In: Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence (IJCAI-93), R. Bajcsy ed, Morgan Kaufman, 1993, pp. 634–639.

Curb Your Theory ! A circumscriptive approach for inclusive interpretation of disjunctive information*

Thomas Eiter, Georg Gottlob
Christian Doppler Lab for Expert Systems
Technical University of Vienna
Paniglgasse 16, A-1040 Wien, Austria
(eiter|gottlob)@vexpert.dbai.tuwien.ac.at

Yuri Gurevich
Electrical Engineering and CS Department
University of Michigan
Ann Arbor, MI 48109–2122, USA

Abstract

We introduce *curbing*, a new nonmonotonic technique of commonsense reasoning that is based on model minimality but unlike circumscription treats disjunction inclusively. A finitely axiomatized first-order theory T is transformed to a formula $Curb(T)$ whose set of models is defined as the smallest collection of models which contains all minimal models of T and which is closed under formation of minimal upper bounds with respect to inclusion. We first give an intuitive definition of *Curb* in third-order logic and then show how *Curb* can be equivalently expressed in second-order logic. We study the complexity of inferencing from a curbed propositional theory and present a PSPACE algorithm for this problem. Finally, we address different possibilities to approximate the curb of a theory.

1 Introduction

Circumscription [McCarthy, 1980] looks to many as one of the most promising principles for formalizing commonsense reasoning. However, as recently pointed out by Raymond Reiter [Reiter, 1992], it runs into problems in connection with disjunctive information. The minimality principle of circumscription often enforces the *exclusive* interpretation of a disjunction $a \vee b$ by adopting the models in which either a or b is true but not both. There are situations in which an *inclusive* interpretation is desired and seems more natural. Consider the following example due to Reiter. Suppose you throw a coin into an area which is divided into a black and a white field. Circumscription applied to

$$black\_field(coin) \vee white\_field(coin)$$

excludes that the coin falls into both fields and tells you that the coin is either in the white or in the black field. This is certainly not satisfactory. An extension of this example is even more impressive. Imagine a handful of coins thrown onto a chessboard; circumscription tells us that no coin touches both a black and a white field. It is not clear whether any of the well-known variants of circumscription (cf. [Etherington, 1988]) can suitably handle inclusive disjunction of positive information.

In this paper we suggest an approach to tackle the problem pointed out by Reiter. We present the *curb* method, which is a generalization of circumscription able to handle inclusive disjunctions of positive information properly. Our method relies on the new concept of minimal upper bound models. Intuitively, a minimal upper bound model corresponds to a minimal model for the inclusive interpretation of disjuncts. Using this concept, we develop the notion of a "good" model of a first-order theory. Although the concept is more involved than circumscription, we show that "good" models can be captured like circumscriptive models in second-order logic.[1] Furthermore, inferencing from the good models of a propositional sentence is feasible in quadratic space; we show that for a reasonable approximation of the good models this problem is no harder than inference from circumscriptive models.

The rest of this paper is organized as follows. Section 2 considers additional examples and introduces informally the concept of good models. Section 3 provides a formal definition and a logical description. Section 4 examines aspects of computational complexity, and Section 5 addresses approximation issues. Section 6 reviews related work and concludes the paper.

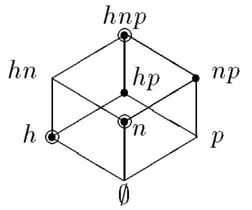
2 Good models

Let us consider two additional scenarios in which an inclusive interpretation of disjunction is desirable. Models are represented by their positive atoms.

Example 1: Suppose there is a man in a room with a painting, which he hangs on the wall if he has a hammer and a nail. It is known that the man has a hammer or a nail or both. This scenario is represented by the theory T_1 in Figure 1. The desired models are h, n, and hnp, which are encircled. Circumscribing T_1 by minimizing all variables yields the two minimal models h and n (see Figure 1). Since p is false in the minimal models, circumscription tells us that the man does not hang the painting up. One might argue that the variable p should not be minimized but fixed when applying circumscription. However, starting with the model of T_1 where h, n

*This is a short version not providing detailed proofs of all results. Full proofs and more results will appear in an extended report.

[1] By using a second-order predicate constant. However, it is quantification which counts for the order of a formula [van Benthem and Doets, 1983].

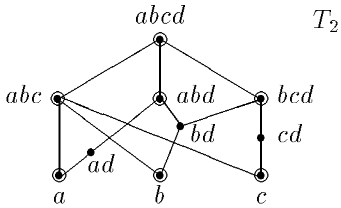


$$T_1 = \{ h \vee n, \; (h \,\&\, n) \rightarrow p \}$$

Figure 1: The hammer-nail-painting example

and p are all true and then circumscribing with respect to h and p while keeping p true, we obtain the models hp and np, which are not very intuitive. If we allow p to vary in minimizing h and n, the outcome is the same as for minimizing all variables. On the other hand, the model hnp seems plausible. This model corresponds to the inclusive interpretation of the disjunction $h \vee n$. □

