On Open Defaults

Vladimir Lifschitz

Stanford University, Stanford, California 94305, USA
and
University of Texas, Austin, Texas 78712, USA

Abstract

In Reiter's default logic, the parameters of a default are treated as metavariables for ground terms. We propose an alternative definition of an extension for a default theory, which handles parameters as genuine object variables. The new form of default logic may be preferable when the domain closure assumption is not postulated. It stands in a particularly simple relation to circumscription. Like circumscription, it can be viewed as a syntactic transformation of formulas of higher order logic.

1 Introduction

Default logic [Reiter, 1980] is one of the most expressive and most widely used nonmonotonic formalisms. In one respect, however, the main definition of default logic, that of an *extension*, is not entirely satisfactory.

Recall that a default

$$\alpha : \beta_1, \ldots, \beta_m / \gamma \tag{1}$$

is *open* if it contains free variables, and *closed* otherwise. The concept of an extension is defined in two steps: It is first introduced, by means of a fixpoint construction, for default theories without open defaults, and then generalized to arbitrary default theories. Since interesting cases usually involve open defaults, the second step is crucial. Its main idea is that a default with free variables has the same meaning as the set of all its ground instances.[1] In other words, free variables in a default are viewed as metavariables for ground terms.

In many cases, this treatment of free variables makes the effect of a default surprisingly weak. Consider the default theory with one axiom $P(a)$ and one default,

$$: \neg P(x) / \neg P(x). \tag{2}$$

[1] The actual reduction of the general case to the case of closed defaults is more complex, because it involves the Skolemization of all axioms and of the consequents of all defaults—a detail which is irrelevant for this discussion.

Intuitively, this default expresses that $P(x)$ is assumed to be false whenever possible. We can expect that it will allow us to prove

$$\forall x(P(x) \equiv x = a). \tag{3}$$

But all that this default gives is the literals $\neg P(t)$ for the ground terms t different from a. Notice that the behavior of circumscription[2] is quite different. The circumscription of P in $P(a)$, which expresses the same idea of making $P(x)$ false whenever possible, is equivalent to (3).

In applications to formalizing commonsense reasoning, this weakness of open defaults is sometimes undesirable. Consider the following example.[3] Suppose that, for any two blocks x and y, the default is that x is not on y. If there is no evidence that any blocks are on the block B_1 then, for each individual block B_i, we will be able to conclude that it is not on B_1. But we may be unable to justify the conclusion that B_1 is clear, in the sense that there are no blocks on B_1. Indeed, the set of conclusions

$$\neg on(B_1, B_1), \ldots, \neg on(B_n, B_1)$$

is weaker than the universally quantified formula

$$\forall x \neg on(x, B_1),$$

unless we accept the "domain closure assumption"

$$\forall x(x = B_1 \vee \ldots \vee x = B_n),$$

expressing that every block is represented by one of the constants B_i. The domain closure assumption is sometimes unacceptable: We may be unable or unwilling to design the language in such a way that each object in the domain of reasoning be represented by a ground term.

In this paper we propose a modification of default logic in which free variables in defaults are treated as genuine object variables, rather than metavariables for ground terms. The new form of default logic is better suited for formalizing default reasoning in the absence of the domain closure assumption. Another reason why this modification of Reiter's system can be of interest is that it stands in a particularly simple relation to circumscription and consequently sheds some light on the important and difficult problem of connecting various approaches to default reasoning.

As our starting point, we take a characterization of extensions for the case of closed defaults based on [Guerreiro and Casanova, 1990]. This characterization uses a fixpoint construction which is similar to Reiter's, except that the fixpoints in question are *classes of models*, rather than sets of sentences. It is not particularly surprising that extensions can be defined in such a manner, because there is a natural correspondence between classes of models and sets of sentences: For any class V of models, we can consider its *theory*, that is, the set of sentences that are true in all models from V. But for our purposes the

[2] The definition of circumscription can be found in [McCarthy, 1986] or [Lifschitz, 1985].

[3] [McCarthy, 1980], Section 8, Remark 2; [Poole, 1987b], Example 3. See also the discussion of "default reasoning in an open domain" (the note to Example A5) in [Lifschitz, 1989a].

transition from formulas to models is essential. In the realm of formulas, the only way to refer to objects in the domain of reasoning is through their syntactic representations, that is, ground terms. In the realm of models, we can talk about elements of the universe directly.

In Section 2 we review, for motivation and further reference, the fixpoint constructions from [Reiter, 1980] and [Guerreiro and Casanova, 1990]. The new definition of an extension is given in Section 3. Then we discuss its relation to traditional default logic (Section 4) and to circumscription (Section 5). In Section 6 we outline an extension of the formalism in which some object, function or predicate constants are treated as "fixed." In Section 7 we show that the new form of default logic, like circumscription, can be viewed as a syntactic transformation of formulas of higher order logic. Related work is surveyed in Section 8. Most proofs are deferred to the appendix.

2 Extensions According to Reiter and Guerreiro–Casanova

According to [Reiter, 1980], a *default theory* (in a given first order language) is a pair (D, W), where D is a set of defaults of the form (1), and W is a set of sentences. A default theory (D, W) is *closed* if all defaults from D are closed. Let (D, W) be a closed default theory. For any set of sentences S, consider the smallest set of sentences S' which includes W, is closed under classical logic, and satisfies the condition:

($*$) For any default (1) from D, if $\alpha \in S'$ and $\neg\beta_1, \ldots, \neg\beta_m \notin S$ then $\gamma \in S'$.

