Default-Reasoning with Models

Roni Khardon* **Dan Roth†**

Aiken Computation Laboratory,
Harvard University,
Cambridge, MA 02138.
{roni,danr}@das.harvard.edu

Abstract

Reasoning with model-based representations is an intuitive paradigm, which has been shown to be theoretically sound and to possess some computational advantages over reasoning with formula-based representations of knowledge. In this paper we present more evidence to the utility of such representations.

In real life situations, one normally completes a lot of missing "context" information when answering queries. We model this situation by augmenting the available knowledge about the world with context-specific information; we show that reasoning with model-based representations can be done efficiently in the presence of varying context information. We then consider the task of default reasoning. We show that default reasoning is a generalization of reasoning within context, in which the reasoner has many "context" rules, which may be conflicting. We characterize the cases in which model-based reasoning supports efficient default reasoning and develop algorithms that handle efficiently fragments of Reiter's default logic. In particular, this includes cases in which performing the default reasoning task with the traditional, formula-based, representation is intractable.

Further, we argue that these results support an incremental view of reasoning in a natural way.

1 Introduction

The generally accepted framework for studying reasoning in intelligent systems is the knowledge-based system approach. The idea is to store the knowledge in some *representation language* with a well defined meaning assigned to its sentences. The sentences are stored in a Knowledge Base (KB) which is combined with a reasoning mechanism that can be used to determine what can be inferred from the sentences in the KB. There are many knowledge representations that can be used to represent the knowledge in a knowledge-based system. Different representation systems (e.g., a set of logical rules, a probabilistic network) are associated with corresponding reasoning mechanisms, each with its own merits and range of applications. Given a logical knowledge base, for example, reasoning can be abstracted as a deduction task: determine whether a sentence, assumed to capture the situation at hand, is logically implied by the knowledge base. In all cases, the emphasis of this approach is on comprehensibility [McCarthy and Hayes, 1969; Pearl, 1988]: knowledge should be encoded so that it is readily accessible.

It is widely acknowledged today that a large part of our everyday reasoning involves arriving at conclusions that are not entailed by our "theory" of the world. Many conclusions are derived in the absence of information that is sufficient to imply them. This type of reasoning is naturally non-monotonic since further evidence may force us to revise our conclusions. Within the knowledge-based systems approach this situation is handled by theories for reasoning with "defaults" (see e.g. [Reiter, 1987a]). The true knowledge about the world is augmented by a set of *default rules* that capture only "typical" cases. The quest is for a reasoning system that, given a query, responds in a way that agrees with what we know about the world and the default assumptions and at the same time supports our intuition about a plausible conclusion.

Computational considerations, however, render this self-contained approach to reasoning inadequate for commonsense reasoning. This is true not only for the task of deduction, but also for many other forms of reasoning which have been developed. All those were shown to be even harder to compute than the original formulation [Selman, 1990; Roth, 1993]. Of particular interest in this context are the results on default reasoning tasks [Selman and Kautz, 1990; Kautz and Selman, 1991; Papadimitriou, 1991], where the increase in complexity is clearly at odds with the intuition that reasoning with defaults should somehow reduce the complexity of reasoning. This remains true, even when we severely restrict the expressivity of the knowledge base, the default rules and the queries allowed. For example, when the knowledge base is Horn, all the default rules are positive

*Research supported by Center for Intelligent Control Systems, ARO contract DAAL03-92-G-0115.

†Research supported by NSF grant CCR-92-00884 and by DARPA AFOSR-F4962-92-J-0466.

literals, and the query is a single positive literal, the default reasoning task is NP-Hard [Selman and Levesque, 1990]. This should be contrasted with the case of deductive reasoning, where Horn theories are distinguished by the existence of linear time satisfiability algorithms.

An alternative approach to the study of commonsense reasoning is developed in [Kautz *et al.*, 1995; Khardon and Roth, 1994c]. There, the knowledge base is represented as a set of models (satisfying assignments) of the world rather than a logical formula describing it. It is not hard to motivate a model-based approach to reasoning from a cognitive point of view and indeed, most of the proponents of this approach to reasoning have been cognitive psychologists [Johnson-Laird, 1983; Johnson-Laird and Byrne, 1991; Kosslyn, 1983], who have alluded to the notion of "reasoning from examples" on a qualitative basis. In the AI community this approach can be seen as an example of Levesque's notion of "vivid" reasoning [Levesque, 1986; 1992], and is somewhat related to Minsky's frames-theory [Minsky, 1975].