Example 2: Suppose you have invited some friends to a party. You know for certain that one of Alice, Bob, and Chris will come, but you don't know whether Doug will come. You know in addition the following habits of your friends. If Alice and Bob go to a party, then Chris or Doug will also come; if Bob and Chris go, then Alice or Doug will go. Furthermore, if Alice and Chris go, then Bob will also go. This is represented by theory T_2 in Figure 2. Now what can you say about who will



$$T_2 = \{ a \vee b \vee c, \; (a \,\&\, b) \rightarrow (c \vee d), \; (b \,\&\, c) \rightarrow (a \vee d), \; (a \,\&\, c) \rightarrow b \}$$

Figure 2: The party example

come to the party? Look at the models of T_2 in Figure 2. Circumscription yields the minimal models a, b, and c, which interpret the clause $a \vee b \vee c$ exclusively in the sense that it is minimally satisfied. However, there are other plausible models. For example, abc. This model embodies an inclusive interpretation of a and b within $a \vee b \vee c$; it is also minimal in this respect. abd is another model of this property. Similarly, bcd is a minimal model for an inclusive interpretation of b and c. The models ad, bd, and cd are not plausible, however, since a scenario in which Doug and only one of Alice, Bob or Chris are present does not seem well-supported. □

In the light of these examples, the question arises how circumscription can be extended to work satisfactory. An important insight is that such an extension must take disjunctions of positive events seriously and allow inclusive (hence nonminimal) models, even if such models contain positive information that is not contained in any minimal model. On the other hand, the fruitful principle of minimality should not be abandoned by adopting models that are intuitively not concise. Our idea is the synthesis of both: adopt the minimal inclusive models. That is, adopt for minimal models M_1, M_2 any model M which includes both M_1 and M_2 and is a minimal such model; in other words, M is a *minimal upper bound* (*mub*) for M_1 and M_2.

To illustrate, in Example 1 hnp is a mub for h and n (notice that hn is not a model), and in Example 2 abc is a mub for a and c; abd is another one, so several mub's can exist. In order to capture general inclusive interpretations, mub's of arbitrary collections $M_1, M_2, M_3, \ldots$ of minimal models are adopted.

It appears that in general not all "good" models are obtainable as mub's of collections of minimal models. The good model $abcd$ in Example 2 shows this. It is, however, a mub of the good models a and bcd (as well as of abc and abd). This suggests that not only mub's of collections of minimal models, but mub's of any collection of good models should also be good models.

Our approach to extend circumscription for inclusive interpretation of disjunctions is thus the following: adopt as good models the least set of models which contains all circumscriptive (i.e. minimal) models and which is closed under including mub's. Notice that this approach yields in Examples 1 and 2 the sets of intuitively good models, which are encircled in Figs. 1 and 2.

3 Capturing good models formally

We give in this section a formal semantical definition of good models of a first-order sentence, and we provide a logical sentence which describes the good models. We assume ZFC, i.e. Zermelo-Fraenkel Set Theory with the Axiom of Choice, as a standard metamathematical frame [Suppes, 1972].

As for circumscription, we need a language of higher-order logic (cf. [van Benthem and Doets, 1983]) over a set of predicate and function symbols, i.e. variables and constants of finite arity $n \geq 0$ of suitable type. Recall that 0-ary predicate symbols are identified with propositional symbols.

We use set notation for predicate membership and inclusion. $q \in p$ means that $p(q)$ is true, $p \subseteq q$ that p implies q, and $p \subset q$ that p strictly implies q. $(\exists p \theta q)(\ldots)$ denotes $\exists p(p \theta q \;\&\; \ldots)$ where θ is $\in, \subseteq$, or $\subset$; similarly $(\forall p \theta q)(\ldots)$ denotes $\forall p(p \theta q \rightarrow \ldots)$. A sentence is a formula in which no variable occurs free; it is of order $n+1$ if the order of any quantified symbol occurring in it is $\leq n$ [van Benthem and Doets, 1983].

A structure M consists of a nonempty set $|M|$ and an assignment $\mathcal{I}(M)$ of predicates, i.e. relations (resp. functions), of suitable type over $|M|$ to the predicate (resp. function) constants. The object assigned to constant C, i.e. the extension of C in M, is denoted by $[\![C]\!]_M$ or simply C if this is clear from the context. Equality is interpreted as identity. A model for a sentence φ is any structure M such that φ is true in M ($M \models \varphi$).

