Artificial Intelligence, 54:397–410, 1992.

An Efficient Method for Eliminating Varying Predicates from a Circumscription

Marco Cadoli*

Dipartimento di Informatica e Sistemistica,

Università di Roma "La Sapienza"

Thomas Eiter, Georg Gottlob†

Christian Doppler Labor für Expertensysteme,

Institut für Informationssysteme, Technische Universität Wien

Abstract

Circumscription appears to be the most powerful and well-studied technique used in formalizing common-sense reasoning. The general form of predicate circumscription allows for fixed and varying (floating) predicates. We show that the inference problem under this form of circumscription is efficiently reducible to inferencing under circumscription without varying predicates. In fact, we transform this problem even into circumscription without fixed and varying predicates, that is where all predicates are minimized. Thus any theorem prover or algorithm for inferencing under circumscription without fixed and varying predicates is able to handle inferencing under the general form of predicate circumscription. As a consequence, algorithms that compute circumscription for an inference task can be simplified.

*Mailing address: Via Salaria 113, I-00198 Roma, Italia. E-mail: `cadoli@vaxrma.infn.it`

†Mailing address: Paniglgasse 16, A-1040 Wien, Austria.
Internet e-mail: {`eiter,gottlob`}`@vexpert.dbai.tuwien.ac.at`

1 Introduction

Circumscription is a powerful non-monotonic technique that allows us to deal with the representation of incomplete information in knowledge bases. The basic idea of circumscription is to represent implicitly an assumption of normality through the *minimization* of the extension of predicates formalizing abnormality in a first-order formula. For example, we know that birds normally fly, but exceptions to this general rule can arise. The circumscriptive approach translates this general rule into the formula $\forall x.bird(x) \wedge \neg abnormal(x) \supset flies(x)$ and considers only those models of the formula in which the extension of the predicate $abnormal$ is minimal, that is cannot be reduced without leading to an inconsistent interpretation.

The original definition of circumscription by McCarthy [12] did not take into account the possibility of *varying* the extension of predicates. Variability of predicates means that the extensions of some predicates (the varying ones) can be affected in any way (enlarged or reduced) in order to decrease the extensions of those predicates that are minimized. Etherington, Mercer and Reiter [3] noticed that with the original definition it is impossible to derive new positive facts from a knowledge base. For example, it is impossible to infer $flies(Tweety)$ from the fact $bird(Tweety)$ and the rule $\forall x.bird(x) \wedge \neg abnormal(x) \supset flies(x)$ minimizing *abnormal* while *bird* and *flies* are fixed.

A revised definition of circumscription has been provided by McCarthy in [13], which allows the extensions of some of the predicates to vary, while the extensions of other predicates are minimized. With this new definition of circumscription it is possible to infer new positive facts. If we allow in the previous example the predicate $flies$ to vary, then we can infer $flies(Tweety)$. Recently, several theorem provers and inference algorithms for circumscription have been designed [1, 7, 15].

Generally, the set of the predicates of a formula to be circumscribed is partitioned into three subsets: The predicates to be minimized (denoted by P), those that are allowed to vary (denoted by Z) and those whose extension should remain fixed (denoted by Q). Fixed predicates are used to model definite entities that we do not want to be altered by circumscription. More general formulations of circumscription allow arbitrary expressions $E(P, \mathbf{x})$ to be minimized [13] or functions to be varied [10, 11]. These extensions are inessential, cf. Etherington [4, p.132,Theorem 6.3] and de Kleer and Konolige [2], however, and will be disregarded in this paper.

Several authors have addressed the question of simplifying the definition of circumscription, in particular by providing equivalent formulations of predicate circumscription in terms of simpler forms of circumscription. de Kleer and Konolige [2] show that it is possible to translate any first-order formula $\mathcal{T}$ into another first-order formula $\mathcal{T}'$ such that the circumscription of $\mathcal{T}$ with fixed predicates is equivalent to the circumscription of $\mathcal{T}'$ without fixed predicates. Lifschitz [10] performed a similar transformation for the elimination of varying predicates, but introduced second-order formulae. Yuan and Wang [20] defined a method for eliminating the varying predicates without introducing second-order formulae. The two major drawbacks of their method are that it is incomplete and that it yields in general an exponential-size formula.

In this paper we are concerned with *inferring* formulae from a circumscription rather than transforming a circumscription into an equivalent one. In particular we exhibit an efficient method for reducing the problem of *inference* of a formula under circumscription with varying predicates to the problem of inferencing a formula under circumscription where all predicates are minimized.

In this way we show that any algorithm for performing inference under circumscription without varying predicates (for instance the method of Bossu and Siegel [1]) can be used to perform any inference under circumscription with fixed and varying predicates. Since fixed and varying predicates can be bypassed efficiently combining the method in [2] and our method, well-known algorithms for inference under predicate circumscription (for example Przymusinski's MILO-resolution [15]) can be noticeably simplified.