Given a model-based representation of the knowledge base KB and a query α, the deduction task $KB \models \alpha$ can be answered in a straightforward way: Evaluate α on all the models in the representation. If you find a model of KB which does not satisfy α, then $KB \not\models \alpha$, otherwise conclude $KB \models \alpha$. Clearly, if the model-based representation contains all the models of KB this approach yields correct deduction, but representing KB by explicitly holding *all* the possible models is not plausible. A model-based approach becomes feasible if KB can be replaced by a small model-based representation and still support correct deduction.

The theory of model-based representations developed in [Khardon and Roth, 1994c] (generalizing the theory developed in [Kautz *et al.*, 1995] for the case of Horn expressions) characterizes the propositional languages for which model-based representations support efficient deduction and abduction. It is shown that in many cases in which the deduction and abduction tasks are NP-Hard in the formula-based setting, the model-based representation is small (polynomial in the number of propositional variables in the domain), and reasoning with it yields correct and efficient reasoning algorithms.

In this paper, we extend the work presented in [Khardon and Roth, 1994c] and present some more computational advantages of reasoning with model-based representations. As a basic computational task we consider the problem of *reasoning within a varying context*. In real life situations, one normally completes a lot of missing context information when answering queries [Levesque, 1986]. We model this situation by augmenting the knowledge we have about the world with context-specific information. Reasoning within context is therefore a deduction task, where some additional constraining information is added to the knowledge base. We show how to solve this task efficiently using a model-based representation, for a variety of propositional languages as context information.

We then consider the task of *default reasoning*. There, given a representation of the world, a set of (sometimes conflicting) default rules and an assertion q, one is trying to asses whether q can be concluded "by default" from the available information. We show that default reasoning is a generalization of reasoning within context, in which the reasoner has many context rules, which may be conflicting. We provide an efficient algorithm for the default reasoning task, for various classes of world knowledge, default rules and queries, based on the algorithm developed for reasoning within context.

As in the case of deductive and abductive reasoning [Khardon and Roth, 1994c], we present an efficient default reasoning algorithm for cases where the formula based reasoning is hard. For example, in contrast to the hardness result mentioned above, we show that if the knowledge base is any propositional language with a polynomial size DNF[1], the default rules are arbitrary monotone functions and the query is a Horn query, the default reasoning task can be solved correctly and efficiently.

Equally important for the plausibility of model based reasoning is the view that it suggests about reasoning. While we do not consider in the paper the question of how the knowledge base is acquired, this issue is clearly an important one, and the plausibility of any theory for reasoning hinges on it. It is important therefore to mention that it has been shown, within the Learning to Reason framework [Khardon and Roth, 1994b; 1995], that model-based representations that are suitable for the reasoning tasks considered here can be learned efficiently. The model based approach to default reasoning can therefore be incorporated within an inductive setting. The model based representation can be efficiently learned, context specific default rules can be acquired in various learning processes, and these can be combined to work together in a plausible and efficient way. Furthermore, we show how knowledge available within a specific context can be used to reason within this context. Therefore, our treatment of reasoning within context supports the view that an intelligent agent constructs a representation of the world incrementally by pasting together many "narrower" views from different contexts.

The inductive nature of non-monotonic reasoning is also at the heart of the approach developed in [Valiant, 1994; Roth, 1995], where a different view on dealing with incomplete information is taken.

2 Preliminaries

We consider problems of reasoning where the "world" is modeled as a Boolean function $W : \{0,1\}^n \rightarrow \{0,1\}$. We use interchangeably the terms propositional expression and Boolean function, and likewise for propositional language and a class of Boolean functions. We denote classes of Boolean functions by $\mathcal{F}$, $\mathcal{G}$, and functions by f, g.

We consider a set $X = \{x_1, \ldots, x_n\}$ of *variables*, each of which is associated with a world's attribute and can

[1]The size of the model-based representation of KB is related to the size of its minimal DNF. Thus, we do not assume that the DNF representation is known but only require that a polynomial size representation exists.

take the value 1 or 0 to indicate whether the associated attribute is true or false in the world.

