Value Minimization in Circumscription

Chitta Baral **Alfredo Gabaldon**
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968 U.S.A.
{*chitta, alfredo*}*@cs.utep.edu*

Alessandro Provetti
C.I.R.F.I.D.
Università di Bologna
Via Galliera 3, I-40121 Bologna, Italy
provetti@cirfid.unibo.it

Abstract

Minimization in circumscription has focussed on minimizing the extent of a set of predicates (with or without priorities among them), or of a formula. Although most circumscription formalisms allow varying of functions and other constants, no formalism to the best of our knowledge minimized functions. In this paper we introduce and motivate the notion of *value minimizing* a function in circumscription. In value minimizing we do not minimize the extent of the function; rather we minimize the value of the function. We show how Lifschitz's nested abnormality theories can be used to do value minimization.

1 Introduction

In this paper we deal with the circumscription of terms in default theories for knowledge representation. While circumscription of predicates minimizes their extent, we describe circumscription of terms as minimizing the interpretation of terms w.r.t. a given *weight* of objects in the domain.

To the best of our knowledge minimization in circumscription has focussed on minimizing the extent of a set of predicates (with or without priorities among them) [Grosof, 1991, Lifschitz, 1993], or of a formula [McCarthy, 1986]. Also, although minimization of functions is not ruled out, we did not find examples that dealt with minimizing functions. Most circumscription formalisms do allow varying of functions and other constants while circumscribing though.

To visualize our goal, consider a database of employees containing the relation

$$Employee(emp\_id, emp\_name, salary)$$

The logical counterpart of this database is a first-order axiomatization with Closed World Assumption, e.g. with circumscription of the predicate *Employee*. That is, the *extent* of the predicate *Employee* is minimal compatible with the constraints.

Now suppose we would like to add additional information that the employer tries his best to minimize the salary of each of his employees. To formalize this we need axioms that will allow us to prefer models of the theory where employees get paid lesser. For example, let us assume that the employees in our database are *John* and *Mike* with *100* and *200* as their corresponding *emp_id*. Now suppose our theory has two models:

$$M_1 = \{ Employee(100, John, 30000), \\ Employee(200, Mike, 40000)\}$$

$$M_2 = \{ Employee(100, John, 35000), \\ Employee(200, Mike, 43000)\}$$

The employer will of course prefer $\mathcal{M}_1$ over $\mathcal{M}_2$. The goal of this paper is to formalize the minimization used in the above example where $\mathcal{M}_1$ is minimal among the set of models $\{\mathcal{M}_1, \mathcal{M}_2\}$.

The formalization of the above example includes the constraint that an employee has a unique salary and a unique name. Hence, we can intuitively view the above minimization as minimizing a function F_{sal} that maps employee ids into salaries. Notice, however, that we are not minimizing the extent of the function[1], rather we are minimizing the value of the function.

Recently in [Baral et al., 1996a], we have discussed the relationship between specifications in the action description language $\mathcal{L}$ [Baral et al., 1996b] and circumscriptive theories. There we were faced with the minimization of a particular term, $sit\_map(S_N)$, that mapped the current-situation symbol S_N onto a sequence of actions, understood as the history of the

[1]To stay close to the usual circumscription terminology, we use "extent of a function" to refer to the domain of a function.

domain. This was required to formalize the assumption

"no actions occurred except those needed to explain the facts in the theory"

which is present in the semantics of $\mathcal{L}$.

The main focus of [Baral et al., 1996a] was in showing the equivalence of specifications in $\mathcal{L}$ and their axiomatization based on Nested Abnormality Theories, a novel circumscription schema proposed by Lifschitz [Lifschitz, 1995]. Minimizing the term $Sit\_map(S_N)$ was part of this axiomatization. We believe function minimization and the minimization of ground terms (term minimization) are interesting in their own right and need to be discussed explicitly and independently. This will allow other researchers to easily use them in their applications using circumscription.

With that in mind, in this paper we generalize and expand on our observations in [Baral et al., 1996a] and formalize the notion of minimizing functions[2] where the minimization is done not with respect to the extent of the function, but with respect to the value of the function. The minimization is relative to an ordering $\mathcal{R}$ over the elements of the universe. To distinguish it from the standard minimization in circumscription, we refer to this as *value minimization*.

Our discussion is carried out in the framework of Nested Abnormality Theories, and we believe it represents further evidence of the suitability of this circumscription schema for knowledge representation.

To start with, we use the translation of specifications in $\mathcal{L}$ to circumscriptive theories as a motivational example of the necessity of value minimization.

