Circ*(T : P;**Q**) is P;**Q**-minimal.

[1] As an anonymous referee pointed out, this theorem in its present form has already been given in [1]. P.Besnard quotes the Perlis and Minker-Paper as his source for the theorem but does not comment on the significance of the discrepancy between the theorem as given by him and as found in [8]

**Proof:** Let $\mathrm{M} \models \mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$. Assume that there exists a model $\mathrm{N}_0$ of T with $\mathrm{N}_0 \prec_{\mathrm{P};\mathbf{Q}} \mathrm{M}$.

From the well-foundedness of T we know that either $\mathrm{N}_0$ is already P;**Q**-minimal itself, or there exists a P;**Q**-minimal model $\mathrm{N}_1$ with $\mathrm{N}_1 \prec_{\mathrm{P};\mathbf{Q}} \mathrm{N}_0$. In both cases we conclude that $\mathrm{N} \prec_{\mathrm{P};\mathbf{Q}} \mathrm{M}$ for some P;**Q**-minimal model N. As N is also a model of $\mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$, and $\mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ is explicitly-P;**Q**-defining, this leads to a contradiction: In both M and N the sentence $\forall \mathbf{x}(\mathrm{P}\mathbf{x} \leftrightarrow \Phi(\mathbf{x}))$ holds, with a formula $\Phi$ that does not contain $\mathrm{P}, \mathrm{Q}_1, \ldots, \mathrm{Q}_n$. Consequently, the extension of P is the same in M and N, a contradiction to $\mathrm{N} \prec_{\mathrm{P};\mathbf{Q}} \mathrm{M}$. Hence there is no model that P;**Q**-reduces M, i.e. M is P;**Q**-minimal.

Note that the theory $\mathrm{Qb} \wedge \forall x(\mathrm{Q}x \rightarrow \mathrm{P}x)$ which Perlis and Minker used to illustrate their theorem is well-founded with respect to P, and therefore is indeed P-complete.

Perlis and Minker further present a generalization of their theorem on explicitly-P;**Q**-defining theories to a class of theories called *disjunctively*-P;**Q**-*defining*. This generalization naturally is also prey to our counterexample, but adding the condition of well-foundedness here too is sufficient to restore the theorem.

## 4 Completeness and Reducibility

$\mathit{Circ}_{\mathrm{II}}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ is called *reducible* if there is a first-order sentence $\phi$ such that $\mathit{Circ}_{\mathrm{II}}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ and $\phi$ are equivalent, i.e. have the same models. Reducibility and Completeness have generally been treated as two quite seperate problems. It is an interesting phenomenon however, that in many cases these two notions coincide: For many theories T that can be shown to be complete $\mathit{Circ}_{\mathrm{II}}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ is reducible and vice versa. As an example we mention the class of theories for which Rabinov proves reducibility in [9]. The theories in this class are identified by a syntactic structure generalized from the seperable formulas of [5]. It is not difficult to prove that these theories also satisfy the condition of lemma 3.1, which then establishes their completeness.

To shed some more light on the relation between completeness and reducibility we propose the following:

**Theorem 4.1** The following are equivalent:

(i) Every model of $\mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ is P;**Q**-minimal.

(ii) T is P;**Q**-complete and $\mathit{Circ}_{\mathrm{II}}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ is reducible.

**Proof:** The implication (ii)⇒(i) is easy to prove: As $\mathit{Circ}_{\mathrm{II}}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ is reducible, there is a first-order sentence $\phi$ such that a model M of T is P;**Q**-minimal iff $\mathrm{M} \models \phi$. In particular $\mathrm{T} \models_{\mathrm{P};\mathbf{Q}} \phi$.

Let $\mathrm{M} \models \mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$. From P;**Q**-completeness of T and $\mathrm{T} \models_{\mathrm{P};\mathbf{Q}} \phi$ follows $\mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q}) \vdash \phi$, hence $\mathrm{M} \models \phi$, and M is P;**Q**-minimal.

For the implication (i)⇒(ii) it remains to show that (i) implies reducibility. This follows easily with the following lemma.

**Lemma 4.2** If every model of $\mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ is P;**Q**-minimal, then there exists a finite subset $\Theta$ of $\mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ that is equivalent to $\mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$.

Informally speaking, lemma 4.2 states that whenever (i) holds $\mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ contains only a finite amount of information.