Assignments are mappings from X to $\{0,1\}$, and we treat them as elements in $\{0,1\}^n$ with the natural mapping. Assignments in $\{0,1\}^n$ are denoted by x, y, z, and $weight(x)$ denotes the number of 1 bits in the assignment x. A *clause* is a disjunction of literals, and a CNF formula is a conjunction of clauses. For example $(x_1 \vee \overline{x_2}) \wedge (x_3 \vee \overline{x_1} \vee x_4)$ is a CNF formula with two clauses. A *term* is a conjunction of literals, and a DNF formula is a disjunction of terms. For example $(x_1 \wedge \overline{x_2}) \vee (x_3 \wedge \overline{x_1} \wedge x_4)$ is a DNF formula with two terms. A CNF formula is *monotone* if all the literals in it are positive (unnegated). A CNF formula is Horn if every clause in it has at most one positive literal. A CNF formula is k-quasi-Horn if there are at most k positive literals in each clause. It is a k-quasi-reversed-Horn if there are at most k negative literals in each clause. A DNF formula is k-quasi-monotone DNF if there are at most k negative literals in each term.

Every Boolean function has many possible representations, and in particular both a CNF representation and a DNF representation. By the DNF size of f, denoted $|DNF(f)|$, we mean the number of terms in the minimal DNF representation of f. (Similarly, for $|CNF(f)|$.)

An assignment $x \in \{0,1\}^n$ satisfies f if $f(x) = 1$. (x is also called a model of f.) If f is a theory of the "world", a satisfying assignment of f is sometimes called a *possible world*. By "f implies g", denoted $f \models g$, we mean that every model of f is also a model of g. Throughout the paper, when no confusion can arise, we identify a Boolean function f with the set of its models, namely $f^{-1}(1)$. Observe that the connective "implies" ($\models$) used between Boolean functions is equivalent to the connective "subset or equal" ($\subseteq$) used for subsets of $\{0,1\}^n$. That is, $f \models g$ if and only if $f \subseteq g$.

3 Reasoning with Models

Consider a propositional knowledge base W and let α be a propositional query. The deduction problem $W \models \alpha$ can be approached using the following model-based strategy:

Algorithm MBR(Γ, α):

Test Set: A set $\Gamma \subseteq W$ of possible assignments.

Test: If there is an element $x \in \Gamma$ which does not satisfy α, return "NO". Otherwise, return "YES".

Clearly, this approach solves the inference problem if Γ is the set of *all* models (satisfying assignments) of W, but this set might be too large. A model-based approach becomes useful if one can show that it is possible to use a fairly small set of models as the Test Set, and still perform reasonably good inference.

This section briefly introduces the monotone theory of Boolean functions [Bshouty, 1993], and the theory of reasoning with models[2] (see [Khardon and Roth, 1994c] for more details).

[2] We note that this direction was studied independently in the Relational Data Base community [Beeri *et al.*, 1984; Mannila and Raiha, 1986]. The results on model-based reasoning have immediate implications in this domain which are described elsewhere [Khardon *et al.*, 1995].

Definition 1 (Order) *We denote by* $\leq$ *the* usual partial order *on the lattice* $\{0,1\}^n$*, the one induced by the order* $0 < 1$*. That is, for* $x, y \in \{0,1\}^n$*,* $x \leq y$ *if and only if* $\forall i, x_i \leq y_i$*. For an assignment* $b \in \{0,1\}^n$ *we define* $x \leq_b y$ *if and only if* $x \oplus b \leq y \oplus b$ *(Here* $\oplus$ *is the bitwise addition modulo 2). We say that* $x > y$ *if and only if* $x \geq y$ *and* $x \neq y$*.*

Intuitively, if $b_i = 0$ then the order relation on the ith bit is the normal order; if $b_i = 1$, the order relation is reversed, that is, $1 <_{b_i} 0$.