2 The Reduction

In this section we exhibit a transformation of a pair of first-order formulae $\mathcal{T}$ and $\mathcal{F}$ into a first-order formula $\Gamma(\mathcal{T};\mathcal{F};P;Z)$ such that $\mathcal{F}$ follows from the circumscription of $\mathcal{T}$ where the P predicates are minimized and the Z predicates vary if and only if $\neg u$ follows from the circumscription of $\Gamma(\mathcal{T};\mathcal{F};P;Z)$ under minimization of all predicates, where u is a new 0-ary predicate symbol not occurring in $\mathcal{T}$ or $\mathcal{F}$.

We introduce some notation first. In the following, predicate constants and predicate variables are both referred to as *predicate symbols*. If A denotes a list $a_1,\ldots,a_n$ of predicate symbols occurring in a formula $\mathcal{G}$, then A' and A'' denote the lists $a'_1,\ldots,a'_n$ and $a''_1,\ldots,a''_n$, respectively, where it is assumed that for $1 \le i \le n$, a'_i and a''_i are new distinct predicate symbols not occurring in $\mathcal{G}$ such that a_i, a'_i, a''_i are of the same arity. AB denotes the concatenation of lists A and B of predicate symbols. If A, B are the respective lists $a_1,\ldots,a_n$ and $b_1,\ldots,b_n$ of predicate symbols where a_i and b_i are of the same arity, $1 \le i \le n$, we use the following notation:

$$(A \equiv \neg B) \text{ stands for } \bigwedge_{1\le i\le n} (\forall \mathbf{x}_i.a_i(\mathbf{x}_i) \equiv \neg b_i(\mathbf{x}_i)),$$

$$(A < B) \text{ stands for } [\bigwedge_{1\le i\le n} (\forall \mathbf{x}_i.a_i(\mathbf{x}_i) \supset b_i(\mathbf{x}_i))] \wedge [\bigvee_{1\le i\le n} (\exists \mathbf{x}_i.\neg a_i(\mathbf{x}_i) \wedge b_i(\mathbf{x}_i))],$$

$$(A \equiv \mathit{true}) \text{ stands for } [\bigwedge_{1\le i\le n} (\forall \mathbf{x}_i.a_i(\mathbf{x}_i))].$$

We follow Lifschitz's well-known notion of parallel circumscription in [10]. For the definition of second order structures and models see [10]; as Lifschitz, we denote the

universe of a structure M by $|M|$. For each list A of predicate constants, we denote by $[\![A]\!]_M$ the extension of the A-predicates in structure M. We write $[\![A]\!]_M \leq [\![A]\!]_{M'}$ if and only if for each a in A the extension of a in M is a subset of the extension of a in M'; $[\![A]\!]_M < [\![A]\!]_{M'}$ means $[\![A]\!]_M \leq [\![A]\!]_{M'}$ and $[\![A]\!]_M \neq [\![A]\!]_{M'}$. We write $M <_{P;Z} M'$ for models M, M' iff $|M| = |M'|$, the interpretation of functions and predicates not in PZ is identical and $[\![P]\!]_M < [\![P]\!]_{M'}$, where P, Z are disjoint lists of predicate constants. Model M is $P;Z$-*minimal* iff there is no model M' such that $M' <_{P;Z} M$. If P covers all predicates, *minimal* means $P;\emptyset$-minimal and $M < M'$ stands for $M <_{P;\emptyset} M'$.
$CIRC(\mathcal{T};P;Z)$ is defined by the second-order formula

$$\mathcal{T}[P;Z] \wedge (\forall P'Z'.\mathcal{T}[P';Z'] \supset \neg(P' < P)),$$

where P and Z are disjoint lists of predicate constants and $\mathcal{T}[P';Z']$ is $\mathcal{T}$ with each predicate symbol p in PZ replaced by p'. All remaining predicates not in PZ (the Q predicates) are fixed. Note that unlike in [10, 11], functions are not allowed to vary. For convenience, we denote the circumscription of $\mathcal{T}$ under minimization of all predicates ($Q = Z = \emptyset$) by $CIRC(\mathcal{T})$.

de Kleer and Konolige have shown how to eliminate fixed predicates from a circumscription efficiently (even in linear time) [2]; thus we may assume that no fixed predicates occur in $\mathcal{T}$. Further, we may assume without loss of generality that in any instance of $CIRC(\mathcal{T};P;Z) \models \mathcal{F}$ all predicates in the formula $\mathcal{F}$ occur in $\mathcal{T}$: if a predicate p does not occur in $\mathcal{T}$, simply join a tautology involving only p (e.g., $(\forall x.p(x) \supset p(x))$) to $\mathcal{T}$.

The transformation is as follows:

Definition 2.1 (reduction) *Let $\mathcal{T}, \mathcal{F}$ be first-order formulas, and let the lists P, Z partition the predicate symbols in $\mathcal{T}$. The first-order formula $\Gamma(\mathcal{T};\mathcal{F};P;Z)$ is defined as follows:*

$$\Gamma(\mathcal{T};\mathcal{F};P;Z) = (PZ \equiv \neg P''Z'') \;\wedge\; [(u \wedge (P'Z' \equiv true)) \;\vee\; (\neg u \wedge \Theta(P,Z,P',Z'))]$$

where $\Theta(P,Z,P',Z')$ is the formula $(\mathcal{T} \supset \mathcal{F}) \vee (\mathcal{T}[P';Z'] \wedge (P' < P))$.