This set S' is denoted by $\Gamma(S)$. S is said to be an *extension* for (D, W) if it is a fixpoint of Γ, that is, if $\Gamma(S) = S$.

The Guerreiro–Casanova approach to closed default theories can be described as follows. For any class V of structures for the language of (D, W), let $Th(V)$ stand for the theory of V—the set of sentences which are true in all structures from V. Let V' be the largest class of models of W which satisfies the condition:

($**$) For any default (1) from D, if $\alpha \in Th(V')$ and $\neg\beta_1, \ldots, \neg\beta_m \notin Th(V)$ then $\gamma \in Th(V')$.

This largest V' always exists:

Proposition 1. *The union of all classes V' of models of W which satisfy ($**$) satisfies ($**$) also.*

This class V' is denoted by $\Sigma(V)$. As essentially established in [Guerreiro and Casanova, 1990], extensions can be characterized in terms of the fixpoints of Σ:

Proposition 2. *A set of sentences is an extension for (D, W) if and only if it has the form $Th(V)$ for some fixpoint V of Σ.*

Intuitively, if we think of the extension $Th(V)$ as a set of "beliefs," then V is the class of "worlds" that are possible according to these beliefs.

Our objective is to generalize the definition of Σ to open defaults. Then the characterization of extensions given by Proposition 2 will serve as the basis for a new definition of an extension.

Consider an open default

$$\alpha(x) : \beta_1(x), \ldots, \beta_m(x) / \gamma(x), \tag{4}$$

where x is a list of variables. Instead of replacing the variables from x by ground terms, as in Reiter's logic, we want to replace them by arbitrary elements of the universe, or, more precisely, by symbols that serve as "names" of arbitrary elements of the universe. The condition $(**)$ will turn into something like this: For any default (4) from D and any tuple ξ of "names," if $\alpha(\xi) \in Th(V')$ and $\neg\beta_1(\xi), \ldots, \neg\beta_m(\xi) \notin Th(V)$ then $\gamma(\xi) \in Th(V')$.

There is a problem, however, with this idea: Different structures from V have, generally, different universes, and "names" appropriate for one structure from V will generally make no sense for another. In order to generalize the condition $(**)$ to open defaults, we need to modify it so that all structures from V have the same universe.

3 Default Logic with a Fixed Universe

Let (D, W) be a default theory, not necessarily closed, and let U be a nonempty set. By a *world* we understand any model of W with the universe U. Extend the language of (D, W) by object constants representing all elements of U; these constants will be called *names*. For any set of worlds V, $Th^*(V)$ is the set of sentences in the extended language which are true in all worlds from V. (Thus $Th(V)$ is the set of the sentences from $Th^*(V)$ that do not contain names.)

For any set of worlds V, consider the largest set of worlds V' which satisfies the condition:

$(***)$ For any default (4) from D and any tuple of names ξ, if $\alpha(\xi) \in Th^*(V')$ and $\neg\beta_1(\xi), \ldots, \neg\beta_m(\xi) \notin Th^*(V)$ then $\gamma(\xi) \in Th^*(V')$.

This largest V' always exists:

Proposition 3. *The set of sets V' satisfying $(***)$ is closed under union.*

This set V' will be denoted by $\Delta(V)$. The operator Δ is the "fixed universe" counterpart of Σ. Notice that Δ depends not only on the default theory (W, D), but also on the universe U.

A *U-extension* for (D, W) is any set of sentences of the form $Th(V)$, where V is a fixpoint of Δ. Notice that U-extensions, just like extensions in Reiter's logic, consist of sentences in the language of (D, W); they do not contain names.

It is clear that the U-extensions for a given default theory (D, W) are completely determined by the cardinality of U. For any positive integer n, let $Card_n$ be a standard sentence expressing that there are exactly n objects; for instance, we can take $Card_1$ to be $\forall xy(x = y)$. By $Card_U$ we denote $Card_n$ if the cardinality of U is n, and $\{\neg Card_1, \neg Card_2, \ldots\}$ if U is infinite. Any U-extension contains W and $Card_U$, and is closed under classical logic.

As an example, consider the operator Δ for the default theory discussed in the introduction:

$$W = \{P(a)\},\ D = \{: \neg P(x)/\neg P(x)\}. \tag{5}$$

We will see in Section 5 that, if M is a model of (3) with the universe U, then $\{M\}$ is a fixpoint of Δ, and, conversely, every fixpoint of Δ has this form. The U-extension $Th(V)$ corresponding to a fixpoint V of Δ is the deductive closure of (3) and $Card_U$. Thus, for every U, there is exactly one U-extension.

In Reiter's logic, we can say that a sentence A is a *consequence* of (D, W) if A is in the intersection of all extensions for (D, W). For instance, the consequences of the default theory (5) are the logical consequences of $P(a)$. Let us say that A is a *fixed universe consequence (F-consequence)* of (D, W) if A is in the intersection of all U-extensions for (D, W) over all nonempty sets U. In other words, an F-consequence is a sentence which is true in every world which belongs to some fixpoint of Δ. It is clear that the set of F-consequences of (D, W), like the set of its consequences, contains W and is closed under classical logic. In case of (5), the F-consequences are the logical consequences of (3), which is exactly what we wanted to achieve.

If D is empty, then the condition $(***)$ is trivially true, so that, for every V, $\Delta(V)$ is the set of all worlds, and this set is the only fixpoint of Δ. Consequently, the only U-extension of $(\emptyset, W)$ consists of the sentences that are true in all models of W with the universe U. It follows that the F-consequences of this theory are the sentences logically entailed by W.