Let $\mathbf{p} = p_1, \ldots, p_n$ be a list of first-order predicate constants and $\mathbf{z} = z_1, \ldots, z_m$ a list of first-order predicate or function constants disjoint with $\mathbf{p}$. For any structure M, let $\mathcal{M}^M_{\mathbf{p};\mathbf{z}}$ be the class of structures M' such that $|M| = |M'|$, and $[\![C]\!]_M = [\![C]\!]_{M'}$ for every constant C not occurring in $\mathbf{p}$ or $\mathbf{z}$. The pre-order $\leq^M_{\mathbf{p};\mathbf{z}}$ on $\mathcal{M}^M_{\mathbf{p};\mathbf{z}}$ is defined by $M_1 \leq^M_{\mathbf{p};\mathbf{z}} M_2$ iff $[\![p_i]\!]_{M_1} \subseteq [\![p_i]\!]_{M_2}$ for all $1 \leq i \leq n$. The pre-order $\leq_{\mathbf{p};\mathbf{z}}$ is the union of all $\leq^M_{\mathbf{p};\mathbf{z}}$

over all structures. We write $\mathcal{M}_{\mathbf{p}}^{M}$ etc. if $\mathbf{z}$ is empty; $\leq_{\mathbf{p}}^{M}$ and $\leq_{\mathbf{p}}$ are partial orders on $\mathcal{M}_{p}^{M}$ resp. all structures.

The circumscription of $\mathbf{p}$ in a first-order sentence $\varphi(\mathbf{p},\mathbf{z})$ with $\mathbf{z}$ floating is the second-order sentence [Lifschitz, 1985]

$$\varphi(\mathbf{p},\mathbf{z}) \;\&\; \neg\exists \mathbf{p}',\mathbf{z}'(\varphi(\mathbf{p}',\mathbf{z}') \;\&\; (\mathbf{p}' \subset \mathbf{p}))$$

which will be denoted by $Circ(\varphi(\mathbf{p},\mathbf{z}))$ ($\mathbf{p}$ and $\mathbf{z}$ will be always presupposed). Here $\mathbf{p}'$, $\mathbf{z}'$ are lists of predicate and function variables matching $\mathbf{p}$ and $\mathbf{z}$ and $\mathbf{p} \subset \mathbf{p}'$ stands for $(\mathbf{p}' \subseteq \mathbf{p})$ & $(\mathbf{p}' \neq \mathbf{p})$, where $(\mathbf{p}' \subseteq \mathbf{p})$ is the conjunction of all $(p_i' \subseteq p_i)$, $1 \leq i \leq n$. The following is a straightforward consequence of the definitions.

Proposition 3.1 [Lifschitz, 1985] $M \models Circ(\varphi(\mathbf{p},\mathbf{z}))$ *iff* M *is* $\leq_{\mathbf{p};\mathbf{z}}$*-minimal among the models of* $\varphi(\mathbf{p},\mathbf{z})$.

3.1 formal definition and logical description

We formally define the concept of a "good" model as follows. First define the property that a set of models is closed under minimal upper bounds.

Definition 3.1 *Let* $\varphi(\mathbf{p},\mathbf{z})$ *be a first-order sentence. A set* $\mathcal{M}$ *of models of* $\varphi(\mathbf{p},\mathbf{z})$ *is* $\leq_{\mathbf{p};\mathbf{z}}$*-closed iff, for every* $\mathcal{M}' \subseteq \mathcal{M}$ *and any model* M *of* $\varphi(\mathbf{p},\mathbf{z})$*, if* M *is* $\leq_{\mathbf{p};\mathbf{z}}$*-minimal among the models of* $\varphi(\mathbf{p},\mathbf{z})$ *which satisfy* $M' \leq_{\mathbf{p};\mathbf{z}} M$ *for all* $M' \in \mathcal{M}'$ *then* $M \in \mathcal{M}$.

Clearly the set of all models is closed. Further, every closed set must contain all $\leq_{\mathbf{p};\mathbf{z}}$-minimal models of $\varphi(\mathbf{p},\mathbf{z})$ (let $\mathcal{M}' = \emptyset$); the empty set is closed iff $\varphi(\mathbf{p},\mathbf{z})$ has no minimal model. We define goodness as follows.

Definition 3.2 *A model* M *of* $\varphi(\mathbf{p},\mathbf{z})$ *is good with respect to* $\mathbf{p};\mathbf{z}$ *iff* M *belongs to the least* $\mathbf{p};\mathbf{z}$*-closed set of models of* $\varphi(\mathbf{p},\mathbf{z})$.

Notice that good models only exist if a unique smallest closed set exists. The latter is immediately evident from the following characterization of goodness.

Proposition 3.2 *A model* M *of* $\varphi(\mathbf{p},\mathbf{z})$ *is good with respect to* $\mathbf{p};\mathbf{z}$ *iff* M *belongs to the intersection of all* $\mathbf{p};\mathbf{z}$*-closed sets.*

In the rest of this section we show how to capture goodness by a logical sentence $Curb(\varphi(\mathbf{p},\mathbf{z});\mathbf{p},\mathbf{z})$ whose models are precisely the good models of $\varphi(\mathbf{p},\mathbf{z})$. Similar to circumscription, $\mathbf{p}$ are the minimized predicates (here under the *inclusive* interpretation of disjunction), $\mathbf{z}$ are the floating predicates, and all other predicates are fixed.

For ease of exposition, we restrict ourselves here to the simplest case where $\mathbf{p}$ contains a single unary predicate p (hence the domain D of p is $|M|^1 = |M|$) and $\mathbf{z}$ is empty. The generalization to arbitrary $\mathbf{p}$ and $\mathbf{z}$ is straightforward.