Let us first give a rough intuitive explanation of this transformation. The predicates P'' and Z'' and the formula $PZ \equiv \neg P''Z''$ are added in order to make structures with different extensions of PZ incomparable. Hence, minimization has no effect on the predicates from $PZP''Z''$. This allows us to "simulate" standard semantics within a circumscription where all predicates are minimized. The only predicates in $CIRC(\Gamma(\mathcal{T};\mathcal{F};P;Z))$ that are effectively minimized are those in $P'Z'$ and u. Therefore, $CIRC(\Gamma(\mathcal{T};\mathcal{F};P;Z)) \models \neg u$ holds iff in each $P;Z$-minimal model of $\Gamma(\mathcal{T};\mathcal{F};P;Z)$, the formula $\Theta(P, Z, P', Z')$ is satisfied. This means that for each extension of P and Z, either $\mathcal{T} \supset \mathcal{F}$, or the extension is not $P; Z$-minimal; in other words, $\mathit{CIRC}(\mathcal{T}; P; Z) \models \mathcal{F}$ holds.

For examples of this transformation, refer to the appendix, where its application is demonstrated for two well-known examples for inferencing under circumscription.

In the sequel, we formally prove that $\mathit{CIRC}(\mathcal{T}; P; Z) \models \mathcal{F}$ iff $\mathit{CIRC}(\Gamma(\mathcal{T}; \mathcal{F}; P; Z)) \models \neg u$. To this end, we first show a lemma stating that testing the validity of formulas belonging to a certain fragment of second order logic can be reduced to an inference problem under circumscription where all predicates are minimized. In order to show our main result, we then make use of this transformation.

Definition 2.2 *Let Ψ be a second-order formula of the form $\forall A \exists B\, \Phi(A, B)$ where A, B are disjoint lists of predicate variables and $\Phi(A, B)$ is a first-order formula. Denote by $\tau(\Psi)$ the following first-order formula:*

$$(A \equiv \neg A'') \wedge [(u \wedge (B \equiv \mathit{true})) \vee (\neg u \wedge \Phi(A, B))].$$

Lemma 2.1 *Let Ψ be a second-order formula of the form $\forall A \exists B\, \Phi(A, B)$ where A, B are disjoint lists of predicate variables and $\Phi(A, B)$ is a first-order formula containing no other predicate symbols except those in A and B. Then, Ψ is valid if and only if $\mathit{CIRC}(\tau(\Psi)) \models \neg u$.*

Proof:

We first observe that for any possible universe U, $\tau(\Psi)$ has for any interpretation of the function symbols and any extension $\mathcal{E}$ of the A-predicates over U, a unique model M such that $|M| = U$, $[\![A]\!]_M = \mathcal{E}$, $M \models u$, and $M \models (\forall \mathbf{x}. b(\mathbf{x}))$ for every b from B. Furthermore, if any two models M, M' of $\tau(\Psi)$ satisfy $[\![A]\!]_M \neq [\![A]\!]_{M'}$, then neither $M < M'$ nor $M' < M$.

Now assume that Ψ is valid but $\mathit{CIRC}(\tau(\Psi)) \models \neg u$ is false. This means that there is a minimal model M of $\tau(\Psi)$ such that u is true in M. Observe that $[\![B]\!]_M$ is the maximum possible extension of B over $|M|$ because M satisfies $(B \equiv \mathit{true})$. Since Ψ is valid, for the same extension of all function symbols, there must exist an extension $\mathcal{E}$ of B over $|M|$ such that $\Phi([\![A]\!]_M, \mathcal{E})$ is true over $|M|$. Now consider the structure M' with the same interpretation of functions as M and $[\![AA'']\!]_{M'} = [\![AA'']\!]_M$, $[\![B]\!]_{M'} = \mathcal{E}$, and u is false. Since $M' \models \neg u$, and $M' \models \Phi([\![A]\!]_M, \mathcal{E})$, it is a model of $\tau(\Psi)$. As $[\![AA'']\!]_{M'} = [\![AA'']\!]_M$, $[\![B]\!]_{M'} \leq [\![B]\!]_M$ and $[\![u]\!]_{M'} < [\![u]\!]_M$, we have $M' < M$. This is a contradiction to the minimality of M, however.