2 Motivating example: narratives

The motivations and the theoretical framework of our approach to actions are discussed in detail in [Baral et al., 1996b], which is a continuation of the proposal in [Gelfond and Lifschitz, 1992] for *provably correct theories of actions*. Let us just point out that we have a language $\mathcal{L}$ for describing a sequence of actions and observations, regarding the evolution of a domain. Unlike the basic situation calculus, $\mathcal{L}$ allows observations about actual happenings and values of fluents in actual situations (i.e., allows narratives). Unlike the basic event calculus [Kowalski and Sergot, 1986], $\mathcal{L}$ allows reasoning about actual and hypothetical situations, possibly combined, e.g. to represent "adaptive plans." The entailment of $\mathcal{L}$ theories is nonmonotonic and thus apt for reasoning with incomplete information. Moreover, new observations are assimilated by simple insertion of new facts into the theory.

[2]Our notion of minimization is also applicable to terms.

Recently, [Baral et al., 1996a] we have defined a translation from narrative description specifications in $\mathcal{L}$ into Nested Abnormality Theories [Lifschitz, 1995][3]. In the rest of this section we will focus on this translation, which we describe informally as made up of:

- a subtheory describing effects of actions performed in a domain which is described by means of boolean fluents;
- the axiomatization of *common-sense inertia*;
- observations, i.e. facts describing fluent-values at the initial situation as well as at intermediate stages of the evolution of the domain; and
- facts describing which actions have been previously performed, and their ordering in time.

These axioms are completed with an assumption[4] specifying that

> models of the theory are those that satisfy all the observations and —other things being equal— entail that a minimal number of actions have occurred.

The above assumption, which is truly a model preference criterion, has been captured in our NAT formalization in the following way.

Let us introduce some notation. Let sequences of actions be represented by terms like

$$A_n \ldots \circ A_2 \circ A_1 \circ \epsilon$$

read as "A_1 then A_2 then ... A_n" (constant ϵ represents the empty sequence). Next, we have a set of constants $S_0, S_1 \ldots$ for denoting states of affairs the domain has passed through. The function *Sit_map* maps each situation constant into a sequence of actions which, intuitively, describes what has happened up to the situation itself. Of course the situations are totally ordered along a time line, although the ordering itself can vary from model to model. However, the special constant symbol S_N is always associated to the *late-most* situation, so that $Sit\_map(S_N)$ is interpreted on the longest sequence, i.e. the complete *history* of the domain. For instance:

$\mathcal{M} \models Sit\_map(S_0) = \epsilon$
$\mathcal{M} \models Sit\_map(S_1) = A_1 \circ \epsilon$
...
$\mathcal{M} \models Sit\_map(S_N) = A_k \circ A_{k-1} \circ \ldots \epsilon$

[3]See Appendix A for a brief introduction to Nested Abnormality Theories.

[4]This assumption is not present elsewhere, in approaches that either entail no extra actions, e.g. [Shanahan, 1995] and [Kakas and Miller, 1996], or make no assumptions at all about action occurrences, e.g. [Lin and Reiter, 1995].

Now it should be clear how minimizing $Sit\_map(S_N)$, i.e. mapping it on the shortest possible sequence of actions, captures the assumption discussed above. Here is a sketch of how we formalized this notion with NATs. First of all, assume the following two NATs, which define predicates $Prefix\_eq$ and $Subsequence$ with their usual meaning (universal quantification is implicit on all variables):

$$
\begin{aligned}
&B_{Subsequence} = \\
&\{min\ Subsequence\ : \\
&\qquad Subsequence(\epsilon, \epsilon) \\
&\qquad Subsequence(\alpha, \alpha_1) \supset Subsequence(\alpha, a \circ \alpha_1) \\
&\qquad Subsequence(\alpha, \alpha_1) \supset Subsequence(a \circ \alpha, a \circ \alpha_1) \\
&\}
\end{aligned}
$$

$$
\begin{aligned}
&B_{prefix\_eq} = \\
&\{min\ Prefix\_eq\ : \\
&\qquad Sequence(\alpha) \supset Prefix\_eq(\alpha, \alpha) \\
&\qquad Prefix\_eq(\alpha, \alpha_1) \supset Prefix\_eq(\alpha, a \circ \alpha_1) \\
&\}
\end{aligned}
$$

The above two NATs are *nested* within the outermost theory $T_{\mathcal{L}}$, defined as follows.

$$
\begin{aligned}
T_{\mathcal{L}} = \quad &\{\ Sit\_map\ : \\
&\qquad Sit\_map(S_0) = \epsilon \\
&\qquad Prefix\_eq(Sit\_map(s), Sit\_map(S_N)) \\
&(\star)\ \ Subsequence(\alpha, Sit\_map(S_N)) \supset Ab(\alpha) \\
&\qquad B_{prefix\_eq} \\
&\qquad B_{subsequence} \\
&\qquad \text{other axioms} \\
&\}
\end{aligned}
$$

The first two axioms make sure that S_0 is always mapped on the empty sequence and that all situations are mapped on sequences which are prefixes of the complete history $Sit\_map(S_N)$. Axiom $(\star)$ postulates that each model of $T_{\mathcal{L}}$ contains a set of instances $Ab(\alpha)$, where αs are all the possible subsequences of $Sit\_map(S_N)$. By including axiom $(\star)$ we exploit the fact that subsequence is a partial order on sequences, i.e. if α is a subsequence of β then all subsequences of the former are also subsequences of the latter.