We consider in the following an arbitrary but fixed assignment to all constant symbols except p over a fixed universe, i.e. we consider structures in some fixed $\mathcal{M}_p^M$. We assume that $p^1, q^1, r^1, \ldots$ range over the extensions $[\![p]\!]_{M'}$ in the models M' of $\varphi(p)$ from $\mathcal{M}_p^M$, and that $p^2, q^2, r^2, \ldots$ range over families of such extensions. Technically, this can be asserted by adding $\varphi(p^1)$, $\varphi(q^1)$ etc. and $\forall q \in p^2(\varphi(q))$, $\forall q \in q^2(\varphi(q))$ etc. in suitable places.

We now express a mub of the family p^2 as the formal expression Mub, which says that q^1 is a minimal upper bound of p^2.

$$\begin{aligned} Mub(p^2,q^1) \;\; =_{df} \;\; & (\forall r^1 \in p^2)(r^1 \subseteq q^1) \;\& \\ & \forall s^1(s^1 \subset q^1 \rightarrow (\exists t^1 \in p^2)(t^1 \not\subseteq s^1)). \end{aligned}$$

The formula $Cl(p^2)$ says that p^2 is closed under minimal upper bounds.

$$Cl(p^2) =_{df} \forall r^2 \forall q^1\,(r^2 \subseteq p^2 \;\&\; Mub(r^2,q^1) \;\rightarrow\; q^1 \in p^2).$$

<u>Remark</u>. $M \models Cl(p^2)$ & $Circ(\varphi(p^1))$ implies $p^1 \in p^2$ as $Mub(\emptyset, p^1)$ holds. □

Using $Cl(p^2)$, we define a formula $Good(p^2)$ as the smallest $\mathbf{p};\mathbf{z}$-closed set, which intuitively captures the good models of $\varphi(p)$ from $\mathcal{M}_p^M$.

$$Good(p^2) =_{df} Cl(p^2) \;\&\; \forall r^2(Cl(r^2) \rightarrow p^2 \subseteq r^2).$$

$Good(p^2)$ is true within $\mathcal{M}_p^M$ for a unique p^2.

Lemma 3.3 *For every* M,

$$M \models \forall p^2\,\forall q^2(Good(p^2) \;\&\; Good(q^2) \rightarrow p^2 = q^2).$$

Proof. Assume this is false for some M, i.e. $M \models Good(p^2)$ & $Good(q^2)$ for some p^2, q^2 such that $p^2 \neq q^2$. Let $r^2 = p^2 \cap q^2$. Clearly $M \models Cl(r^2)$. Since $r^2 \subset p^2$ or $r^2 \subset q^2$, it holds that $M \models Cl(r^2)$ & $\neg(p^2 \subseteq r^2)$, say. This implies $M \not\models Good(p^2)$. This gives the desired contradiction. □

Using $Good(p^2)$, we define the *curb of* $\varphi(p)$ as the sentence $Curb(\varphi(p);p)$:

$$Curb(\varphi(p);p) =_{df} \varphi(p) \;\&\; \exists p^2(Good(p^2) \;\&\; p \in p^2).$$

Theorem 3.4 *A model* M *of* $\varphi(p)$ *is good with respect to* p *iff* $M \models Curb(\varphi(p);p)$.

It should be clear how $Curb(\varphi(p);p)$ generalizes to $Curb(\varphi(\mathbf{p};\mathbf{z});\mathbf{p},\mathbf{z})$ for arbitrary $\mathbf{p}$ and $\mathbf{z}$.

3.2 second-order definition of $Curb(\varphi(\mathbf{p},\mathbf{z});\mathbf{p},\mathbf{z})$

Notice that $Curb(\varphi(p);p)$ is a third-order sentence (due to $\exists p^2(\cdots)$ and $\forall r^2(\cdots)$ within $Good(p^2)$), while $Circ(\varphi(p))$ is second-order. We present in this section an equivalent definition in terms of a second-order sentence $Curb^*(\varphi(p);p)$ which characterizes the good models of $\varphi(p)$. A second-order sentence $Curb^*(\varphi(\mathbf{p},\mathbf{z});\mathbf{p},\mathbf{z})$ for arbitrary $\mathbf{p}$ and $\mathbf{z}$ is obtained analogously.

The rationale of what we are going to do in this section is as follows. We first define the metalinguistic notion of α-goodness for all ordinal numbers α. Roughly speaking, a model of a first-order theory T is α-good if it can be obtained from the minimal models by iteratively adding α times all minimal upper bounds. The notion of α-goodness of a model can be expressed in a two-sorted second-order logic whose one sort are the ordinal numbers. We show that in each structure, it is sufficient to consider only those ordinals α for which it holds that $\|\alpha\| \leq \|D\|$, where D is the domain of p. But precisely these ordinals are isomorphic to well-orderings of parts of D and are thus expressible in second-order logic. Our last step thus consists in converting the two-sorted definition of goodness into plain second-order logic (with no specially interpreted predicates) by replacing ordinals with well-orderings.