Next we define: The *monotone extension of* $z \in \{0,1\}^n$ with respect to b:

$$\mathcal{M}_b(z) = \{x \mid x \geq_b z\}.$$

The *monotone extension of* f with respect to b:

$$\mathcal{M}_b(f) = \{x \mid x \geq_b z, \ for\ some\ z \in f\}.$$

The set of *minimal assignments of* f with respect to b:

$$\min_b(f) = \{z \mid z \in f, \ \ such\ that\ \forall y \in f, z \not>_b y\}.$$

Definition 2 (Basis) *A set* B *is a* basis for f *if* $f = \bigwedge_{b \in B} \mathcal{M}_b(f)$*.* B *is a basis for a class of functions* $\mathcal{F}$ *if it is a basis for all the functions in* $\mathcal{F}$*.*

The importance of these definitions is that one can show that every Boolean function has a basis B, and can be represented as follows:

$$f = \bigwedge_{b \in B} \mathcal{M}_b(f) = \bigwedge_{b \in B} \bigvee_{z \in \min_b(f)} \mathcal{M}_b(z) \qquad (1)$$

This representation yields a necessary and sufficient condition describing when $x \in \{0,1\}^n$ is positive for f:

Corollary 1 *Let* B *be a basis for* f*,* $x \in \{0,1\}^n$*. Then,* $x \in f$ *(i.e.,* $f(x) = 1$*) if and only if for every basis element* $b \in B$ *there exists* $z \in min_b(f)$ *such that* $x \geq_b z$*.*

It is known that for every b, the size of $\min_b(f)$ is bounded by the size of its DNF representation. Further, a set of assignments which falsify every clause in a CNF representation of f is a basis for f. Therefore, f has a basis whose size is bounded by $|CNF(f)|$. Some important function classes have a small *fixed* basis, irrespective of the CNF size of the function:

Horn-CNF formulas: The basis for this class is $B_H = \{u \in \{0,1\}^n \mid \text{weight}(u) \geq n-1\}$, since every Horn clause is falsified by an assignment in B_H. Clearly, $|B_H| = n+1$.

k-quasi-Horn formulas:
$B_{H_k} = \{u \in \{0,1\}^n \mid \text{weight}(u) \geq n-k\}$ is a basis. Clearly, $|B_{H_k}| = O(n^k)$. Similarly, there is a basis for k-quasi-reversed-Horn formulas.

$\log n$CNF formulas: CNF in which the clauses contain at most $O(\log n)$ literals. The basis for this class is derived using a combinatorial construction called an (n,k)-universal set [Alon *et al.*, 1992]. It can be shown that $|B_{\log n-CNF}| = O(n^3)$.

Common queries: A function is *common* if every clause in its CNF representation is taken from one of the above classes. The union of the bases for these classes is a basis, B_C, for all common functions. We refer to this class as the class of common queries.

3.1 Deduction

We can now to characterize a model-based knowledge base for which the algorithm **MBR** is successful.

Definition 3 *For a knowledge base $f \in \mathcal{F}$ the set $\Gamma = \Gamma_f^B$ of* characteristic models *of f, is the set of all minimal assignments of f with respect to the basis B. Formally,*

$$\Gamma_f^B = \cup_{b \in B}\{z \in min_b(f)\}.$$

The following are the basic theorem of the theory of reasoning with models, its application to common queries, and a bound on the size of the model-based representation.

Theorem 1 *Let f be any Boolean function, and $\alpha \in \mathcal{G}$, where B be a basis for $\mathcal{G}$. Then $f \models \alpha$ if and only if for every $u \in \Gamma_f^B$, $\alpha(u) = 1$.*

Theorem 2 *Let f be any Boolean function. Then for any common query, model-based deduction using $\Gamma_f^{B_C}$, is correct.*

Theorem 3 *Let f be any Boolean function, and B a basis. Then, the size of the model-based representation of f is*

$$|\Gamma_f^B| \leq \sum_{b \in B} |min_b(f)| \leq |B| \cdot |DNF(f)|.$$

We note that this bound is tight in the sense that for some functions the size of the DNF is indeed needed. It does however allow for an exponential gap in other cases. Namely, there are functions with an exponential size DNF and a linear size model-based representation [Khardon and Roth, 1994c]. It is also interesting to compare the size of this representation to the size of other representations for functions. Examples in [Kautz *et al.*, 1995] show that there are cases where the (Horn CNF) formula representation is small and the model-based representation is exponentially large, and vice versa. For a discussion of these issues see [Khardon and Roth, 1994c].

Example: Let f have the CNF representation:

$$f = (x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_2 \vee x_4) \wedge (\overline{x_1} \vee \overline{x_2} \vee x_3 \vee \overline{x_4})$$

The function f has 12 (out of the 16 possible) satisfying assignments. The non-satisfying assignments of f are[3]: $\{0000, 0001, 0010, 1101\}$.