4 Relation to Reiter's Logic

Often, as in the example (5), a default theory has more F-consequences than consequences. But sometimes this is the other way around. Consider the following example:

$$W = \{P(a)\},\ D = \{: \neg P(b)/\neg P(b)\}. \tag{6}$$

The extension of this theory in Reiter's logic includes $\neg P(b)$. If, on the other hand, the cardinality of U is 1, then the U-extension of this theory includes $P(b)$ (which is a logical consequence of W and $Card_1$), and does not include $\neg P(b)$. Consequently, $\neg P(b)$ is not an F-consequence of (6). The sentence $a \neq b$ is another consequence of (6) which is not an F-consequence.

This example shows also that the two versions of default logic are not equivalent even for closed default theories. For closed defaults, $(***)$ turns into $(**)$, except that both V and V' are assumed to consist of models with a fixed universe U; this distinction is responsible for the difference between consequences and F-consequences.

If, however, the language is propositional, then the choice of U becomes inessential, because models of a propositional theory are simply mappings of propositional symbols into truth values. Consequently, for propositional default theories, Δ coincides with Σ. It follows then by Proposition 2, that, in the propositional case, U-extensions are identical to extensions, and F-consequences are identical to consequences.

Here are some other cases when the two forms of default logic are equivalent:

Proposition 4. *Let (D, W) be a closed default theory. If all models of W have the same finite cardinality, then the F-consequences of (D, W) are identical to its consequences.*

Proposition 5. *Let (D, W) be a closed default theory with at most countably many object, function and predicate constants. If all models of W are infinite, then the F-consequences of (D, W) are identical to its consequences.*

Proposition 4 is applicable when W includes both the domain closure assumption and the unique names assumption. Proposition 5 is applicable when the universe of discourse includes some infinite domain, for instance, natural numbers.

As an illustration of Proposition 5, we can consider a modification of (6) in which axioms expressing the existence of infinitely many objects are added to W. Such a theory will have both $\neg P(b)$ and $a \neq b$ among its F-consequences. Even simpler, this effect can be achieved by assuming two distinct objects:

$$W = \{P(a), \exists xy(x \neq y)\},\ D = \{: \neg P(b)/\neg P(b)\}.$$

If U is a singleton, then the set of worlds for this theory is empty, and $\emptyset$ is the only fixpoint of Δ. Consequently, the only U-extension for a singleton U is the set of all sentences, and the argument made at the beginning of this section regarding (6) does not go through.

5 Normal Defaults Without Prerequisites

In this section we assume that D is

$$\{: \alpha(x)/\alpha(x)\}, \tag{7}$$

so that it consists of a single default, which is a "normal default without a prerequisite." For example, each of the theories (5), (6) satisfies this condition. Intuitively, the effect of the default (7) is to "maximize" α. Proposition 6 below gives a characterization of the U-extensions for such theories which makes this claim precise.

The following notation will be used: For any world M, α^M stands for the set of tuples of names ξ such that $\alpha(\xi)$ is true in M.

Let us say that a world M is *α-maximal* if there is no world M' such that α^M is a proper subset of $\alpha^{M'}$. In other words, an α-maximal world is a model of W with the universe U such that no other model of W with the same universe has more x's satisfying $\alpha(x)$. In particular, if $\alpha(x)$ is $\neg P(x)$, then an α-maximal world is a model in which the extent of P is *minimal* in the sense corresponding to the circumscription of P with all object, function and predicate constants allowed to vary.

In the following proposition we assume that the set of worlds is nonempty, that is, W has at least one model with the universe U.

Proposition 6. *For any default theory with the set of defaults (7), every fixpoint of the corresponding operator Δ is an equivalence class of the set of α-maximal worlds relative to the relation $\alpha^M = \alpha^{M'}$. Conversely, each of these equivalence classes is a fixpoint of Δ.*

This theorem shows that the fixpoints of Δ correspond to the extents of α in α-maximal worlds. The claims made above about the default theories (5) and (6) can be easily justified using Proposition 6. For (5), the α-maximal worlds are the models of (3), and each equivalence class consists of a single model.

Proposition 6 implies that the F-consequences of a theory with the set of defaults (7) can be characterized as the sentences that are true in all α-maximal worlds for all U. In particular:

Corollary. *A sentence $B(P)$ is an F-consequence of the default theory*

$$(\{A(P)\}, \{: \neg P(x)/\neg P(x)\})$$

if and only if $B(P)$ is logically entailed by the circumscription of P in $A(P)$ with all object, function and predicate constants allowed to vary.

In the corresponding result for Reiter's default logic ([Etherington, 1987a], Theorem 2), W is required to include the domain closure assumption and a form of the unique names assumption. The new approach to open defaults makes these conditions redundant.

6 Default Logic with Fixed Constants

As we have seen, circumscription with all constants varied is a special case of default logic with a fixed universe. In this section, a generalization of this form of default logic is defined, which is related to circumscription with some object, function and predicate constants fixed. This generalization provides a new perspective on the relationship between circumscription and default logic. We do not propose it as a serious candidate for AI use.

A *default theory with fixed constants* is a triple (D, W, C), where D and W are as in the standard definition of a default theory, and C is a subset of object, function and predicate constants. The symbols from C are the *fixed* constants of the theory; the remaining constants are *varied*. A default theory (D, W) corresponds to the case when C is empty. By L_C we denote the first order language whose object, function and predicate constants are the members of C; thus L_C is a sublanguage of the language of the theory.