Conversely, assume that $\mathit{CIRC}(\tau(\Psi)) \models \neg u$ holds. Let $\mathcal{E}$ be an arbitrary extension of the A-predicates over some universe U. By our observation from above, there exists for any U and any function interpretation over U a model M of $\tau(\Psi)$ such that $|M| = U$, $[\![A]\!]_M = \mathcal{E}$, and $M \models u$. From $\mathit{CIRC}(\tau(\Psi)) \models \neg u$ it follows that M cannot be minimal; hence there exists a model $M' < M$ with $[\![A]\!]_{M'} = [\![A]\!]_M$ where u is not

satisfied. (Clearly u cannot be satisfied in any model $M' < M$: Any such M' must satisfy $|M'| = |M| = U$ and $[\![A]\!]_{M'} = [\![A]\!]_M$, and since M is for this function interpretation the unique model of $\tau(\Psi)$ over U satisfying $[\![A]\!]_M = \mathcal{E}$ and $M \models u$, $M' \models \neg u$ must hold; hence, M' must satisfy $\neg u \wedge \Phi(A, B)$, the second term of the disjunction in $\tau(\Psi)$.) Consequently, $\Phi([\![A]\!]_{M'}, [\![B]\!]_{M'})$ is true over U under that function interpretation. Since U, the function interpretation and the extension $\mathcal{E} = [\![A]\!]_{M'}$ of the A-predicates over U were arbitrary, it follows that $\forall A \exists B\, \Phi(A, B)$ is true over any universe U under any function interpretation, and hence Ψ is valid.
□

Theorem 2.2 *Let $\mathcal{T}, \mathcal{F}$ be first-order formulas such that all predicates in $\mathcal{F}$ appear in $\mathcal{T}$, and let the lists P, Z partition the predicates in $\mathcal{T}$. Then, CIRC$(\mathcal{T}; P; Z) \models \mathcal{F}$ if and only if CIRC$(\Gamma(\mathcal{T}; \mathcal{F}; P; Z)) \models \neg u$.*

Proof:

CIRC$(\mathcal{T}; P; Z) \models \mathcal{F}$ holds iff $\mathcal{F}$ is valid in every $P; Z$-minimal model of $\mathcal{T}$ [10, Proposition 1], that is *CIRC*$(\mathcal{T}; P; Z) \models \mathcal{F}$ is true iff every model of $\mathcal{T}$ is also a model of $\mathcal{F}$ or not a $P; Z$-minimal model of $\mathcal{T}$. Thus *CIRC*$(\mathcal{T}; P; Z) \models F$ is true iff the second-order formula $(\forall PZ.(\mathcal{T}[P; Z] \supset \mathcal{F}) \vee (\exists P'Z'.(\mathcal{T}[P'; Z'] \wedge (P' < P))))$ is valid, which is equivalent to the following formula $\Psi[\mathcal{T}; \mathcal{F}; P; Z]$:

$$(\forall PZ\, \exists P'Z'.(\mathcal{T}[P; Z] \supset \mathcal{F}) \vee (\mathcal{T}[P'; Z'] \wedge (P' < P))).$$

Note that $\Psi[\mathcal{T}; \mathcal{F}; P; Z]$ is of the form $\forall A \exists B\, \Phi(A, B)$, where $\Phi(A, B)$ is a first-order formula and $A = PZ, B = P'Z'$. Note that $\forall A \exists B \Phi(A, B)$ satisfies the conditions on the formula Ψ in Lemma 2.1. Since $\tau(\forall A \exists B\, \Phi(A, B))$ is the first-order formula $(A \equiv \neg A'') \wedge [(u \wedge (B \equiv \mathit{true})) \vee (\neg u \wedge \Phi(A, B))]$, we thus get by Lemma 2.1 that *CIRC*$(\mathcal{T}; \mathcal{F}; Z) \models \mathcal{F}$ iff $\neg u$ follows from the circumscription of the first-order formula

$$(PZ \equiv \neg P''Z'') \wedge [(u \wedge (P'Z' \equiv \mathit{true})) \ \vee\ (\neg u \wedge ((\mathcal{T} \supset \mathcal{F}) \ \vee\ (\mathcal{T}[P'; Z'] \wedge (P' < P))))].$$

The latter formula is just the formula $\Gamma(\mathcal{T}; \mathcal{F}; P; Z)$; thus *CIRC*$(\mathcal{T}; P; Z) \models \mathcal{F}$ iff *CIRC*$(\Gamma(\mathcal{T}; \mathcal{F}; P; Z)) \models \neg u$. □

Remark. Lemma 2.1 and Theorem 2.2 put together are a simple proof for the fact that the problem of inferencing under circumscription, whether or not varying predicates are used, is precisely equivalent to the problem of checking the validity of a second-order formula belonging to the Π^1_2-level of the prenex hierarchy of second-order logic. By a well-known result of Hintikka [8], this problem (as well as validity checking for *any* higher-order formula) can be effectively reduced to validity checking of a monadic Σ^1_1 second-order formula as long as standard models are considered (see also [18]). Thus, under standard model semantics in the sense of [18], the inference problem under circumscription, with varying predicates as well as without, is complete for the class Σ^1_1. This complements a result of Schlipf [17], who has shown that the same problem restricted to countably infinite models is $\Pi^1_2(\mathcal{N})$-complete, that is Π^1_2-complete over the natural numbers. (Note that the class $\Pi^1_2(\mathcal{N})$ of the analytic hierarchy is strictly more expressive than the class $\Sigma^1_1(\mathcal{N})$.)