If we want to be able to answer all possible Horn queries with respect to f we need to use the Horn basis $B_H = \{1111, 1110, 1101, 1011, 0111\}$. Each of the models 1111,0111,1011,1110 satisfies f and therefore for each of these, $\min_b(f) = b$. For $b = 1101$, the minimal elements can be found by drawing the corresponding lattice and checking which of the satisfying assignments of f are minimal. This yields $\min_{1101}(f) = \{1100, 1111, 1001, 0101\}$. We therefore get that $\Gamma_f^{B_H} = \{1111, 0111, 1011, 1100, 1001, 0101, 1110\}$. Note that it includes only 7 out of the 12 satisfying assignments of f.

Clearly, in general $\Gamma_f^{B_H} \subseteq f$, and therefore model-based deduction never makes mistakes on queries that are implied by f. Furthermore, for the Horn query $\alpha_1 = x_1 \wedge x_3 \rightarrow x_2$, reasoning with $\Gamma_f^{B_H}$ will find the counterexample 1011 and deduce correctly that $f \not\models \alpha_1$.

We note that in general, if f is given in its CNF representation, solving the problem $f \models \alpha$ is co-NP-complete, even when α is a Horn query.

[3] An element of $\{0,1\}^n$ denotes an assignment to the variables $x_1, \ldots, x_n$ (i.e., 0011 means $x_1 = x_2 = 0$, and $x_3 = x_4 = 1$).

4 Reasoning within Context

It has been argued that in real life situations, one normally completes a lot of missing "context" information when answering queries [Levesque, 1986]. For example, if asked at the conference how long it takes to drive to the airport, we would probably assume (unless specified otherwise) that the question refers to the city we are at now, Montreal, rather than where we live (and have been to the airport more times). This corresponds to assigning the value "true" to the attribute "here" for the purpose of answering the question. Sometimes we need a more expressive language to describe our assumptions regarding the current context and assume, say, that some rule applies [Selman and Kautz, 1990]. For example, we may assume (in the "conference" context) that if someone has a car, then it is a rental car.

Let W be a Boolean function that describes our knowledge about the world. A "first principle" way to formalize the above intuition is the following: we want to deduce a query α from W, if α can be inferred from W given that the query refers to the current context. Namely, the instances of W which are relevant to the query must also satisfy the context condition d, a conjunction of some literals and rules. We denote this question by $W \models_d \alpha$.

Notice that it is possible that $W \models_d \alpha$ but $W \not\models \alpha$, if all the satisfying assignments of W that do not satisfy α do not satisfy d. Formalized this way, we get that the problem $W \models_d \alpha$ is equivalent to the problem $W \wedge d \models \alpha$. Thus, a theorem proving approach to reasoning does not give any computational advantage in solving this reasoning problem.

Let $W \in \mathcal{F}$, $\alpha \in \mathcal{G}$ and let B be a basis for $\mathcal{G}$. From Theorem 1 it is clear that given $\Gamma_{W \wedge d}^B$, the set of characteristic models for $W \wedge d$, model-based reasoning can be used to solve the reasoning problem $W \wedge d \models \alpha$. However, we consider here a more general problem: given Γ_W^B we are interested in performing inference according to $\models_d$ with it, where the "context condition" d may vary.

From our model-theoretic definition of the connective $\models_d$ it is clear that if one holds *all* the models of W, then by filtering out all the models that do not satisfy d and then performing the model-based test one can answer $W \wedge d \models \alpha$ correctly. The algorithm C-**MBR** does just that, with the set Γ:

Algorithm C-MBR(Γ, d, α):

- *Test Set:* Consider only those elements of Γ which satisfy d.
- *Test:* If there is such an element which does not satisfy α, return "NO". Otherwise, return "YES".

The following theorems characterize the cases in which this algorithm, C-MBR, provides correct reasoning.