Definition 3.3 $\text{good}(p^1) \leftrightarrow_{df} \exists\alpha.\, \alpha\text{-good}(p^1)$, *where*

$$0\text{-good}(p^1) \quad \leftrightarrow_{df} \quad Circ(\varphi(p^1)), \quad \text{and for } \alpha > 0,$$
$$\alpha\text{-good}(p^1) \quad \leftrightarrow_{df} \quad \exists q^2\,(Mub(q^2, p^1)\ \&\ (\forall r^1 \in q^2)(\exists \beta < \alpha).\, \beta\text{-good}(r^1)).$$

For $\alpha \geq 0, \text{strictly-}\alpha\text{-good}(p^1) \leftrightarrow_{df} \alpha = \inf\{\beta : \beta\text{-good}(p^1)\}$.

We show that this notion of goodness is equivalent to the one captured by $Curb(\varphi(p); p)$.

Lemma 3.5 *A model M of $\varphi(p)$ is good with respect to p iff* $\text{good}([\![p]\!]_M)$ *holds true.*

Proof. (Sketch) Let $q^2 = \{p^1 : \text{good}(p^1)\}$. (i) We show that $M \models Cl(r^2)$ implies $q^2 \subseteq r^2$. Prove by induction on α, that for every r^2 such that $M \models Cl(r^2)$, r^2 contains every α-good r^1. (ii) We show that $M \models Cl(q^2)$. Suppose $r^2 \subseteq q^2$ and $M \models Mub(r^2, q^1)$. Define $\beta = \sup\{\alpha : r^1 \in r^2,\ r^1 \text{ is strictly-}\alpha\text{-good}\}$. It follows that q^1 is $(\beta+1)$-good, hence $q^1 \in q^2$. (i) and (ii) imply $M \models Good(q^2)$. The result follows from Lemma 3.3 and Theorem 3.4. □

An important observation is that goodness of any p^1 is equivalent to α-goodness for an α which depends on the domain D of M.

Lemma 3.6 *For every domain D there exists an ordinal α_D such that* $\text{good}(p^1) \leftrightarrow \alpha_D\text{-good}(p^1)$ *for every p^1. Moreover, $\|\alpha_D\| \leq \|D\|$.*

Proof. (Sketch) Show by induction on α, that if p^1 is strictly-α-good then $\|p^1\| \geq \|\alpha\|$. The lemma follows. □

Now consider the extension of second-order logic where it is allowed to use ordinals of cardinality $\leq \|D\|$. Formally, one can think about a two-sorted structure. Greek letters range over ordinals. We define recursively a second-order predicate constant $G(\alpha, p^1)$, which captures α-goodness within the logical language.

$$G(\alpha, p^1) \quad \leftrightarrow_{df} \quad (\forall q^1 \subset p^1)\,(\exists \beta < \alpha)\ \exists r^1 (G(\beta, r^1)\ \&\ \neg(r^1 \subseteq q^1)\ \&\ (r^1 \subseteq p^1)).$$

In particular, $G(0, p^1) \leftrightarrow Circ(\varphi(p^1))$. If $G(\beta, r^1)$ is equivalent to β-goodness of r^1 for $\beta < \alpha$, then $G(\alpha, p^1)$ says that there is no $q^1 \subset p^1$ such that every $(< \alpha)$-good subset r^1 of p^1 is also a subset of q^1.

Lemma 3.7 *$G(\alpha, p^1)$ is true iff p^1 is α-good.*

Proof. By induction on α. The case $\alpha = 0$ is trivial. Suppose $\alpha > 0$.

"$\leftarrow$": Suppose p^1 is α-good. By definition of α-goodness, there is a collection q^2 that witnesses this fact. In particular, $Mub(q^2, p^1)$ holds. Suppose that q^1 satisfies $q^1 \subset p^1$. By the definition of $Mub(q^2, p^1)$, p^1 is a minimal extension of p such that $r^1 \subseteq p^1$ for every $r^1 \in q^2$. Thus there exists $r^1 \in q^2$ such that $r^1 \not\subseteq q^1$. Since $r^1 \in q^2$, r^1 is β-good for some $\beta < \alpha$. By the hypothesis, $G(\beta, r^1)$ is true. Hence $G(\alpha, p^1)$ is true.

"$\rightarrow$": Suppose $G(\alpha, p^1)$ is true. Take $q^2 = \{r^1 : r^1 \subseteq p^1,\ r^1 \text{ is } \beta\text{-good for some } \beta < \alpha\}$. Check that $Mub(q^2, p^1)$ holds true. By contradiction suppose $Mub(q^2, p^1)$ fails. Clearly p^1 is an upper bound of q^2. Consequently, p^1 is not a minimal upper bound. Hence there exists q^1 such that $q^1 \subset p^1$ and for every $r^1 \in q^2$, $r^1 \subseteq q^1$. By the hypothesis, for every such r^1 there exists a $\beta < \alpha$ such that $G(\beta, r^1)$ is true. Thus q^1 contradicts $G(\alpha, p^1)$. It follows from the definition of α-goodness that p^1 is α-good. □

Corollary 3.8 *p^1 is* good *iff there exists an ordinal α such that $\|\alpha\| \leq \|D\|$ and $G(\alpha, p^1)$ is true.*

Ordinals of cardinality $\leq \|D\|$ can be expressed as well-orderings of a part of D within second-order logic, however. A well-ordering R on a set S is a total ordering which is well-founded, i.e. R satisfies

$$(\forall X \subseteq S)[X \neq \emptyset \rightarrow (\exists x \in X)\,(\forall y \in X).\, R(x, y)]$$

(every nonempty subset X of S has the first element.) The following proposition is well-known.