The definitions of the operator Δ and of a U-extension for this generalization of default theories are the same as in Section 3, except that now we take U to be a structure for the language L_C (rather than merely a universe), and define a *world* to be any model of W obtained from U by assigning interpretations to the varied constants. It is clear that the U-extensions for a default theory with fixed constants remain the same if U is replaced by an isomorphic structure.

An *F-consequence* of a default theory with fixed constants is a sentence that belongs to all its U-extensions for all structures U.

Consider, for instance, the default theory with

$$W = \{\forall x(Q(x) \supset P(x))\},\ D = \{: \neg P(x)/\neg P(x)\},\ C = \{Q\}. \qquad (8)$$

Let U be a structure for the language whose only nonlogical constant is Q (that is, U is a universe along with its subset representing Q). A world is defined by an interpretation of P that makes the sentence $\forall x(Q(x) \supset P(x))$ true, that is, by a subset of the universe that contains the set representing Q. The only fixpoint of Δ consists of one world, in which the extent of P is the same as the extent of Q. The F-consequences of (8) are the sentences logically entailed by

$$\forall x(P(x) \equiv Q(x)). \tag{9}$$

The definition of an α-maximal world (Section 5), Proposition 6 and its proof apply to default theories with fixed constants, without any changes whatsoever. The counterpart of Corollary to Proposition 6 can be stated as follows:

Proposition 7. *A sentence $B(P)$ is an F-consequence of the theory*

$$(\{A(P)\}, \{: \neg P(x)/\neg P(x)\}, C),$$

where C does not include P, if and only if $B(P)$ is logically entailed by the circumscription of P in $A(P)$ with the fixed constants C.

We see that default theories with fixed constants subsume a rather general form of circumscription. The fact that the F-consequences of (8) are the sentences logically entailed by (9) can serve as an illustration of this theorem, because (9) is the result of circumscribing P in $\forall x(Q(x) \supset P(x))$ with Q fixed.

7 Default Logic as a Syntactic Transformation

Our next goal is to show that the definition of a U-extension can be expressed by a higher order logical formula, so that the modification of default logic proposed in this paper, like circumscription, can be viewed as a syntactic transformation of sentences. Since the definition of a U-extension involves not only worlds, but also *sets* of worlds, we will need not only second order, but also *third order* variables.

Let (D, W, C) be a default theory with fixed constants, in a language which has finitely many object, function and predicate constants, and with both D and W finite. Let Z be the list of all varied constants (that is, the constants that do not belong to C). We will explicitly show the occurences of the varied constants in formulas, so that an arbitrary sentence will be written as $F(Z)$, and a formula with the list of parameters x will be written as $F(Z, x)$. The set W will be identified with the conjunction of its elements and written as $W(Z)$.

Take a list of variables z of the same length as Z, such that if the i-th member of Z is an object constant then the i-th member of z is an object variable, and if the i-th member of Z is a function (predicate) constant then the i-th member of z is a function (predicate) variable of the same arity. Given a structure U for the language L_C, the structures obtained from U by assigning interpretations to the varied constants can be identified with combinations of values of the variables z in U. In particular, the models of W that are obtained in this way ("worlds") correspond to the values of z for which $W(z)$ is true.

Let v be a variable such that $v(z)$ is a well-formed formula (so that v is third order if Z contains at least one function or predicate constant). Values of v can be identified with sets of structures obtained from U by assigning interpretations to the varied constants. Then the values of v satisfying the condition

$$\forall z[v(z) \supset W(z)] \tag{10}$$

represent sets of worlds. More generally, for any formula $F(Z,x)$, the formula

$$\forall z[v(z) \supset F(z,x)]$$

expresses that $F(Z,x)$ is true in each structure from v. We will denote this formula by $F^{\forall}(v,x)$, so that (10) will be written as $W^{\forall}(v)$. Similarly, the formula

$$\exists z[v(z) \wedge F(z,x)],$$

expressing that $F(Z,x)$ is true in at least one structure from v, will be denoted by $F^{\exists}(v,x)$.

Using this notation, it is easy to encode the relation between sets of worlds V, V' expressed by $(***)$ (Section 3). If d is a default

$$\alpha(Z,x) : \beta_1(Z,x), \ldots, \beta_m(Z,x)/\gamma(Z,x)$$

from D, then by $d(v,v')$ we denote the formula

$$\forall x[\alpha^{\forall}(v',x) \wedge \beta_1^{\exists}(v,x), \ldots, \beta_m^{\exists}(v,x) \supset \gamma^{\forall}(v',x)].$$

Then the condition $(***)$ is expressed by the conjunction

$$\bigwedge_{d \in D} d(v,v').$$

Since $\Delta(V)$ is the union of all sets V' of worlds satisfying $(***)$, it is the represented by the abbreviation

$$\Delta(v) = \lambda z \exists v'[W^{\forall}(v') \wedge \bigwedge_{d \in D} d(v,v') \wedge v'(z)].$$

Now we can define the "default logic operator" DL. By $\mathrm{DL}(D,W,C)$ we denote the sentence

$$\exists v[\Delta(v) = v \wedge v(Z)].$$

A structure obtained from U by assigning values to the varied constants is a model of this sentence if and only if it belongs to some fixpoint of the corresponding operator Δ. Consequently, the class of models of $\mathrm{DL}(D,W,C)$ is the union of all fixpoints of the operators Δ corresponding to all structures U. We conclude:

Proposition 8. *A sentence B is an F-consequence of (D,W,C) if and only if B is logically entailed by $DL(D,W,C)$.*

8 Related Work

The correspondence between extensions and classes of models on which our main definition is based was first used by Etherington [1987b], although his construction is less transparent than that of [Guerreiro and Casanova, 1990].