Now, suppose interpretation $\mathcal{I}$ satisfies all the axioms of $T_{\mathcal{L}}$ and maps $Sit\_map(S_N)$ on α while interpretation $\mathcal{I}'$ maps it —other things being equal— onto β. As a result, the extent of Ab under $\mathcal{I}$ is a proper subset of the extent of Ab under $\mathcal{I}'$. Therefore, $\mathcal{I}'$ is not a circumscriptive model of $T_{\mathcal{L}}$. This kind of considerations are the subject of next section.

3 Generalizing the Approach

In this section the theory of value minimization by NATs is introduced and discussed in model-theoretic terms. First of all, the concept of value minimization is defined precisely. Second, we describe the structure of NATs that *implements* value minimization and show its correctness.

Our discussion is carried out in a theoretical framework where several assumptions are present. We argue that these assumptions are common to most applications of knowledge representation and reasoning we are familiar with, and therefore they do not seem to be limiting the scope of application.

Domain closure (DCA). Each object of the material domain is represented by a ground term, i.e., we restrict to models whose universe is isomorphic to the Herbrand domain.

Unique names (UNA). The usual sets of constant inequalities $C_1 \neq C_2,\ C_1 \neq C_3, \ldots$ are always included in the theories we examine.

Notice how both DCA and UNA are easily added or removed from theories. See [Lifschitz, 1995] for a discussion on which axioms/blocks are needed to implement DCA.

Definition 1 *(Explicit domain)*

A NAT is said to have an explicit domain if it contains axioms and blocks for DCA and UNA.

Now we can define value minimization as a partial order on models of explicit domain theories. To do so, let us introduce the following notation. Let

$$\mathcal{I}[[\pi]]$$

stand for the set of tuples which belong to the extent of predicate π in interpretation $\mathcal{I}$. For functions we use $\mathcal{I}[[\phi]]$ to denote the codomain of ϕ in $\mathcal{I}$. Also, we use

$$\mathcal{I}[[\phi]](\tau)$$

to denote the object which —according to interpretation $\mathcal{I}$— function ϕ maps term τ into. Since we are considering Herbrand models only, we can have expressions like $\mathcal{I}[[\phi]](\tau) = \nu$ where ν is a term of the language of the object theory.

3.1 Value Minimization

Definition 2 *(Value-minimal)*

Let T be a theory, ϕ be a function and Z a tuple of predicate/function constants in the language of T. Let $\mathcal{R}$ be a partial order defined between the elements of the universe. For two structures $\mathcal{M}$ and $\mathcal{M}'$ of T, we say $\mathcal{M} \leq^{(\phi,\mathcal{R});Z} \mathcal{M}'$ if

1. $|\mathcal{M}\ | = |\mathcal{M}'|$;

2. $\mathcal{M}[[\sigma]] = \mathcal{M}'[[\sigma]]$
For each constant σ *s.t.* $\sigma \neq \phi$, $\sigma \notin Z$;

3. $\forall x.\mathcal{M}[[\phi]](x)\ \mathcal{R}\ \mathcal{M}'[[\phi]](x)$.

A model $\mathcal{M}$ *of* T *is minimal relative to* $\leq^{(\phi,\mathcal{R});Z}$ *if there is no model* $\mathcal{M}'$ *of* T *such that* $\mathcal{M}' <^{(\phi,\mathcal{R});Z} \mathcal{M}$.

The definition above can be tailored to minimization of ground terms, as we did in Section 2 when we minimized the ground term $Sit\_map(S_N)$ rather than function constant Sit_map as such. In these cases we speak of *term minimization.* In general, when the minimization is relative to a function ϕ and terms $t_1, \ldots, t_n$ the third condition in Definition 2 becomes

$$\mathcal{M}[[\phi]](t_i)\ \mathcal{R}\ \mathcal{M}'[[\phi]](t_i) \text{ for each } t_i,$$

and we write minimality as relative to the ordering $\leq^{(\phi,t_1,\ldots,t_n,\mathcal{R});Z}$ between structures.

Example 1 *(Exponentiation) Consider the following two interpretations of function power on the domain of naturals with* $\leq$ *as the usual "less than or equal" relation on naturals. By abuse of language, we define:*

$$\mathcal{M}[[power]](x) = x^2,\ \mathcal{M}'[[power]](x) = x^3.$$

Of course, for all x $\mathcal{M}[[power]](x) \leq \mathcal{M}'[[power]](x)$ *and* $\mathcal{M} <^{(power,\leq)} \mathcal{M}'$

Definition 2 above concerns minimization relative to a fixed, external criterion $\mathcal{R}$, which is a partial ordering between elements of the universe.