Proposition 3.9 *Let S be a set. For every ordinal α such that $\|\alpha\| \leq \|S\|$ there is a well-ordering $R(x, y)$ of a subset $E \subseteq S$ which is isomorphic to α, i.e. $(E, R) \cong (\{\beta : \beta < \alpha\}, \leq)$.*

Now let $WO(R)$ be a formula which states that R is a well-ordering of a part of the domain, and let $IS(R_1, R_2)$ be a formula which says that the well-ordering R_2 is a strict initial segment of the well-ordering R_1, hence $\alpha(R_2) < \alpha(R_1)$ where $\alpha(R)$ is the ordinal represented by R. Formulating WO and IS in second-order logic is easy. Using WO and IS, the predicate constant G^2, which is informally equivalent to G, is defined as follows.

$$G^2(R, q) \quad \leftrightarrow_{df} \quad WO(R)\ \&\ \varphi(q)\ \&\ (\forall q^1 \subset q).\ \exists R_1\, \exists r^1\, (G^2(R_1, r^1)\ \&\ IS(R, R_1)\ \&\ \neg(r^1 \subseteq q^1)\ \&\ (r^1 \subseteq p^1)).$$

Now $Curb^*(\varphi(p); p)$ is defined by

$$Curb^*(\varphi(p); p) =_{df} \exists R.\, G^2(R, p).$$

Theorem 3.10 *$M \models Curb^*(\varphi(p); p)$ iff M is a model of $\varphi(p)$ which is good with respect to p.*

Proof. M is a good model of $\varphi(p) \leftrightarrow \text{good}([\![p]\!]_M)$ holds true $\leftrightarrow$ there exists an α such that $\|\alpha\| \leq \|D\|$ and $G(\alpha, p)$ holds $\leftrightarrow$ there exists a well-ordering R of $E \subseteq D$, where D is the domain of p, such that $G(\alpha(R), p)$ is true $\leftrightarrow$ there exists an R such that $G^2(R, p)$ is true $\leftrightarrow$ $Curb^*(\varphi(p); p)$ is true. □

Remark: $Curb^*(\varphi(p); p)$ involves a second-order predicate constant G^2; it can be eliminated by a more involved construction, at least over all infinite structures.

4 Computational complexity

An important aspect of any reasoning method is, of course, its computational complexity (cf. [Garey and Johnson, 1979] for basic concepts and definitions). It is clear that our method is in the full first-order case highly undecidable, just as any of the well-known methods in nonmonotonic reasoning. We give here a more detailed account of the propositional case. Notice that in this case, a structure M is a truth-value assignment to the propositional variables and $\mathbf{p}$ and $\mathbf{z}$ are sets of propositional variables. In particular, we consider the inference problem.

Problem:	CURB-INFERENCE
Instance:	Disjoint sets P, Z of propositional variables and propositional formulas F, G.
Question:	Does $Curb(F; P, Z)$ imply G ?

This problem can be easily shown to be coNP-hard, as here $Curb(F; P, Z)$ is consistent iff F is consistent; hence, it is intractable. It may seem at first glance that the problem needs exponential space, in particular since checking whether a truth-value assignment is a good model seems difficult. The straightforward algorithm, trying to generate this model starting from the minimal models by iteratively including minimal upper bounds, is clearly exponential in both time and space. However, the algorithm **GOOD** in Table 1 below shows that model checking is feasible in polynomial space. The correctness of **GOOD** follows from Theorem 5.2 of the next section, which implies that we can limit ourselves to consider minimal upper bounds of *pairs* of models. **GOOD** can be straightforwardly implemented such that its body uses only space linear in the input size. Furthermore, it is easily shown by induction that the recursive depth is bounded by $|P|$. Consequently, the algorithm runs in quadratic space.

```
ALGORITHM GOOD(M, F, P, Z) : boolean;
  input: truth-value assignment M, propositional formula
         F, sets P, Z of propositional variables, P ∩ Z = ∅.
  output: "true" iff M is wrt to P; Z a good model of F.

    if (M ⊭ F) then return false;
    minimal := true;
    for each M1 <P;Z M do
      if (M1 ⊨ F) then minimal := false;
    if minimal then return true;
    for each M1, M2 <P;Z M, M1 ≰P;Z M2, M2 ≰P;Z M1 do
      if GOOD(M1, F, P, Z) then
        if GOOD(M2, F, P, Z) then
        begin mub := true;
          for each M3 with M1, M2 <P;Z M3 <P;Z M do
            if GOOD(M3, F, P, Z) then mub := false;
          if mub then return true;
        end;
    return false;
```