The transformation as a syntactic reduction is feasible in linear time and hence very efficient. The question of how it affects the complexity of the *derivation* is rather intricate, however. On the one hand, our transformation adds new predicate symbols to the initial formula that lead to additional computational overhead. Also, some methods require the input to be in a special form. As an example, Przymusinski's method requires the circumscribed formula to be in clausal form. Hence, in order to use our transformation, the first step is to compute the clausal form of $\Gamma(\mathcal{T};\mathcal{F};P;Z)$. This computation may be very expensive, sometimes more expensive than transforming $\mathcal{T}$ itself into clause form. On the other hand, our transformation eliminates the handling of varying predicates, which constitutes a substantial part of most derivation algorithms. In Przymusinski's algorithm for instance, this step may cause exponential overhead. In summary, our transformation shifts the burden of handling varying predicates to predicate minimization.

An application of our transformation to closed-world reasoning is also possible. This technique for non-monotonic reasoning is closely related to circumscription [3, 10, 5] and has been defined in several ways. In its previous versions (*CWA*, *GCWA*, *EGCWA* [16, 14, 19]), the closed-world assumption was applied to all the predicates. In its more recent definitions (*CCWA*, *ECWA* [6, 5]) it is possible to vary a predicate, and the intended meaning of this variability is the same as in circumscription. Notice that the *Extended Closed World Assumption* (*ECWA*), which like circumscription minimizes some predicates P while varying other predicates Z and fixing the remaining predicates Q, is the most general rule for closed-world reasoning, extending all other versions of closed world-reasoning mentioned above.

Lifschitz demonstrated in [10] that for function-free first-order theories, the *CWA* under the assumptions of domain-closure and unique names is equivalent to predicate circumscription where all predicates are minimized, provided *CWA* is consistent. Gelfond et al. show in [5] that under the same assumptions the closure of a theory under *ECWA* for P, Z is equivalent to predicate circumscription under identical P, Z. Thus our method can fruitfully be applied for simplifying closed-world reasoning. In particular, it is possible to reduce inferencing under *ECWA* to inferencing under the *Extended Generalized Closed World Assumption* (*EGCWA*) defined in [19], since this rule is identical to *ECWA* for $Q = Z = \emptyset$ [5].

3 Conclusions

We have introduced an efficient method to reduce inference under predicate circumscription to inferencing under circumscription where all predicates are minimized, and we proved its correctness. This reduction can be performed in linear time. We believe that the method provides new insights into the relation between the general form of predicate circumscription and circumscription without varying predicates. The same applies to different versions of closed-world reasoning. Furthermore, the method can be practically used to build theorem provers for predicate circumscription out of existing simpler inference engines or procedures for circumscription without varying predicates (e.g., the algorithm by Bossu and Siegel [1]). The new method even allows noticeable simplifications to familiar algorithms to answer queries under predicate circumscription. For instance, Przymusinski's well-known query answering algorithm based on MILO-resolution [15] involves a separate treatment for variable predicates that can be bypassed by our transformation.

A Appendix

To illustrate the transformation, we consider the classic examples of Tweety the bird and Sam the penguin.

Example 1
Let our knowledge base be the following:

$$\mathcal{T} = b(Tweety) \wedge (\forall x.b(x) \wedge \neg ab(x) \supset f(x)).$$

(The intended meaning of b is "is a bird", of ab is "is abnormal", and of f is "flies"). The list P of the predicates to be minimized contains only ab, while the remaining predicates b and f are the varying predicates Z. It is well-known that $CIRC(\mathcal{T}; P; Z) \models f(Tweety)$ (see, for example [6, 9]). The transformation we defined previously yields the following formula:

$$\begin{aligned}\Gamma(\mathcal{T}; f(Tweety); P; Z) \;\; = \;\; & [(ab, b, f) \equiv \neg(ab'', b'', f'')] \wedge \\ & [(u \wedge (ab', b', f' \equiv true)) \;\vee\; (\neg u \wedge \Theta(ab, b, f, ab', b', f'))],\end{aligned}$$

where $\Theta(ab, b, f, ab', b', f') = [(\mathcal{T} \supset f(Tweety)) \;\vee\; (\mathcal{T}[ab'; b', f'] \wedge (ab' < ab))]$. Substituting for our notation, we get

$$\Gamma(\mathcal{T}; f(Tweety); P; Z) = C_1 \wedge C_2 \wedge C_3 \wedge [(u \wedge D_1) \vee (\neg u \wedge (D_2 \vee D_3))],$$

where

$$C_1 \;\; = \;\; (\forall x.b(x) \equiv \neg b''(x)),$$

$$
\begin{aligned}
C_2 &= (\forall x.f(x) \equiv \neg f''(x)), \\
C_3 &= (\forall x.ab(x) \equiv \neg ab''(x)), \\
D_1 &= (\forall x.b'(x)) \wedge (\forall x.f'(x)) \wedge (\forall x.ab'(x)), \\
D_2 &= [b(Tweety) \wedge (\forall x.b(x) \wedge \neg ab(x) \supset f(x))] \supset f(Tweety), \\
D_3 &= [b'(Tweety) \wedge (\forall x.b'(x) \wedge \neg ab'(x) \supset f'(x)) \wedge \\
&\qquad (\forall x.ab'(x) \supset ab(x)) \wedge (\exists x.\neg ab'(x) \wedge ab(x))].
\end{aligned}
$$