The possibility of introducing "nonground instances of defaults" is discussed by Poole [1987b] for his formulation of default logic, equivalent to the normal subset of Reiter's system. In [Poole, 1987a], the distinction between varied and fixed predicates is added to that formalism.

Przymusinski [1989] defends the use of non-Herbrand models in logic programming. Since logic programs can be viewed as a special case of default theories [Bidoit and Froidevaux, 1988], this issue is related to the problem of open defaults.

The counterpart of an open default theory in autoepistemic logic is an autoepistemic theory with “quantifying-in.” One way of defining a semantics for such theories is proposed by Levesque [1990]. Konolige [1989] discusses another approach; he also studies the problem of reducing circumscription to autoepistemic logic without assuming domain closure. Introspective circumscription [Lifschitz, 1989b] is a system analogous to autoepistemic logic, in which unrestricted quantification is allowed. Like the formalism of this paper, it subsumes some forms of “minimizing” circumscription.

Levesque's formulation of autoepistemic logic is based on a mapping into a monotonic system, like the characterization of our default logic given in Section 7. One difference is that Levesque uses a simple mapping into a rather involved modal logic, and we use a rather complicated transformation whose target language is classical.

9 Conclusion

Recent research shows that the main ideas of different nonmonotonic formalisms are more compatible with each other than one might think. Perhaps we will not have to select any one of the classical approaches ([McCarthy, 1980], [McDermott and Doyle, 1980], [Reiter, 1980]) as a basis for the nonmonotonic formalism of the future, and reject the others; we may be able to combine the advantages of different models of nonmonotonic reasoning in the same system.

Konolige [1988] noticed, for instance, that modal nonmonotonic languages, such as autoepistemic logic, are close in their expressiveness to the language of default logic. Unfortunately, the problem of finding precise equivalence results for Konolige's translation turned out to be quite difficult. Several “groundedness” conditions have been proposed in order to filter out the autoepistemic extensions that have no counterparts in default logic. The first attempt [Konolige, 1988] was unsuccessful, and this has led to the invention of “supergrounded” and “robust” extensions (see ([Marek and Truszczyński, 1989], Sections 2.3 and 3). Further work in this direction is described in [Marek and Truszczyński, 1990].

The available reductions of defaults to epistemic formulas are not completely satisfactory. What makes the situation even more complicated is the fact that autoepistemic logic is merely one point in the whole spectrum of nonmonotonic modal systems introduced in [McDermott, 1982], and possibly not the best for AI applications [Shvarts, 1990]. On the other hand, the availability of an epistemic modal operator is an attractive feature of a knowledge representation language, in connection with the problem of representing integrity constraints [Reiter, 1988]. Hopefully, future research will lead to the invention of an elegant modal system, such that default logic will be linked to it by means of Konolige's translation or a similar mechanism.

The results of this paper suggest that such a system may very well be a superset of some forms of circumscription. It is also possible that this system, like circumscription, will be defined by a syntactic transformation with a clear model-theoretic meaning—and this is what makes the definition of circumscription so attractive in the first place.

Since some logic programming languages with negation as failure can be easily embedded into default logic ([Bidoit and Froidevaux, 1988], [Gelfond and Lifschitz, 1990], [Kowalski and Sadri, 1990]), we can expect that they, too, will become subsets of the nonmonotonic system of the future. These subsets will be important in view of their good computational properties. It may be possible to automate reasoning in more complex nonmonotonic theories by compiling them into logic programs [Gelfond and Lifschitz, 1989] or by constructing tractable "approximations" to them.

Acknowledgments

I would like to thank David Etherington, Michael Gelfond, Ramiro Guerreiro, Hector Levesque, and John McCarthy for comments on earlier versions of this paper. This research was supported in part by NSF grant IRI-8904611 and by DARPA under Contract N00039-84-C-0211.

References

[Bidoit and Froidevaux, 1988] Nicole Bidoit and Christine Froidevaux. Negation by default and nonstratifiable logic programs. Technical Report 437, Université Paris XI, 1988.

[Etherington, 1987a] David Etherington. Relating default logic and circumscription. In *Proc. IJCAI-87*, pages 489–494, 1987.

[Etherington, 1987b] David Etherington. A semantics for default logic. In *Proc. IJCAI-87*, pages 495–498, 1987.

[Gelfond and Lifschitz, 1989] Michael Gelfond and Vladimir Lifschitz. Compiling circumscriptive theories into logic programs. In Michael Reinfrank, Johan de Kleer, Matthew Ginsberg, and Erik Sandewall, editors, *Non-Monotonic Reasoning: 2nd International Workshop (Lecture Notes in Artificial Intelligence 346)*, pages 74–99. Springer-Verlag, 1989.

[Gelfond and Lifschitz, 1990] Michael Gelfond and Vladimir Lifschitz. Logic programs with classical negation. In David Warren and Peter Szeredi, editors, *Logic Programming: Proc. of the Seventh Int'l Conf.*, pages 579–597, 1990.

[Guerreiro and Casanova, 1990] Ramiro Guerreiro and Marco Casanova. An alternative semantics for default logic. Preprint, The Third International Workshop on Nonmonotonic Reasoning, South Lake Tahoe, 1990.

[Konolige, 1988] Kurt Konolige. On the relation between default and autoepistemic logic. *Artificial Intelligence*, 35:343–382, 1988.