Table 1: Algorithm for the good model property

Theorem 4.1 CURB-INFERENCE *is feasible in quadratic space.*

Proof. $Curb(F(P, Z); P, Z) \not\models G$ iff there exists a model M of F which is good with respect to $P; Z$, such that $M \not\models G$. Cycling through all truth assignments to find such an M using **GOOD** is clearly possible in quadratic space; hence the result. □

5 Approximation

In this section we briefly address possibilities to approximate the full set of good models by a subset. A suitable approximation may even be more intuitive than the full set of good models. The assumption that a "good" set of models accepted by an agent is closed under minimal upper bounds is, of course, an idealization. Which inclusive models the agent accepts will depend on the particular situation; finding a general rationale for this is cleary difficult. We consider here the two possibilities of building inclusive models within bounded depth and from a bounded number of disjuncts.

5.1 bounded depth

The first approximation is to limit iterated inclusion of minimal upper bounds. Informally, we choose only the models that are α-good for some α such that $\|\alpha\| \leq \|\delta\|$, where the ordinal δ is a limit on the depth in building minimal upper bounds. The results of Section 3.2 make it clear that the good models within depth δ can be captured by a second-order sentence $Curb^\delta(\varphi(\mathbf{p}, \mathbf{z}); \mathbf{p}, \mathbf{z})$ as long as δ is second-order expressible. Notice that circumscription appears as the case $\delta = 0$, i.e. $Curb^0(\varphi(\mathbf{p}, \mathbf{z}); \mathbf{p}, \mathbf{z})$ is equivalent to $Circ(\varphi(\mathbf{p}, \mathbf{z}))$.

Concerning computational complexity, the inference problem is in the propositional case for finite constant δ as easy (and as hard) as circumscription.

Theorem 5.1 *The δ-bounded version of* CURB-INFERENCE, *i.e. deciding if $Curb^\delta(F(P, Z); P, Z)$ implies G, for δ a finite constant is complete for the Π_2^P level of the polynomial hierarchy* [Meyer and Stockmeyer, 1972].

Proof. We sketch a proof of membership in Π_2^P. $Curb^\delta(F(P, Z); P, Z) \not\models G$ iff there is a δ-good model M of F such that $M \not\models G$. Guess M and $\leq |P|$ witnesses $M_1, \ldots, M_m$ for M (i.e., good models such that M is a minimal upper bound), and for each M_i witnesses etc. to a recursive depth $\leq \delta$. The guess consists of $\leq |P|^\delta$ truth-value assignments, and is thus polynomial in the input size. Checking whether M is a minimal upper bound of $M_1, \ldots, M_m$ etc. is possible in polynomial time with an NP oracle. Consequently, deciding if $Curb^\delta(F(P, Z); P, Z) \not\models G$ is in $\mathrm{NP}^{\mathrm{NP}} = \Sigma_2^P$, which implies that "Inference" is in Π_2^P. Π_2^P-hardness for $\delta = 0$ (i.e. circumscription) was shown in [Eiter and Gottlob, 1991]. The construction used in the proof can be easily generalized to arbitrary finite constant δ. □

5.2 bounded disjunctions

Another attempt is to limit the cardinality of model sets from which minimal upper bounds are formed. Intuitively, this corresponds to limiting the number of inclusively interpreted disjuncts by a cardinal $\kappa > 0$. The concept of closed$_\kappa$ set is defined by adding in the definition of closed set the condition "$\|\mathcal{M}'\| \leq \kappa$"; goodness$_\kappa$ is the relative notion of goodness.

Clearly, goodness$_1$ is equivalent to circumscription. For $\kappa \geq 2$, we obtain over finite structures (i.e. $|M|$ is finite) the following result.

Theorem 5.2 *Over finite structures, for every $\kappa \geq 2$ a model of $\varphi(\mathbf{p}, \mathbf{z})$ is good$_\kappa$ with respect to $\mathbf{p}; \mathbf{z}$ iff it is good with respect to $\mathbf{p}; \mathbf{z}$.*

Proof. We show this for $\mathbf{p} = p$, p unary, and empty $\mathbf{z}$ under above notational conventions; the proof is easily generalized. Show by induction on $\|p^1\|$ that every α-good p^1 is α-good$_2$. The case $\|p^1\| = 0$ is trivial. For $\|p^1\| > 0$, let $q^2 = \{q^1 : q^1 \subset p^1, \mathrm{good}(q^1)\}$.