Theorem 4 *Given Γ_W^B, the algorithm C-MBR provides an exact solution to the reasoning problem $W \models_d \alpha$ for every d such that B is a basis for $d \rightarrow \alpha$.*

Proof: Clearly, $W \models_d \alpha \equiv W \wedge d \models \alpha \equiv W \models \overline{d} \vee \alpha \equiv W \models (d \rightarrow \alpha)$. Therefore, from Theorem 1, when B is a basis for $d \rightarrow \alpha$, Γ_W^B can be used for model-based reasoning with it. Models of W that do not satisfy d are useless as counterexamples since $d \rightarrow \alpha$ always holds and therefore, using the Test Set of *Algorithm C-MBR* produces the correct inference. ■

Theorem 5 *The following conditions on α, B and d guarantee that C-MBR supports correct reasoning within context.*
(i) Let α be a Horn query and $B = B_H$, a basis for Horn theories. If d is a monotone Boolean function then B is a basis for $d \rightarrow \alpha$.
(ii) Let α be a k-quasi-Horn query and $B = B_{H_{k+l}}$, a basis for $(k+l)$-quasi-Horn theories. If d is a Boolean function that can be represented as a l-quasi-monotone DNF then B is a basis for $d \rightarrow \alpha$.
(iii) Let α be a $\log n$CNF query and B a basis for $2\log n$CNF theories. If d is a conjunction of up to $\log n$ arbitrary rules (or disjunctions) then B is a basis for $d \rightarrow \alpha$.

Proof: For (i), consider first the case in which $d = \bigwedge_{i \in I} x_i$. Let $\alpha = \bigwedge_{j \in J}(m_j \rightarrow x_j)$ be a Horn query (that is, m_j is a monotone conjunction). Then

$$d \rightarrow \alpha \equiv \overline{d} \vee \alpha \equiv \overline{d} \vee \bigwedge_{j \in J} (\overline{m_j} \vee x_j) \equiv \bigwedge_{j \in J} ((d \wedge m_j) \rightarrow x_j),$$

which means that $d \rightarrow \alpha$ is a Horn expression. In general, if d' is a monotone function, $d' = \bigvee_\ell d_\ell$ where d_ℓ are monotone conjunctions. Since $\forall f, g, h,\ (f \vee g) \rightarrow h \equiv (f \rightarrow h) \wedge (g \rightarrow h)$, we get that $d' \rightarrow \alpha \equiv \bigwedge_\ell (\bigwedge_{j \in J}((d_\ell \wedge m_j) \rightarrow x_j))$ is a Horn theory[4]. For (ii), in the same way, we get that the body of the rule might contain up to $k+l$ non-negated literals. For (iii), similar manipulations show that $d \rightarrow \alpha$ is a $2\log n$CNF. ■

The approach presented in this section can be viewed as a process of *augmenting* a model-based representation Γ with a set of rules. Given a model-based representation Γ_W^B of W, any rule that holds in W cannot help in answering queries, since it does not filter out any assignment of W, and is thus redundant. However, the context rules do not hold in W and thus augmenting W with them modifies the set of conclusions. As we have shown, in order to reason within context, we need to maintain a model-based representation with respect to a basis that is slightly larger than in the pure deductive case.

5 Default Reasoning with Models

In the previous section we assumed that the context information is given. In general, one might have many context rules, which may be conflicting, and the goal is to derive plausible conclusions in the face of this information.

This is the situation modeled in default reasoning. Given a representation of the world, a set of (sometimes conflicting) default rules, and an assertion q, one is trying to asses whether q can be concluded by default from the available information. We will see that this is a generalization of the situation discussed in the previous section, and that it can also be dealt with efficiently using model-based representations. We shall concentrate here on a special case of Reiter's Default logic [Reiter, 1980], applied to propositional logic and with some restrictions on the default rules.

In Reiter's default logic, default rules have the form $\alpha : \beta/\gamma$, which should read as "if α holds and it is consistent to assume β then conclude γ". The case with $\beta = \gamma$ is called *normal defaults*, and α is called a prerequisite. The discussion below considers normal defaults with empty prerequisites (i.e. of the form $: \beta/\beta$). In this case, we denote by D the set of Boolean functions β, and say that D is the set of default rules. We sometimes treat a collection of rules as their conjunction. That is, $D(x) = 1$ means $\wedge_{d \in D} d(x) = 1$.

Definition 4 *A default rule is* simple *if it is of the form $: \beta/\beta$, and β is a single literal. The rule is* positive *if β is any monotone function. The rule is* positive simple *if β is a positive literal.*

Notice that the theory for diagnosis [Reiter, 1987b] and the closed world defaults [Reiter, 1980] can be described using simple defaults.