We show that $CIRC(\Gamma(\mathcal{T}; f(Tweety); P; Z)) \models \neg u$: Suppose that there exists a minimal model M of $\Gamma(\mathcal{T}; f(Tweety); P; Z)$ such that $M \models u$; clearly $M \not\models D_2 \vee D_3$, otherwise the interpretation N of $\mathcal{T}$ which is equivalent to M everywhere but not in u ($N \not\models u$) would be a model of $\Gamma(\mathcal{T}; f(Tweety); P; Z)$ such that $N < M$, but this contradicts the hypothesis that M is a minimal model of $\Gamma(\mathcal{T}; f(Tweety); P; Z)$. Notice that $M \not\models D_2$ implies $M \models ab(Tweety)$. Moreover $M \models D_1$, since M is a model of $\Gamma(\mathcal{T}; f(Tweety); P; Z)$. We can define a new interpretation L of $\Gamma(\mathcal{T}; f(Tweety); P; Z)$ such that L is equivalent to M everywhere but not in u or in ab' such that $L \models \neg u \wedge \forall x.\neg ab'(x)$. Since $M \models D_1$, it follows that $L \models (\forall x.b'(x)) \wedge (\forall x.f'(x))$. Since $M \models ab(Tweety)$, it follows that $L \models ab(Tweety)$. It follows from this $L \models D_3$; this means that L is a model of $\Gamma(\mathcal{T}; f(Tweety); P; Z)$. But $L < M$, which contradicts the hypothesis that M is minimal.

Playing devil's advocate, let us see if we can prove more from $CIRC(\Gamma(\mathcal{T}; \mathcal{F}; P; Z))$ than what is available from $CIRC(\mathcal{T}; P; Z)$. Suppose we add $walk(George)$ to $\mathcal{T}$ (let $\mathcal{T}' = \mathcal{T} \wedge walk(George)$) and try to prove $walk(Tweety)$ from $CIRC(\Gamma(\mathcal{T}'; walk(Tweety); P; Z))$, where $P = ab$ and $Z = b, f, walk$. Assuming $Tweety \neq George$, choose the following extensions for $ab, b, f, walk$: $\forall x.ab(x) \equiv \mathit{false}$, $\forall x.b(x) \equiv x = Tweety$, $\forall x.f(x) \equiv x = Tweety$, and $\forall x.walk(x) \equiv x = George$. This induces a minimal model of $\mathcal{T}'$ in which $walk(Tweety)$ is false. Therefore, $\Theta(ab, b, f, walk, ab', f', walk')$ in the reduction of $\mathcal{T}'$ is false in that case; hence, if u is true, if $ab', b', f', walk' \equiv \mathit{true}$ holds, and if we select the double primed predicates appropriately, we obtain a minimal model of $\Gamma(\mathcal{T}'; walk(Tweety); P; Z)$ in which u is true. Therefore, $CIRC(\Gamma(\mathcal{T}'; walk(Tweety); P; Z)) \not\models \neg u$. $\Box$

Example 2

We will consider the classic example of *Sam* the penguin. Our knowledge base is the following:

$$
\begin{aligned}
\mathcal{T} &= p(Sam) \wedge (\forall x.p(x) \supset b(x)) \wedge \\
&\quad (\forall x.b(x) \wedge \neg ab_1(x) \supset f(x)) \wedge \\
&\quad (\forall x.p(x) \wedge \neg ab_2(x) \supset \neg f(x))
\end{aligned}
$$

(Here p is intended to mean "is a penguin", ab_1 and ab_2 grasp different kinds of abnormality, and b, f are as above). P contains ab_1, ab_2, while Z is p, b, f. $CIRC(\mathcal{T}; P; Z) \models$

$f(Sam)$ does not hold (indeed, the extensions $\forall x.ab_1(x) \equiv x = Sam$, $\forall x.ab_2(x) \equiv \mathit{false}$, $\forall x.p(x) \equiv x = Sam$, $\forall x.b(x) \equiv x = Sam$, $\forall x.f(x) \equiv \mathit{false}$ induce a $P;Z$-minimal model of $\mathcal{T}$ in which $f(Sam)$ is not satisfied). Our transformation yields the following formula:

$$\begin{aligned}\Gamma(\mathcal{T}; f(Sam); P; Z) \quad = \quad & [(ab_1, ab_2, p, b, f) \equiv \neg(ab_1'', ab_2'', p'', b'', f'')] \wedge \\ & [(u \wedge (ab_1', ab_2', p', b', f' \equiv \mathit{true})) \ \vee \ (\neg u \wedge \Theta(P, Z, P', Z'))],\end{aligned}$$

where

$$\Theta(P, Z, P', Z') = [(\mathcal{T} \supset f(Sam)) \ \vee \ (T[ab_1', ab_2'; p', b', f'] \wedge (ab_1', ab_2' < ab_1, ab_2))].$$