$Mub(q^2, p^1)$ holds. If $q^2 = \emptyset$, the statement trivially holds. In the other case, let $r^1 \in q^2$ be maximal with respect to $\subseteq$. As $|M|$ is finite, such an r^1 exists. Let $s^1 \in q^2$ such that $s^1 \not\subseteq r^1$. Such an s^1 must exist. Then $Mub(\{r^1, s^1\}, p^1)$ holds, and by the induction hypothesis r^1 and s^1 are good$_2$; hence, p^1 is good$_2$ and the statement holds. Clearly, every p^1 which is good$_\kappa$ is good. Hence the result follows. □

Theorem 5.2, which fails for arbitrary structures, implies a dichotomous result on the expressivity of κ-bounded disjuncts: Either we get only the minimal models or all models obtainable by unbounded disjuncts. Limiting *simultaneously* the number of disjuncts and and the depth in building minimal upper bounds can be used to cut down the set of good models.

6 Related work and conclusion

Inclusive interpretation of disjunctive models has been investigated in logic programming [Ross and Topor, 1988; Sakama, 1989; Ross, 1989; Chan, 1991].

The Disjunctive Database Rule (DDR) [Ross and Topor, 1988] has been proposed to allow cautious derivation of negative literals from a disjunctive database. The DDR allows all models in Examples 1 and 2, hence also the unintuitive ones. Moreover, it depends on syntactical representation [Chan, 1991], which however is customary for semantics of logic programming. Thus, the DDR is basically different from our method.

A more sophisticated approach that allows also to deal with negative clauses was introduced in [Sakama, 1989; Chan, 1991] by the equivalent concepts "Possible Models Semantics" (PMS) and "Possible Worlds Semantics" (PWS). This approach has recently been generalized to cover Negation by Failure [Sakama and Inoue, 1993]. In Examples 1 and 2, PMS and PWS coincide with *curb*. However, if the clause $h \vee n \vee p$ is added to T_1, which has no effect on the models, then PMS and PWS adopt all models. Hence, PWS and PMS are syntax-dependent and basically differ from our method.

Another approach to treat disjunction inclusively is the weak well-founded semantics (WWF) for disjunctive logic programs in [Ross, 1989]. In case of negation-free programs, this semantics coincides with the DDR [Ross, 1989], which implies syntax-dependency of WWF. Hence, WWF is basically different from *curb*.

In this paper we presented a new approach to nonmonotonic commonsense reasoning that seems to be more appropriate than circumscription in many cases, namely, when disjunction of positive information is naturally interpreted in an inclusive fashion. Our method of curbing theories differs significantly from all previous approaches to treat disjunction inclusively. It is syntax independent and yields the more intuitive models. We have shown that *Curb* is second-order definable and have derived some relevant complexity results. We also have fostered two ways of approximating the curb of a theory.

We believe that this new approach deserves further investigations. On the one hand, it is tempting to find new and better algorithms for inferencing under *curb* or its approximations. On the other hand, the inclusive interpretation of disjunction is not always desired. Sometimes it seems that a hybrid approach which interprets certain predicates (or certain connectives) inclusively and others exclusively is more appropriate, cf. [Ross, 1989; Sakama, 1989; Chan, 1991]. Our ongoing research deals with all these topics.

Acknowledgments

The authors thank C. Sakama for clarifying comments.

References

[Chan, 1991] E.P.F. Chan. A Possible Worlds Semantics for Disjunctive Databases. *IEEE Transactions on Data and Knowledge Engineering*, 1991. To appear.

[Eiter and Gottlob, 1991] Thomas Eiter and Georg Gottlob. Propositional Circumscription and Extended Closed World Reasoning are $\mathbf{\Pi}_2^P$-complete. *Theoretical Computer Science*, vol. 113, June 1993.

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

[Garey and Johnson, 1979] Michael Garey and David S. Johnson. *Computers and Intractability – A Guide to the Theory of NP-Completeness*. 1979.

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

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

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

[Meyer and Stockmeyer, 1972] A.R. Meyer and L.J. Stockmeyer. The equivalence problem for regular expressions with squaring requires exponential time. In *Proc. 13^{th} Ann. Symp. on Switching and Automata Theory*, pp. 125–129. IEEE Computer Society, 1972.

[Reiter, 1992] R. Reiter. Talk given at the Christian Doppler Lab for Expert Systems, March 31, 1992.

[Ross and Topor, 1988] K.A. Ross and R.W. Topor. Inferring Negative Information From Disjunctive Databases. *J. of Automated Reasoning*, 4(2):397–424, 1988.

[Ross, 1989] K. Ross. The Well-Founded Semantics for Disjunctive Logic Programs. In *Proceedings DOOD-89*, pp. 352–369, 1989.

[Sakama and Inoue, 1993] C. Sakama and K. Inoue. Negation in Disjunctive Logic Programs. In *Proceedings ICLP-93*, Budapest, Hungary, June 1993. MIT-Press.

[Sakama, 1989] C. Sakama. Possible Model Semantics for Disjunctive Databases. In *Proceedings 1st Intl. Conf. DOOD-89*, pp. 337–351, Kyoto Japan, 1989.

[Suppes, 1972] P. Suppes. *Axiomatic Set Theory*. Dover Pub., New York, 1972.

[van Benthem and Doets, 1983] Johan van Benthem and Kees Doets. Higher Order Logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, Vol.I*, chapter I.4, pp. 275–329. 1983.