[Konolige, 1989] Kurt Konolige. On the relation between autoepistemic logic and circumscription. In *Proc. of IJCAI-89*, pages 1213–1218, 1989.

[Kowalski and Sadri, 1990] Robert Kowalski and Fariba Sadri. Logic programs with exceptions. In David Warren and Peter Szeredi, editors, *Logic Programming: Proc. of the Seventh Int'l Conf.*, pages 598–613, 1990.

[Levesque, 1990] Hector Levesque. All I know: a study in autoepistemic logic. *Artificial Intelligence*, 42(2,3):263–310, 1990.

[Lifschitz, 1985] Vladimir Lifschitz. Computing circumscription. In *Proc. of IJCAI-85*, pages 121–127, 1985.

[Lifschitz, 1989a] Vladimir Lifschitz. Benchmark problems for formal non-monotonic reasoning, version 2.00. In Michael Reinfrank, Johan de Kleer, Matthew Ginsberg, and Erik Sandewall, editors, *Non-Monotonic Reasoning: 2nd International Workshop (Lecture Notes in Artificial Intelligence 346)*, pages 202–219. Springer-Verlag, 1989.

[Lifschitz, 1989b] Vladimir Lifschitz. Between circumscription and autoepistemic logic. In Ronald Brachman, Hector Levesque, and Raymond Reiter, editors, *Proc. of the First Int'l Conf. on Principles of Knowledge Representation and Reasoning*, pages 235–244, 1989.

[Marek and Truszczyński, 1989] Wiktor Marek and Mirosław Truszczyński. Relating autoepistemic and default logic. In Ronald Brachman, Hector Levesque, and Raymond Reiter, editors, *Proc. of the First Int'l Conf. on Principles of Knowledge Representation and Reasoning*, pages 276–288, 1989.

[Marek and Truszczyński, 1990] Wiktor Marek and Mirosław Truszczyński. Modal logic for default reasoning. To appear, 1990.

[McCarthy, 1980] John McCarthy. Circumscription—a form of non-monotonic reasoning. *Artificial Intelligence*, 13:27–39,171–172, 1980. Reproduced in [McCarthy, 1990].

[McCarthy, 1986] John McCarthy. Applications of circumscription to formalizing common sense knowledge. *Artificial Intelligence*, 26(3):89–116, 1986. Reproduced in [McCarthy, 1990].

[McCarthy, 1990] John McCarthy. *Formalizing common sense: papers by John McCarthy.* Ablex, Norwood, NJ, 1990.

[McDermott and Doyle, 1980] Drew McDermott and Jon Doyle. Nonmonotonic logic I. *Artificial Intelligence*, 13:41–72, 1980.

[McDermott, 1982] Drew McDermott. Nonmonotonic logic II: Nonmonotonic modal theories. *Journal of the ACM*, 29(1):33–57, 1982.

[Poole, 1987a] David Poole. Fixed predicates in default reasoning. Manuscript, 1987.

[Poole, 1987b] David Poole. Variables in hypotheses. In *Proc. IJCAI-87*, pages 905–908, 1987.

[Przymusinski, 1989] Teodor Przymusinski. On the declarative and procedural semantics of logic programs. *Journal of Automated Reasoning*, 5:167–205, 1989.

[Reiter, 1980] Raymond Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[Reiter, 1988] Raymond Reiter. On integrity constraints. In Moshe Vardi, editor, *Theoretical Aspects of Reasoning about Knowledge: Proc. of the Second Conf.*, pages 97–111, 1988.

[Shvarts, 1990] Grigori Shvarts. Autoepistemic modal logics. In Rohit Parikh, editor, *Theoretical Aspects of Reasoning about Knowledge: Proc. of the Third Conf.*, pages 97–110, 1990.

Appendix. Proofs of Theorems

Proposition 1. *The union of all classes V' of models of W which satisfy $(**)$ satisfies $(**)$ also.*

Proof. Let V_0' be the union of all classes V' of models which satisfy $(**)$, and let (1) be a default from D such that $\alpha \in Th(V_0')$ and $\neg\beta_1, \ldots, \neg\beta_m \notin Th(V)$. Take any model $M \in V_0'$. By the choice of V_0', there exists a class V' of models which satisfies the conditions $(**)$ and $M \in V' \subset V_0'$. Since $\alpha \in Th(V_0') \subset Th(V')$, we can conclude that $\gamma \in Th(V')$. Consequently γ is true in M. Thus γ is true in every model from V_0', that is, $\gamma \in Th(V_0')$.

The proof of Proposition 2 is based on two lemmas.

Lemma 1. *For any class of structures V, $\Gamma(Th(V)) = Th(\Sigma(V))$.*

Proof. According to the definition of Γ, $\Gamma(Th(V))$ is the smallest class of sentences S' such that (i) S' is closed under classical logic, (ii) $W \subset S'$, and (iii) for any default (1) from D, if $\alpha \in S'$ and $\neg\beta_1, \ldots, \neg\beta_m \notin Th(V)$ then $\gamma \in S'$. Notice that a set of sentences S' is closed under classical logic if and only if it can be represented in the form $Th(V')$ for some class of structures V'. (Proof: Take V' to be the class of all models of S'.) Furthermore, $W \subset Th(V')$ means that every member of V' is a model of W. Consequently, $\Gamma(Th(V))$ can be characterized as the smallest class of sentences of the form $Th(V')$, where V' is a class of models of W, such that $(**)$ holds. Since the operator Th is monotone decreasing, this is the same as $Th(V')$ for the largest class V' of models of W satisfying $(**)$, that is, the same as $Th(\Sigma(V))$.