An interesting feature of NATs is that $\mathcal{R}$ can actually be *implemented* within the theory, simply by adding a block containing its axiomatization. Other circumscription methods (such as prioritized circumscription) would bring about the additional complication of possible undesired interaction between the part of the theory where $\mathcal{R}$ is supposed to be defined and the rest of the theory where $\mathcal{R}$ may appear[5].

Definition 3 *(Term ordering)*

We say that a NAT T has a term ordering $\mathcal{R}$ *with respect to a function* ϕ *if it contains a block defining a partial ordering* $\mathcal{R}$ *on the elements of the codomain of* ϕ*, i.e. For all models* $\mathcal{M}$ *of* T*, and for all* x *and* y*,*

$$\mathcal{M}[[\phi]](x)\ \mathcal{R}\ \mathcal{M}[[\phi]](y) \text{ iff } (\mathcal{M}[[\phi]](x), \mathcal{M}[[\phi]](y)) \in \mathcal{M}[[\mathcal{R}]]$$

□

[5]Notice that there is no extra complexity involved in this process, since transitive closures are always defined as a second-order theory, as indeed the block defining $\mathcal{R}$ is.

Now, let us show that for each value minimization criterion $\leq^{(\phi,\mathcal{R});Z}$ an equivalent NAT formulation is found. This result is the counterpart of fundamental Proposition 2.5.1 of [Lifschitz, 1993].

Theorem 1 *(Value-minimal equivalence)*

Let T *be an explicit domain NAT with term ordering* $\mathcal{R}$ *and let* ϕ *be a function constant and* Z *be a tuple of predicate/function constants in the language of* T*. Then, a model* $\mathcal{M}$ *of* T *is a model of*

$$\begin{array}{l} \{\phi, Z\ : \\ \qquad \forall x, y.\ \mathcal{R}(y, \phi(x)) \supset Ab(x,y) \\ \\ \qquad T \\ \} \end{array} \tag{1}$$

if and only if $\mathcal{M}$ *is minimal relative to* $\leq^{(\phi,\mathcal{R});Z}$.

Proof:

Consider a theory T as described in the above proposition. Let $T_{\phi,\mathcal{R}}$ be the NAT of the form (1) built from T.

($\Leftarrow$)

Assume $\mathcal{M}$ to be a model of T minimal relative to $\leq^{(\phi,\mathcal{R});Z}$. Let us define the structure $\mathcal{M}_{Ab}$ from $\mathcal{M}$ by letting:

1. $|\mathcal{M}_{Ab}| = |\mathcal{M}|$;
2. $\mathcal{M}_{Ab}[[\sigma]] = \mathcal{M}[[\sigma]]$ for every constant σ in the language of T[6] which is not included in Z;
3. $\mathcal{M}_{Ab}[[Ab]] = \{(x,y) : (y, \phi(x)) \in \mathcal{M}[[\mathcal{R}]]\}$.

Since $\mathcal{M}_{Ab}$ is a model of T, $\mathcal{M}_{Ab}$ is a model of $T_{\phi,\mathcal{R}}$ iff the extent of Ab is minimal. We now show $\mathcal{M}_{Ab}$ is a model of $T_{\phi,\mathcal{R}}$ by contradiction.

Suppose there is a model $\mathcal{M}'$ of $T_{\phi,\mathcal{R}}$ such that $|\mathcal{M}'| = |\mathcal{M}_{Ab}|$, $\mathcal{M}'[[\sigma]] = \mathcal{M}_{Ab}[[\sigma]]$ for every constant σ different from Ab and not in Z, and $\mathcal{M}'[[Ab]] \subset \mathcal{M}_{Ab}[[Ab]]$. Then, there exist x, y s.t.

$$(x,y) \notin \mathcal{M}'[[Ab]] \text{ and } (x,y) \in \mathcal{M}_{Ab}[[Ab]].$$

For the same x and y, since

$$\mathcal{M}[[\phi]](x)\ \mathcal{R}\ \mathcal{M}'[[\phi]](x) \Rightarrow y\ \mathcal{R}\ \mathcal{M}'[[\phi]](x)$$

by transitivity, we conclude that $\mathcal{M} \not\leq^{(\phi,\mathcal{R});Z} \mathcal{M}'$. On the other hand, since for all x

$$\mathcal{M}'[[\phi]](x)\ \mathcal{R}\ \mathcal{M}'[[\phi]](x) \Rightarrow (x, \mathcal{M}'[[\phi]](x)) \in \mathcal{M}'[[Ab]]$$

[6]Notice Ab does not belong to the language of T.

we can conclude that $(x, \mathcal{M}'[[\phi]](x)) \in \mathcal{M}_{Ab}[[Ab]]$. Now, since it follows that $\forall x. \mathcal{M}'[[\phi]](x)\ \mathcal{R}\ \mathcal{M}[[\phi]](x)$ we have established that $\mathcal{M}' \leq^{(\phi,\mathcal{R});Z} \mathcal{M}$. This implies $\mathcal{M}$ is not minimal relative to $\leq^{(\phi,\mathcal{R});Z}$ which contradicts our assumption, therefore $\mathcal{M}_{Ab}$ is a model of $T_{\phi,\mathcal{R}}$.