Now suppose (i) holds for $\mathrm{T}(\mathrm{P},\mathbf{Q})$, and $\Theta \subset \mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ as provided by lemma 4.2. We receive the following list of equivalences:

$$\begin{array}{ll} & \mathrm{M} \models \mathit{Circ}_{\mathrm{II}}(\mathrm{T} : \mathrm{P};\mathbf{Q}) \\ \text{iff} & \mathrm{M} \text{ is } \mathrm{P};\mathbf{Q} - \text{minimal} \\ \text{iff} & \mathrm{M} \models \mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q}) \qquad \text{(by (i))} \\ \text{iff} & \mathrm{M} \models \Theta \end{array}$$

Hence $\mathit{Circ}_{\mathrm{II}}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ can be reduced to the conjunction of the sentences in $\Theta$.

The proof of lemma 4.2 relies on an ultraproduct construction. We give a very brief summary of this method here. A good introduction to this subject can be found in section 4.1 of [2].

The ultraproduct construction is an important method of constructing new models from a given (usually infinite) set $\{\mathrm{M}_i \mid i \in \mathrm{I}\}$ of models for a language L. It relies crucially on the concept of ultrafilters which is defined as follows.

Given some set I a subset D of $2^{\mathrm{I}}$ is called an *ultrafilter over* I iff
(i) $\emptyset \notin \mathrm{D}$ (ii) If $\mathrm{x} \in \mathrm{D}$ and $\mathrm{x} \subset \mathrm{y}$ then $\mathrm{y} \in \mathrm{D}$
(iii) For all $\mathrm{x},\mathrm{y} \in \mathrm{D}$: $\mathrm{x} \cap \mathrm{y} \in \mathrm{D}$ (iv) For all $\mathrm{x} \in 2^{\mathrm{I}}$: $\mathrm{x} \in \mathrm{D}$ or $\mathrm{I} \setminus \mathrm{x} \in \mathrm{D}$.

An ultrafilter is called *principal* iff there is a x$\subset$I such that for all y$\subset$I: y $\in$ D iff x $\subset$ y.

Elements of an ultrafilter over I can be thought of as 'large' subsets of I or, somewhat more precisely, subsets of I 'including the significant part of I'.

Given a set $\{\mathrm{M}_i \mid i \in \mathrm{I}\}$ and an ultrafilter D over I the ultraproduct $\prod_{\mathrm{D}} \mathrm{M}_i$ is defined as follows:

- To obtain the domain of $\prod_{\mathrm{D}} \mathrm{M}_i$ we define an equivalence relation $\sim_{\mathrm{D}}$ on the Cartesian product of the domains of the individual $\mathrm{M}_i$'s by

$$(\mathrm{a}_i)_{i\in\mathrm{I}} \sim_{\mathrm{D}} (\mathrm{b}_i)_{i\in\mathrm{I}} \text{ iff } \{i \mid \mathrm{a}_i = \mathrm{b}_i\} \in \mathrm{D}.$$

  The domain of $\prod_{\mathrm{D}} \mathrm{M}_i$ then is the set of equivalence classes with respect to this relation.

- Let us assume that L contains a relation symbol R of arity one. We give the definition of $\prod_{\mathrm{D}} \mathrm{M}_i[\mathrm{R}]$ as an example for how the extensions of symbols in L ar defined: The equivalence class of $(\mathrm{a}_i)_{i\in\mathrm{I}}$ with respect to $\sim_{\mathrm{D}}$ is in $\prod_{\mathrm{D}} \mathrm{M}_i[\mathrm{R}]$ iff $\{i \mid \mathrm{a}_i \in \mathrm{M}_i[\mathrm{R}]\} \in \mathrm{D}$.

This somewhat tedious set of definitions rewards us with the fundamental theorem of ultraproducts which, in its essence, states that for all first-order sentences $\phi$ in L:

$$\prod_{\mathrm{D}} \mathrm{M}_i \models \phi \text{ iff } \{i \mid \mathrm{M}_i \models \phi\} \in \mathrm{D}.$$

**Proof of lemma 4.2:**

Suppose (i) holds for T(P,$\mathbf{Q}$). If *Circ*(T : P;$\mathbf{Q}$) is inconsistent, then compactness of first-order logic yields the existence of a finite inconsistent subset $\Theta$ of *Circ*(T : P;$\mathbf{Q}$). Assume therefore that *Circ*(T : P;$\mathbf{Q}$) is consistent. Assume that $\Theta \not\models$ *Circ*(T : P;$\mathbf{Q}$) for all finite $\Theta \subset$ *Circ*(T : P;$\mathbf{Q}$). We then have: For all finite $\Theta \subset$ *Circ*(T : P;$\mathbf{Q}$) there is a model $\mathrm{M}_\Theta$ and a sentence $\psi_\Theta \in$ *Circ*(T : P;$\mathbf{Q}$) such that $\mathrm{M}_\Theta \models \Theta \cup \{\neg\psi_\Theta\}$.