Lemma 2. *The class of models of $Th(\Sigma(V))$ coincides with $\Sigma(V)$.*

Proof. If V_1' is any class of models of W which satisfies $(**)$, and V_2' is any class of structures such that $Th(V_2') = Th(V_1')$, then V_2' is a class of models of W, and it satisfies $(**)$ also. By applying this to $\Sigma(V)$ as V_1' and to the class of models of $Th(\Sigma(V))$ as V_2', we conclude that the class of models of $Th(\Sigma(V))$ is a class of models of W and satisfies $(**)$. Since it contains $\Sigma(V)$, which is the largest such class, the two classes coincide.

Proposition 2. *A set of sentences is an extension for (D, W) if and only if it has the form $Th(V)$ for some fixpoint V of Σ.*

Proof. If V is a fixpoint of Σ, then, by Lemma 1,

$$\Gamma(Th(V)) = Th(\Sigma(V)) = Th(V),$$

so that $Th(V)$ is an extension. To show that any extension can be represented in this form, consider any extension S, and let V be the class of its models. Since S is closed under first-order logic, $S = Th(V)$, and it remains to check that V is a fixpoint of Σ. Since S is a fixpoint of Γ, $Th(V) = \Gamma(Th(V))$. From the last two equalities and Lemma 1, $S = Th(\Sigma(V))$. By Lemma 2, it follows that $\Sigma(V)$ is class of models of S, that is, $\Sigma(V) = V$.

Proposition 3. *The set of sets V' satisfying $(***)$ is closed under union.*

Proof. Let V_0' be the union of some of the sets V' which satisfy $(***)$, let (4) be a default from D, and let ξ be a tuple of names such that $\alpha(\xi) \in Th^*(V_0')$ and $\neg\beta_1(\xi), \ldots, \neg\beta_m(\xi) \notin$

$Th^*(V)$. Take any world M from V_0'. By the choice of V_0', there exists a set V' of worlds which satisfies the conditions $(***)$ and $M \in V' \subset V_0'$. Since $\alpha(\xi) \in Th^*(V_0') \subset Th^*(V')$, we can conclude that $\gamma(\xi) \in Th^*(V')$. Consequently $\gamma(\xi)$ is true in M. Thus $\gamma(\xi)$ is true in every world from V_0', that is, $\gamma(\xi) \in Th^*(V_0')$.

The proofs of Propositions 4 and 5 use the following terminology. Given a set of sentences W and a nonempty set U, we say that U is *W-complete* if for every model M of W there exists a structure with the universe U that is elementarily equivalent to M. For instance, if all models of W have the same cardinality, then any set of this cardinality is W-complete. The Löwenheim—Skolem theorem shows that if W has at most countably many constants and no finite models, then any infinite set is W-complete.

For any structure M, let $\overline{M}$ be the class of structures elementarily equivalent to M. If V is a class of structures, then $\overline{V}$ stands for $\bigcup_{M \in V} \overline{M}$. Obviously, $V \subset \overline{V}$ and $\overline{\overline{V}} = \overline{V}$.

By V_0 we will denote the set of all worlds (that is, of all models of W with the universe U).

Lemma 3. *If U is W-complete, then, for any class V of models of W,*

$$\overline{\overline{V} \cap V_0} = \overline{V}.$$

Proof. The inclusion left to right is obvious. Take any model $M \in \overline{V}$, and let M' be a world elementarily equivalent to M. Then

$$M' \in \overline{M} \cap V_0 \subset \overline{V} \cap V_0.$$

Consequently,

$$M \in \overline{M'} \subset \overline{\overline{V} \cap V_0}.$$

For any class V of models of W,

$$\overline{\Sigma(V)} = \Sigma(V), \tag{11}$$

because $\Sigma(V)$ is the largest class satisfying $(**)$, and $\overline{\Sigma(V)}$ is its superclass with the same theory. Observe also that, if U is W-complete, then

$$Th(V) = Th(\overline{V}) = Th(\overline{V} \cap V_0). \tag{12}$$

Since $\Sigma(V)$ is invariant with respect to replacing V by another class of structures with the same theory, it follows that

$$\Sigma(V) = \Sigma(\overline{V}) = \Sigma(\overline{V} \cap V_0). \tag{13}$$

Lemma 4. *For any closed default theory (D, W), any W-complete U, and any set of worlds V,*

$$\Delta(V) = \Sigma(V) \cap V_0.$$

Proof. It is clear that $\Delta(V)$ is contained both in $\Sigma(V)$ and in V_0. In order to prove that $\Sigma(V) \cap V_0$ is a subset of $\Delta(V)$, we only need to check that it satisfies $(**)$ as V'. This follows from the fact that, by (11) and (12), $\Sigma(V) \cap V_0$ has the same theory as $\Sigma(V)$.

Lemma 5. *For any closed default theory (D, W), any W-complete U, and any set of worlds V,*

$$\Sigma(V) = \overline{\Delta(V)}.$$

Proof. Obviously, $\Delta(V) \subset \Sigma(V)$. Then, by (11),

$$\overline{\Delta(V)} \subset \overline{\Sigma(V)} = \Sigma(V).$$

To prove that $\Sigma(V)$ is a subclass of $\overline{\Delta(V)}$, we need to check that, for any V' satisfying $(**)$, $V' \subset \overline{\Delta(V)}$. Take any V' satisfying $(**)$. By (12), $Th(\overline{V'} \cap V_0) = Th(V')$. Consequently, $\overline{V'} \cap V_0$ satisfies $(**)$ also, so that $\overline{V'} \cap V_0 \subset \Delta(V)$. Then, by Lemma 3, $V' \subset \overline{\overline{V'} \cap V_0} \subset \overline{\Delta(V)}$.