$(\Rightarrow)$

Assume $\mathcal{M}$ to be a model of $T_{\phi,\mathcal{R}}$. By definition of NATs, $\mathcal{M}$ is a model of T. We show $\mathcal{M}$ is a model of T minimal relative to $\leq^{(\phi,\mathcal{R});Z}$ by contradiction.

Suppose there is a model $\mathcal{M}'$ of T s.t. $\mathcal{M}' <^{(\phi,\mathcal{R});Z} \mathcal{M}$. Let us augment $\mathcal{M}'$ by a predicate Ab defined as follows:

$$\mathcal{M}'[[Ab]] = \{(x,y) : (y, \phi(x)) \in \mathcal{M}'[[\mathcal{R}]]\}$$

Then, for all x $\mathcal{M}'[[\phi]](x)\ \mathcal{R}\ \mathcal{M}[[\phi]](x)$. By transitivity, for all x, y:

$$y\ \mathcal{R}\ \mathcal{M}'[[\phi]](x)\ \Rightarrow\ y\ \mathcal{R}\ \mathcal{M}[[\phi]](x)$$

Thus, $(x,y) \in \mathcal{M}'[[Ab]]$ implies $(x,y) \in \mathcal{M}[[Ab]]$ and therefore

$$\mathcal{M}'[[Ab]] \subseteq \mathcal{M}[[Ab]].$$

On the other hand, there exists an x for which the condition $\mathcal{M}[[\phi]](x)\ \mathcal{R}\ \mathcal{M}'[[\phi]](x)$ does not hold. For such an x, $(x, \mathcal{M}[[\phi]](x)) \notin \mathcal{M}'[[Ab]]$. However, for all x, $\mathcal{M}[[\phi]](x)\ \mathcal{R}\ \mathcal{M}[[\phi]](x)$ holds by definition of $\mathcal{R}$. Hence, $(x, \mathcal{M}[[\phi]](x)) \in \mathcal{M}[[Ab]]$ and we obtain

$$\mathcal{M}[[Ab]] \not\subseteq \mathcal{M}'[[Ab]]$$

Consequently, $\mathcal{M}$ is not a model of $T_{\phi,\mathcal{R}}$, which contradicts our assumption. Therefore $\mathcal{M}$ is a model of T minimal relative to $\leq^{(\phi,\mathcal{R});Z}$. □

3.2 Term Minimality

Let us now move on to define minimization of ground terms, i.e. to generalize the approach we discussed in Section 2. As before, consider an explicit domain NAT T with term ordering $\mathcal{R}$, and let ϕ be a function constant, $t_1, \ldots, t_n$ be ground terms, and Z be a tuple of predicate/function constants in the language of T. The following result is easily established.

Proposition 1 *(Term-minimal equivalence)*

a model $\mathcal{M}$ of T is a model[7] *of*

$$\begin{array}{ll} \{\phi, Z\ : & \\ & \forall y.\ \mathcal{R}(y, \phi(t_1)) \supset Ab(t_1, y) \\ & \ldots \\ & \forall y.\ \mathcal{R}(y, \phi(t_n)) \supset Ab(t_n, y) \\ & \\ & T \\ \} & \end{array} \quad (2)$$

if and only if $\mathcal{M}$ is minimal relative to $\leq^{(\phi,t_1,\ldots,t_n,\mathcal{R});Z}$.

□

Our motivating example in Section 2 is a case of term-minimization. The following corollary of the above proposition is useful in proving the equivalence of the semantics of a domain description in $\mathcal{L}$ and its translation to the NAT $T_{\mathcal{L}}$.