Let *Circ*(T : P;$\mathbf{Q}$)=$\{\phi_0, \phi_1, \phi_2, \ldots\}$. Without loss of generality we may assume that $\phi_0$=T. For $m \geq 0$ we define:

$$\Theta_m := \{\phi_i \mid i \leq m\},\ \mathrm{M}_m := \mathrm{M}_{\Theta_\mathrm{m}} \text{ and } \psi_m := \psi_{\Theta_\mathrm{m}}.$$

As $\mathrm{M}_m \models \neg\psi_m$, there is a model $\mathrm{M}'_m$ of $\mathrm{T}(\mathrm{P},\mathbf{Q})$ with $\mathrm{M}'_m \prec_{\mathrm{P};\mathbf{Q}} \mathrm{M}_m$ for all $m \geq 0$. From the models $\mathrm{M}_m$ and $\mathrm{M}'_m$ ($m \geq 0$) we can now construct models M and M′ of $\mathrm{T}(\mathrm{P},\mathbf{Q})$ with $\mathrm{M}' \prec_{\mathrm{P};\mathbf{Q}} \mathrm{M}$ and $\mathrm{M} \models \mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$, a contradiction to (i).

To obtain the models M and M′ let D be an arbitrary nonprincipal ultrafilter over the natural numbers. Define

$$\mathrm{M} := \prod_{\mathrm{D}} \mathrm{M}_m \quad \text{and} \quad \mathrm{M}' := \prod_{\mathrm{D}} \mathrm{M}'_m.$$

M is a model of $\mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$: This follows from the fundamental theorem of ultraproducts and the fact that $\mathrm{M}_m \models \phi_i$ for all $\phi_i \in \mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ and $m \geq i$. (Note that $\{m \mid m \geq i\} \in \mathrm{D}$ for all $i$.) Similarly $\mathrm{M}' \models \mathrm{T}(\mathrm{P},\mathbf{Q})$. As $\mathrm{M'}_m[P] \subset \mathrm{M}_m[P]$ for all $m \geq 0$ we may finally conclude that $\mathrm{M}'[P] \subset \mathrm{M}[P]$.

**Corollary 4.3** If $\mathrm{T}(\mathrm{P},\mathbf{Q})$ is well-founded with respect to $\mathrm{P};\mathbf{Q}$, and $\mathit{Circ}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ is explicitly-$\mathrm{P};\mathbf{Q}$-defining, then $\mathit{Circ}_{\mathrm{II}}(\mathrm{T} : \mathrm{P};\mathbf{Q})$ is reducible.

**Proof:** Theorem 3.3 and 4.1.

# 5 Conclusion

While circumscription in some cases provides a simple and efficient way of describing minimality, there is no getting around the fact that it must fail in others.

It is very desirable, therefore, to find more precise and meaningful characterizations of the classes of theories for which circumscription is suited or not. Theorem 3.3 gives evidence that well-foundedness might be a crucial property that deserves some attention.

Theorem 4.1 gives additional significance to the class of theories which satisfy (i). Not only is this property sufficient to grant reducibility apart from completeness; it also is a necessary one if not at least one of the two main versions of circumscription, variable- and second-order-circumscription, is to fail as a syntactic counterpart to minimal entailment.

There is another angle under which Theorem 4.1 can be viewed: It tells us that while working inside the class given by (i) one does not have to worry, in principle, which of the two major versions of circumscription to use. They both perform equally well. (This, of course, does not say anything about which version to choose in practice: For a given theory inside this class it may still be much easier to compute the first-order formula to which second-order circumscription can be reduced than to find a meaningful instance of the variable-circumscription schema - or vice versa.)

Outside this class matters are not that simple. There are theories which are reducible but not complete and others that are complete but not reducible. Thus there is no natural preference of one of these two versions of circumscription over the other.

# References

[1] P.Besnard, An Introduction to Default Logic (Springer-Verlag, Berlin-Heidelberg 1989).

[2] Chang,C. and Keisler,J., Model Theory (North-Holland, Amsterdam 1970).

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

[4] Etherington,D.W., Mercer,R.E. and Reiter,R., On the Adequacy of Predicate Circumscription for Closed World Reasoning, *Computational Intelligence* 1 (1985) 11-15.

[5] Lifschitz,V., Computing Circumscription, In Proceedings of the IJCAI (1985) 121-127.

[6] Lifschitz,V., On the Satisfiability of Circumscription, *Artificial Intelligence* 28 (1986) 17-27.

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

[8] Perlis,D. and Minker,J., Completeness Results for Circumscription, *Artificial Intelligence* 28 (1986) 29-42.

[9] Rabinov,A., A Generelization of Collapsible Cases of Circumscription, *Artificial Intelligence* 38 (1989) 111-117.