Lemma 6. *Let (D, W) be a closed default theory. If U is W-complete, then the U-extensions of (D, W) are identical to its extensions.*

Proof. Consider any extension, that is, a set of the form $Th(V)$, where $\Sigma(V) = V$ (Proposition 2). By (11), $\overline{V} = V$; consequently, (12) and (13) give

$$Th(V) = Th(V \cap V_0) \tag{14}$$

and

$$\Sigma(V \cap V_0) = V.$$

From the last formula and Lemma 4, $\Delta(V \cap V_0) = V \cap V_0$, so that $V \cap V_0$ is a fixpoint of Δ, and $Th(V \cap V_0)$ is a U-extension. By (14), this set is identical to $Th(V)$.

Now consider any U-extension, that is, a set of the form $Th(V)$, where $\Delta(V) = V$. By (13) and Lemma 5,

$$\Sigma(\overline{V}) = \Sigma(V) = \overline{\Delta(V)} = \overline{V}.$$

Thus $\overline{V}$ is a fixpoint of Σ, so that $Th(\overline{V})$ is an extension (Proposition 2). By (12), this set is identical to $Th(V)$.

Proposition 4. *Let (D, W) be a closed default theory. If all models of W have the same finite cardinality, then the F-consequences of (D, W) are identical to its consequences.*

Proof. Let the common cardinality of all models of W be n. If the cardinality of U is n, then, by Lemma 6, the U-extensions of the theory are the same as its extensions. If not, then the set of worlds is empty, $\emptyset$ is the only fixpoint of Δ, and the set of all sentences is the only U-extension. It follows that the intersection of all U-extensions coincides with the intersection of all extensions.

Proposition 5. *Let (D, W) be a closed default theory with at most countably many object, function and predicate constants. If all models of W are infinite, then the F-consequences of (D, W) are identical to its consequences.*

Proof. If U is infinite, then, by Lemma 6, the U-extensions of the theory are the same as its extensions. If U is finite, then the set of worlds is empty, $\emptyset$ is the only fixpoint of Δ, and the set of all sentences is the only U-extension. It follows that the intersection of all U-extensions coincides with the intersection of all extensions.

Proposition 6. *For any default theory with the set of defaults (7), every fixpoint of the corresponding operator* Δ *is an equivalence class of the set of* α*-maximal worlds relative to the relation* $\alpha^M = \alpha^{M'}$*. Conversely, each of these equivalence classes is a fixpoint of* Δ*.*

Proof. By the definition of α^M, for any set of worlds V and any tuple of names ξ,

$$\alpha(\xi) \in Th^*(V) \Leftrightarrow \xi \in \bigcap_{M \in V} \alpha^M,$$

$$\neg\alpha(\xi) \in Th^*(V) \Leftrightarrow \xi \notin \bigcup_{M \in V} \alpha^M.$$

Consequently, for a default theory with the set of defaults (7), $(***)$ is equivalent to the condition:

For any tuple of names ξ, if $\xi \in \bigcup_{M \in V} \alpha^M$ then $\xi \in \bigcap_{M' \in V'} \alpha^{M'}$,

that is to say, to the inclusion

$$\bigcup_{M \in V} \alpha^M \subset \bigcap_{M' \in V'} \alpha^{M'}.$$

Since $\Delta(V)$ is the largest set of worlds V' satisfying this condition,

$$\Delta(V) = \{M' : \bigcup_{M \in V} \alpha^M \subset \alpha^{M'}\}. \tag{15}$$

Assume that V is an equivalence class of the set of α-maximal worlds relative to the relation $\alpha^M = \alpha^{M'}$. Then, for some world M_0,

$$V = \{M : \alpha^{M_0} = \alpha^M\} = \{M : \alpha^{M_0} \subset \alpha^M\}.$$

The set union in (15) coincides with α^{M_0}, so that

$$\Delta(V) = \left\{M' : \alpha^{M_0} \subset \alpha^{M'}\right\} = V.$$

Take now any fixpoint V of Δ. If M' is a world from V, then, by (15),

$$\bigcup_{M \in V} \alpha^M \subset \alpha^{M'},$$

so that, for any $M \in V$, $\alpha^M \subset \alpha^{M'}$. We established this inclusion for any pair of worlds $M, M' \in V$, which means that, for every such pair, $\alpha^M = \alpha^{M'}$. Furthermore, V is nonempty, because, by (15), $\Delta(\emptyset)$ is the set of all worlds (which is assumed to be nonempty). For any world $M_0 \in V$,

$$V \subset \{M' : \alpha^{M_0} = \alpha^{M'}\},$$

and, by (15),

$$\Delta(V) = \{M' : \alpha^{M_0} \subset \alpha^{M'}\}.$$

Since V is a fixpoint of Δ, we conclude that

$$\{M' : \alpha^{M_0} \subset \alpha^{M'}\} \subset \{M' : \alpha^{M_0} = \alpha^{M'}\}.$$

This means that M_0 is α-maximal. Consequently, the fixpoint V is a subset of an equivalence class V_0 of α-maximal worlds. We know that V_0 is a fixpoint of Δ also; since Δ is monotone decreasing, it follows that $V = V_0$.