That is,

$$\Gamma(\mathcal{T}; f(Sam); P; Z) = C_1 \wedge C_2 \wedge C_3 \wedge C_4 \wedge C_5 \wedge [(u \wedge D_1) \vee (\neg u \wedge (D_2 \vee D_3))],$$

where

$$\begin{aligned}
C_1 \quad &= \quad (\forall x.p(x) \equiv \neg p''(x)),\\
C_2 \quad &= \quad (\forall x.b(x) \equiv \neg b''(x)),\\
C_3 \quad &= \quad (\forall x.f(x) \equiv \neg f''(x)),\\
C_4 \quad &= \quad (\forall x.ab_1(x) \equiv \neg ab_1''(x)),\\
C_5 \quad &= \quad (\forall x.ab_2(x) \equiv \neg ab_2''(x)),\\
D_1 \quad &= \quad (\forall x.p'(x)) \wedge (\forall x.b'(x)) \wedge (\forall x.f'(x)) \wedge (\forall x.ab_1'(x)) \wedge (\forall x.ab_2'(x)),\\
D_2 \quad &= \quad [p(Sam) \wedge (\forall x.p(x) \supset b(x)) \wedge (\forall x.b(x) \wedge \neg ab_1(x) \supset f(x)) \wedge \\
&\qquad (\forall x.p(x) \wedge \neg ab_2(x) \supset \neg f(x))] \supset f(Sam)\\
D_3 \quad &= \quad p'(Sam) \wedge (\forall x.p'(x) \supset b'(x)) \wedge (\forall x.b'(x) \wedge \neg ab_1'(x) \supset f'(x)) \wedge \\
&\qquad (\forall x.p'(x) \wedge \neg ab_2'(x) \supset \neg f'(x)) \wedge \\
&\qquad (\forall x.ab_1'(x) \supset ab_1(x)) \wedge (\forall x.ab_2'(x) \supset ab_2(x)) \wedge \\
&\qquad [(\exists x.\neg ab_1'(x) \wedge ab_1(x)) \vee (\exists x.\neg ab_2'(x) \wedge ab_2(x))].
\end{aligned}$$

We show that $\mathit{CIRC}(\Gamma(\mathcal{T}; f(Sam); P; Z)) \not\models \neg u$ by identifying a minimal model M of $\Gamma(\mathcal{T}; f(Sam); P; Z)$ such that $M \models u$. M is characterized as follows :

$$M \models u \wedge (\forall x.p'(x) \wedge b'(x) \wedge f'(x) \wedge ab_1'(x) \wedge ab_2'(x))$$

that is, M satisfies D_1,

$$\begin{aligned}M \quad \models \quad & p(Sam) \wedge b(Sam) \wedge \neg f(Sam) \wedge ab_1(Sam) \wedge \neg ab_2(Sam) \wedge \\ & \neg p''(Sam) \wedge \neg b''(Sam) \wedge f''(Sam) \wedge \neg ab_1''(Sam) \wedge ab_2''(Sam)\end{aligned}$$

and

$$M \models (\forall x. x \neq Sam \supset [\neg p(x) \wedge \neg b(x) \wedge \neg f(x) \wedge \neg ab_1(x) \wedge \neg ab_2(x) \wedge \\ p''(x) \wedge b''(x) \wedge f''(x) \wedge ab_1''(x) \wedge ab_2''(x)]),$$

that is, every predicate from the original theory $\mathcal{T}$ is false in M for every constant different from Sam. Note that M is a model of $\Gamma(\mathcal{T}; F; P; Z)$, because $M \models C_1 \wedge C_2 \wedge C_3 \wedge C_4 \wedge C_5 \wedge u \wedge D_1$. Minimality of M can be seen as follows. Assume that M is not minimal, that is there is a model N of $\Gamma(\mathcal{T}; f(Sam); P; Z)$ such that $N < M$. N can only decrease the predicates ab_1', ab_2', p', b', f', because decreasing any other predicate forces another predicate to increase. This implies that $N \not\models D_1$, therefore $N \models D_2 \vee D_3$ must be true; but $M \models \neg D_2$, and since the predicates in D_2 have the same extension in M and N, $N \models \neg D_2$ holds; hence, $N \models D_3$ must hold. Since $ab_1(Sam)$ is the only ground atom of $ab_1(x), ab_2(x)$ that is true in M and hence also in N, we infer from the last conjunct of D_3 that $N \models \neg ab_1'(Sam)$. Since $D_3 \models b'(Sam) \wedge (\forall x. b'(x) \wedge \neg ab_1'(x) \supset f'(x))$, it follows that $N \models f'(Sam)$. Since $D_3 \models p'(Sam) \wedge (\forall x. p'(x) \wedge \neg ab_2'(x) \supset \neg f'(x))$, it follows that $M \models ab_2'(Sam)$. But $D_3 \models (\forall x. ab_2'(x) \supset ab_2(x))$; thus it follows that $N \models ab_2(Sam)$. This leads to a contradiction, however, since $M \models \neg ab_2(Sam)$ and hence $N \models \neg ab_2(Sam)$ must hold. Therefore N cannot exist, and M is truly a minimal model of $\Gamma(\mathcal{T}; f(Sam); P; Z)$. Thus $CIRC(\Gamma(\mathcal{T}; f(Sam); P; Z)) \not\models \neg u$ is proved.