Corollary 1 *Let T be the theory consisting of all the axioms in $T_{\mathcal{L}}$ from Section 2 except for axiom $Subsequence(\alpha, Sit\_map(S_N)) \supset Ab(\alpha)$. Then M is a model of $T_{\mathcal{L}}$ iff M is a model of T and is minimal relative to $\leq^{(Sit\_map, S_N, Subsequence)}$.*

□

4 Relationship With Other Circumscription Schemata

As pointed out earlier, one of the main advantages of using NATs is ease of representation, since the problem of interference between circumscription of different predicates is not present. Although NATs constitute a new circumscription framework, it is a predicate which is circumscribed in each block. Thus, intuitively, one may conclude that it should be possible to cast value minimization of a function into a more traditional circumscription form. Proposition 2 below gives a characterization of value minimization in a form similar to formula circumscription [McCarthy, 1986]. The difference with formula circumscription is that a function, rather than predicates, varies during the circumscription of the formula. In this sense, *this formulation constitutes a generalization of formula circumscription.*

Following the approach in [McCarthy, 1986], let $T_{\mathcal{R}}(\Phi, z)$ be a second order theory where function variable Φ[8] and predicate/function variables in z appear as free variables, and which contains a term ordering $\mathcal{R}$. Intuitively, Φ is the function whose valued is to be minimized with respect to $\mathcal{R}$ and z is the list of predicate/function variables which are allowed to vary during circumscription. Now, for a function constant ϕ and constants Z ($\mathcal{R} \notin Z$) in the language of $T_{\mathcal{R}}$ we can prove the following.

[7] For $n = 1$ we just have $Ab(y)$ in the right hand side of the axiom.

[8] This definition can be extended to allow a tuple of functions.

Proposition 2 *A structure $\mathcal{M}$ is a model of*

$$T_{\mathcal{R}}(\phi, Z) \wedge \forall \phi', z.[T_{\mathcal{R}}(\phi', z) \wedge [\forall x, y.\mathcal{R}(y, \phi'(x)) \supset \mathcal{R}(y, \phi(x))] \supset [\forall x, y.\mathcal{R}(y, \phi'(x)) \equiv \mathcal{R}(y, \phi(x))]] \quad (3)$$

if and only if $\mathcal{M}$ is a model of $T_{\mathcal{R}}(\phi, Z)$ and it is minimal relative to $\leq^{(\phi,\mathcal{R});Z}$. □

Formula (3) above can be simplified to the following:

$$T_{\mathcal{R}}(\phi, Z) \wedge \forall \phi', z.[T_{\mathcal{R}}(\phi', z) \wedge [\forall x.\mathcal{R}(\phi'(x), \phi(x))] \supset [\forall x.\mathcal{R}(\phi'(x), \phi(x)) \wedge \mathcal{R}(\phi(x), \phi'(x))]]$$

and clearly this can be further simplified to the formula

$$T_{\mathcal{R}}(\phi, Z) \wedge \forall \phi', z.[T_{\mathcal{R}}(\phi', z) \wedge [\forall x.\mathcal{R}(\phi'(x), \phi(x))] \supset [\forall x.\mathcal{R}(\phi(x), \phi'(x))]]$$

When minimization is relative to ground terms $t_1, \ldots, t_k$, we just need to replace the inner universally quantified formulae with the following conjuncts:

$$T_{\mathcal{R}}(\phi, Z) \wedge \forall \phi', z.\left[T_{\mathcal{R}}(\phi', z) \wedge \left[\bigwedge_{i=1,\ldots,k} \mathcal{R}(\phi'(t_i), \phi(t_i))\right] \supset \left[\bigwedge_{i=1,\ldots,k} \mathcal{R}(\phi(t_i), \phi'(t_i))\right]\right]$$

Furthermore, given that we only consider theories where objects in the world are represented by ground terms, there is an even simpler formula for term minimization when the list Z is empty. Let $\overline{c}$ stand for a tuple of ground terms $c_1, \ldots, c_k$, not containing the constant ϕ, where k is the number of terms relative to which we want to minimize ϕ. Let $T_{\mathcal{R}}(\overline{c})$ stand for $T_{\mathcal{R}} \bigwedge_{i=1,\ldots,k} \phi(t_i) = c_i$. Then we can write term minimization as:

$$T(\overline{c}) \wedge \left[\forall \overline{c}' \left(T(\overline{c}') \bigwedge_{i=1,\ldots,k} \mathcal{R}(c_i', c_i)\right) \supset \left(\bigwedge_{i=1,\ldots,k} \mathcal{R}(c_i, c_i')\right)\right]$$

Consider again our motivating example from Section(2). In this case, the list Z is empty, and we minimize function *Sit_map* only w.r.t. term S_N. Let T be the theory consisting of all the axioms in $T_{\mathcal{L}}$ except for axiom $Subsequence(\alpha, Sit\_map(S_N)) \supset Ab(\alpha)$. Let α be a sequence of actions and let us shorten *Subsequence* into *Sseq*. Let $T(\alpha)$ denote $T \wedge (Sit\_map(S_N) = \alpha))$. We can prove the following:

Proposition 3 *NAT $T_{\mathcal{L}}$ is equivalent to the formula*

$$T(\alpha) \wedge \forall \beta.[(T(\beta) \wedge Sseq(\beta, \alpha)) \supset Sseq(\alpha, \beta)]$$

□

5 Additional Remarks

We now make some observations about our formulation of value minimization of functions (and terms) using circumscription and possible extensions of it.