A *default theory* is a pair (D, W) where D is a set of default rules, and W is propositional expression. An *extension* of (D, W) is defined as a fixed point of some operator. For our special case the following theorem gives an alternative and simpler definition: (The operator $Th(R)$ denotes the theorem closure of R.)

Theorem 6 ([Reiter, 1987b], page 88) *Let D be a set of normal defaults with empty prerequisite. E is an extension of (D, W) iff $E = Th(W \cup \{\beta | \ : \beta/\beta \in S\})$, and S is a maximal subset of D such that $E = W \cup \{\beta | \ : \beta/\beta \in S\}$ is consistent.*

Using this theorem as the definition for extension we can identify a maximal consistent subset S with each extension E. We denote this subset by S_E. Given that S_E is consistent with W, we get that an extension E includes q iff $W \wedge S_E \models q$.

The *default reasoning task* $DEF(D, W, q)$ is defined as follows: given a default theory (D, W) and a propositional expression q, decide whether there exists an extension E of (D, W) such that $q \in E$.

Clearly, if W is consistent with the set of all rules in D, then there exists only one maximal consistent subset S of D, the one which contains all these rules. This case simply reduces to reasoning within the context D, which we discussed earlier. The main difficulty which arises in the general case is that W may not be consistent with all of D.

Next we present positive results on default reasoning. As in the deductive reasoning case, the efficient results

[4] We note that the size of the resulting Horn theory might be exponentially large, but it only appears in the analysis. We do not actually compute this expression in the algorithm. Rather, filtering examples according to d is sufficient. The same holds for the other cases.

we present hold in cases where the formula based reasoning is hard[5]. The following algorithm, D-MBR, is based on the abduction algorithm[6] developed in [Kautz *et al.*, 1995] and used in [Khardon and Roth, 1994c].

Algorithm D-MBR(Γ, D, q):

Do for all models $z \in \Gamma$ such that $q(z) = 1$

Let $S = \{d \in D \ : \ d(z) = 1\}$

If C-MBR(Γ, S, q) answers "YES", return "YES".

EndDo

Return "NO". /* No extension found */

Assume that $\Gamma = \Gamma_W$, a model based representation of W. The algorithm D-MBR receives Γ, D and a query q as input. It starts by enumerating all the models in Γ. When it finds a model z in which the query holds, (i.e., $q(z) = 1$) it sets S to be the set of all the rules in D that this model satisfies. The algorithm then tests whether $W \wedge S \models q$ by calling the procedure C-MBR to decide whether $W \models_S q$. If the answer is "yes" the algorithm returns "yes" (and possibly the extension S), and otherwise it continues to test the next model in Γ. If all the models in Γ have been scanned and no good extension has been found the algorithm says "no".

Theorem 7 *Let W be any propositional expression, D a set of positive defaults, and $\mathcal{Q}$ the class of Horn queries. Then, for all $q \in \mathcal{Q}$, on input $(D, \Gamma_W^{B_H}, q)$, the algorithm D-MBR solves the default reasoning task $DEF(D, W, q)$ correctly.*

Proof: Assume first that there exists extension E that contains q. We show that the algorithm answers "YES". By definition, the existence of the extension implies that $W \wedge S_E \models q$ and that there exists an assignment $u \in W$ such that $S_E(u) = 1$ and $q(u) = 1$. Let $z \in min_{1^n}(W)$ such that $z \leq_{1^n} u$. Then $S_E(z) = 1$, as S_E is monotone. Since $W \wedge S_E \models q$ we have $q(z) = 1$. Since $z \in \Gamma_W^{B_H}$, the algorithm D-MBR will use it, and set S_E correctly (since S_E is a maximal set of elements from D). By (i) of Theorem 5 the algorithm C-MBR answers "Yes" when called with $\Gamma_W^{B_H}$, S_E and q, and therefore D-MBR answers "YES". Assume now that D-MBR returns "YES". Since q is Horn and S is a set of monotone functions, Theorem 5 implies that C-MBR is correct, that is $W \wedge S \models q$. By construction, S is a maximally consistent set and therefore the corresponding extension contains q. ∎

[5]We note that the default reasoning task is NP-Hard [Selman and Levesque, 1990] when the knowledge base is Horn, all the default rules are positive literals, and the query is a single positive literal. Our results provide an algorithm for this class of problems, which is polynomial in the size of the model based representation. The latter though may be exponential in the size of the Horn expression, and in particular this happens for the problems used in the reduction in [Selman and Levesque, 1990]. So strictly speaking we do not prove an advantage in this special case. Our results, however, provide efficient algorithms in cases where they were not known to exist before.