We close the example considering a small change to our knowledge base $\mathcal{T}$. Suppose the bird $Tweety$ from Example 1 appears in Sam's world, and let $\mathcal{T}' = \mathcal{T} \wedge b(Tweety)$. Under the unique-names assumption ($Tweety \neq Sam$), can we still prove that $Tweety$ flies under minimization of abnormality while the other predicates vary ? Indeed, under this assumption $CIRC(\mathcal{T}'; P; Z) \models f(Tweety)$ holds. This can be seen as follows. There exist models of $\mathcal{T}'$ in which $ab_1(Tweety)$ and $ab_2(Tweety)$ are false, and in each of them $p(Tweety)$ is false. Thus for inferencing $f(Tweety)$ under circumscription, $\mathcal{T}'$ essentially reduces to the knowledge base $\mathcal{T}$ in Example 1 where ab and ab_1 need to be collapsed. Therefore, one can show that $CIRC(\Gamma(\mathcal{T}'; f(Tweety); P; Z)) \models \neg u$ holds along a line of argumentation similar to that in Example 1. $\Box$

Acknowledgment

The authors would like to thank the anonymous referees for their helpful comments and suggestions to improve a previous draft of this paper. They are also grateful to Alex Leitsch for hints on references.

References

[1] G. Bossu and P. Siegel. Saturation, Nonmonotonic Reasoning and the Closed

World Assumption. *Artificial Intelligence*, 25:13–63, 1985.

[2] J. de Kleer and K. Konolige. Eliminating the Fixed Predicates from a Circumscription. *Artificial Intelligence*, 39:391–398, 1989.

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

[4] D. W. Etherington. *Reasoning with Incomplete Information.* Morgan Kaufmann Publishers, Inc., Los Altos, 1988.

[5] M. Gelfond, H. Przymusinksa, and T. Przymusinski. On the Relationship Between Circumscription and Negation as Failure. *Artificial Intelligence*, 38:75–94, 1989.

[6] M. Gelfond and H. Przymusinska. Negation as Failure: Careful Closure Procedure. *Artificial Intelligence*, 30:273–287, 1986.

[7] M. L. Ginsberg. A Circumscriptive Theorem Prover. *Artificial Intelligence*, 39:209–230, 1989.

[8] J. Hintikka. Reductions in the Theory of Types. *Acta Philosophica Fennica*, 8:61–115, 1955.

[9] V. Lifschitz. Closed-World Databases and Circumscription. *Artificial Intelligence*, 27:229–235, 1985.

[10] V. Lifschitz. Computing Circumscription. In *Proceedings IJCAI-85*, pages 121–127, 1985.

[11] V. Lifschitz. Pointwise Circumscription. In *Proceedings AAAI-86*, pages 406–410, 1986.

[12] J. McCarthy. Circumscription – A Form of Non-Monotonic Reasoning. *Artificial Intelligence*, 13:27–39, 1980.

[13] J. McCarthy. Applications of Circumscription to Formalizing Common-Sense Knowledge. *Artificial Intelligence*, 28:89–116, 1986.

[14] J. Minker. On Indefinite Data Bases and the Closed World Assumption. In *Proceedings of the* 6^{th} *Conference on Automated Deduction (CADE)*, pages 292–308, 1982.

[15] T. Przymusinski. An Algorithm to Compute Circumscription. *Artificial Intelligence*, 38:49–73, 1989.

[16] R. Reiter. On Closed-World Databases. In H. Gallaire and J. Minker, editors, *Logic and Data Bases*, pages 55–76. Plenum Press, New York, 1978.

[17] J.S. Schlipf. Decidability and Definability with Circumscription. *Annals of Pure and Applied Logic*, 35:173–191, 1987.

[18] J. van Benthem and K. Doets. Higher Order Logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, Vol.I*, chapter I.4, pages 275–329. D. Reidel Publishing Company, 1983.

[19] A. Yahya and L. Henschen. Deduction in Non-Horn Databases. *Journal of Automated Reasoning*, 1(2):141–160, 1985.

[20] L.Y. Yuan and C.H. Wang. On Reducing Parallel Circumscription. In *Proceedings AAAI-88*, pages 460–464, 1988.