1. NATs' feature of allowing the definition of $\mathcal{R}$ to be in an independent block which is not affected by other axioms in the theory is very important for our purpose.
2. The blocks implementing value-minimization of a function are just another kind of block, and therefore can be part of another NAT. It was only for the sake of definition that we described them as outermost w.r.t. theories.
3. The concepts of minimizing the value of functions and terms presented so far can be extended to predicates, particularly when they intuitively encode functions. Consider again the database example where we have a predicate *Employee* with attributes *emp_id*, *emp_name*, and *emp_salary*. We want to minimize the salary of each employee. This can be done by the following NAT[9]:

   $$\{Employee : \quad Employee(x, n, z) \wedge y \leq z \wedge Salary(y) \wedge Salary(z) \supset Ab(x, y) \quad T \quad \}$$

   where T consists of the definitions for $\leq$ and *Salary* and the rest of the theory about the employee database.

   This case arises in databases, where it is referred to as a *functional dependency* from a subset of the attributes of the predicate to another subset of the attributes.
4. The ordering $\leq^{(\phi,\mathcal{R});Z}$ between structures in Definition 2 can be extended to value minimization of a set of functions and to incorporate possible priorities between them by following the standard approach.
5. Finally, let us point out how Lifschitz's NAT notation is readily adapted to our definitions. We will write

   $$\{C_1, \ldots, C_m,\ min_{\mathcal{R}}\ \phi : A_1, \ldots, A_n, A_{\mathcal{R}}\} \quad (4)$$

   —where $A_{\mathcal{R}}$ defines[10] $\mathcal{R}$— to denote blocks of the form

   $\{C_1, \ldots, C_m, \phi : \mathcal{R}(y, \phi(x)) \supset Ab(x, y), A_1, \ldots, A_n, A_{\mathcal{R}}\}$

   Intuitively, block (4) refers to a theory consisting of blocks $A_1, \ldots, A_n$ and the block $A_{\mathcal{R}}$ defining

[9] Both the schema below and the ordering between structures in Definition 2 can be easily generalized to arbitrary predicates.

[10] We are assuming that $A_{\mathcal{R}}$ is defined using an NAT block such that statements about $\mathcal{R}$ outside of $A_{\mathcal{R}}$ do not affect the definition of $\mathcal{R}$.

$\mathcal{R}$, and where value minimization of function ϕ is performed while predicate/function constants $C_1, \ldots, C_m$ are varying.

Acknowledgements

This work benefited from several discussions with Michael Gelfond and Vladimir Lifschitz.

Support was provided by the National Science Foundation under grant Nr. IRI-9211662 and IRI-9501577.

References

[Baral et al., 1996a] Baral C., Gabaldon A., and Provetti A., 1996. Formalizing Narratives using Nested Circumscription. In Proc. of *14th AAAI National Conference on Artificial Intelligence*, pages 652–657.

Also in the Working notes of the third symposium *Logical Formalizations of Commonsense Reasoning,* pages 15–24.

[Baral et al., 1996b] Baral C., Gelfond M., and Provetti A., 1996. Representing Actions: Laws, Observations and Hypothesis. To appear on *Journal of Logic Programming,* special issue on actions. Also available via *http://cs.utep.edu/chitta/chitta.html*

An earlier version appeared in the Working Notes of AAAI Spring Symposium *Extending Theories of Actions: Formal Theories and Practical Applications.*

[Gelfond and Lifschitz, 1992] Gelfond M. and Lifschitz V., 1992. Representing actions in extended logic programs. In *Joint International Conference and Symposium on Logic Programming*, pages 559–573.

[Grosof, 1991] Grosof B., 1991. Generalizing Priorization. In Proc. of *KR - Principles of Knowledge Representation and Reasoning*, pages 289–300. Morgan Kauffman.

[Kakas and Miller, 1996] Kakas A. and Miller R., 1996. A simple Declarative Language for Describing Narratives With Actions. To appear on *Journal of Logic Programming,* special issue on actions.

[Kowalski and Sergot, 1986] Kowalski R. and Sergot M., 1986. A logic-based calculus of events. *New Generation Computing*, 4:67–95.

[Lifschitz, 1993] Lifschitz V., 1993. Circumscription. In *Handbook of Logic in Artificial Intelligence and Logic Programming,* pages 297–352. Oxford Un. Press.

[Lifschitz, 1995] Lifschitz V., 1995. Nested Abnormality Theories. *Artificial Intelligence*, 74:351-365.

[Lin and Reiter, 1995] Lin F. and Reiter R., 1995. How to Progress a Database II: the STRIPS Connection. *Proc. the of 14th IJCAI Conference*, pages 2001-2007.

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

[McCarthy and Hayes, 1969] McCarthy J. and Hayes P., 1986. Some philosophical problems from the standpoint of artificial intelligence. In Meltzer B. and Michie D., editors, *Machine Intelligence*, 4:463–502. Edinburgh University Press.

[Miller and Shanahan, 1994] Miller R. and Shanahan M., 1994. Narratives in the Situation Calculus. *Journal of Logic and Computation,* 4(5):513–530.

[Pinto and Reiter, 1993] Pinto J. and Reiter R., 1993. Temporal reasoning in logic programming: A case for the situation calculus. In *Proc. of 10th International Conference in Logic Programming*, pages 203–221.

[Shanahan, 1995] Shanahan M., 1995. A Circumscriptive Calculus of Events. *Artificial Intelligence*, 77:249–284.

A Overview of Nested Circumscription

This abstract has been included for referee convenience.

Nested Abnormality Theories (NATs) are a novel circumscription [McCarthy, 1986, Lifschitz, 1993] technique introduced in [Lifschitz, 1995]. NATs allow the use of several abnormality predicates to specify a body of common-sense knowledge without their circumscription conflicting. Lifschitz introduces the notion of *block:* a set of axioms $A_1, \ldots, A_n$—possibly containing the abnormality predicate Ab— "describes" a set of predicates/constants $C_1, \ldots, C_m$, which are allowed to vary during circumscription of Ab. The notation for such a theory is

$$C_1, \ldots, C_m \ : \ A_1, \ldots, A_n. \qquad (5)$$

where each A_i may itself be a block of form (5). A block may be viewed as a set of axioms describing a collection of predicates/constants, possibly depending on other descriptions in embedded blocks. Block (5) can be expressed in terms of the circumscription operator as follows:

$$\text{CIRC}[A_1 \ \wedge \ldots \ \wedge A_n; Ab; C_1, \ldots, C_m]. \qquad (6)$$

In addition to NATs, Lifschitz [Lifschitz, 1995] introduces the idea of replacing the predicate *"Ab"* by an existentially quantified variable. Noticing that Ab plays an auxiliary role and that the interesting consequences of the theory are those which *do not* contain Ab, Ab is replaced by an existentially quantified variable. To put it differently, if $F(Ab)$ denotes (6), and *"ab"* is a predicate variable of the same arity as Ab, we are interested in the consequences of the sentence $\exists ab F(ab)$. The effect of this modification is that abnormality predicates become local to the block where each of them is used.

A.1 Syntax and semantics of NATs

The following definitions are from [Lifschitz, 1995]. Let L be a second order language which does not include Ab. For every natural number k, let L_k be the language obtained by adding the k-ary predicate constant Ab to L. $\{C_1, ..., C_m \ : \ A_1, ..., A_n\}$ is a *block* if each $C_1, \ldots, C_m$ is a predicate or a function constant of L, and each $A_1, \ldots, A_n$ is a formula of L_k or a block.

A *Nested Abnormality Theory* is a set of blocks. The semantics of NATs is characterized by a mapping φ from blocks into sentences of L. If A is a formula of language L_k, φA stands for the universal closure of A, otherwise

$$\varphi\{C_1, ..., C_m \ : \ A_1, ..., A_n\} \ = \ \exists ab F(ab)$$

where

$$F(Ab) = \text{CIRC}[\varphi A_1 \wedge ... \wedge \varphi A_n; Ab; C_1, ..., C_m]$$

For any NAT T, φT stands for $\{\varphi A | A \in T\}$. A *model* of T is a model of φT in the sense of classical logic. A *consequence* of T is a sentence ϕ of language L that is true in all models of T. In this paper, as suggested in [Lifschitz, 1993], we will use the abbreviation

$$\{C_1, \ldots, C_m, \ min \ P : A_1, \ldots, A_n\}$$

to denote blocks of the form

$$\{C_1, \ldots, C_m, P : P(x) \supset Ab(x), A_1, \ldots, A_n\}$$

As the notation suggests, this type of block is used when it is necessary to circumscribe a particular predicate P in a block. In [Lifschitz, 1993] it is shown that

$$\varphi\{C_1, \ldots, C_m, \ min \ P : A_1, \ldots, A_n\}$$

is equivalent to the formula

$$\text{CIRC}[A_1 \wedge \ldots \wedge A_n; P; C_1, \ldots, C_m]$$

when each of A_i is a sentence.

For the sake of our investigation, we are specially interested in NATs for which an equivalent first-order formalization can be given. The idea is to *simplify* a theory of nested blocks by substituting the innermost blocks with their equivalent first-order theories, and repeat the process inside out. The final theory —if it exists— can be used for deduction or comparison purposes. Even when a first-order equivalent exists, the NAT is more suggestive and compact.