[6]Our results were inspired by the connections between abduction and default reasoning developed in [Selman, 1990].

Similarly, we can show:

Theorem 8 *Let W be any propositional expression, D a set of simple defaults with up to k-1 negative literals, and $\mathcal{Q}$ the class of Horn queries. Then, for all $q \in \mathcal{Q}$, on input $(D, \Gamma_W^{B_{H_k}}, q)$, the algorithm D-MBR solves the default reasoning task $DEF(D, W, q)$ correctly.*

Proof: Let E be an extension and S_E its corresponding maximal consistent subset of D. The element b, defined by: $b_i = 0$ when $\overline{x_i} \in S_E$, otherwise, $b_i = 1$, is in the basis B_{H_k}.

Note that if $S_E(u) = 1$ and $z \in min_b(W)$ such that $z \leq_b u$, then $S_E(z) = 1$ since the assignments to variables in S_E are the same in u and z. Therefore the proof of Theorem 7 applies here as well (replacing the basis element 1^n by b, and using (ii) of Theorem 5). ∎

Using Theorem 4 we can find other instances in which correct default reasoning is possible. For example, we could answer queries that are more general than Horn queries. In particular, if D is simple and the literals in D are special (e.g., Abnormal predicates in diagnosis) and do not appear in the queries, then we can set a special basis that will handle every query which is either Horn or in $\log n$-CNF.

The above theorems yield efficient default reasoning whenever the model-based representation Γ is small. (And, in particular, whenever the DNF size of f is small.) This should be contrasted with the hardness of deductive and default reasoning.

6 Discussion

Reasoning with models is an intuitive paradigm, which has been shown to be theoretically sound. In this paper we presented more evidence to the utility of such representations. In particular, these representations support efficient reasoning in the presence of varying context information, as well as some restricted cases of default reasoning. The significance of these results is that they are achieved as natural extensions of exact (deductive) reasoning, and hold in cases in which the traditional formula-based representation does not support efficient reasoning.

These results can be viewed as providing some theoretical support for the usefulness of case-based style reasoning, where a set of "typical cases" is used as a knowledge representation.

We have shown that a model-based representation can be used to reason correctly when some additional constraining context information is supplied. This information augments the agents' knowledge and aids in deriving conclusions relevant to this context. We call this a top-down solution. It is conceivable, though, that an agent would have only *some* of the models, those models that come from some specific context d. In such a case, our results show that the agent reasons correctly within this context (although not within every context). This approach can be shown to work in other scenarios in which the agent constructs a model-based knowledge representation by randomly collecting examples in the environment [Khardon and Roth, 1994a]. Thus, the approach supports the view that an intelligent agent constructs a representation of the world incrementally by pasting together many "narrower" views from different contexts.

In default reasoning, an agent may have many (possibly conflicting) default rules, acquired in different contexts. Default reasoning is thus a generalization of reasoning within context where the additional information may not be consistent, and may not be consistent with the knowledge the agent has about the world. Indeed, a query holds "by default", if there is a plausible context in which it holds. As we have shown, model-based representations efficiently support default reasoning.

Finally, we mention that it has been shown, within the Learning to Reason framework [Khardon and Roth, 1994b], that the model based representations discussed here can be learned efficiently. This can be combined with context specific default rules that are acquired via rote learning or other learning processes [Schuurmans and Greiner, 1994] to work in a plausible way.

References

[Alon *et al.*, 1992] N. Alon, J. Bruck, J. Naor, M. Naor, and R. Roth. Construction of asymptotically good low-rate error-correcting codes through pseudo-random graphs. *IEEE Transactions on information theory*, 38(2):509–516, 1992.

[Beeri *et al.*, 1984] C. Beeri, M. Dowd, R. Fagin, and R. Statman. On the structure of Armstorng relations for functional dependencies. *Journal of the ACM*, 31(1):30–46, 1984.

[Bshouty, 1993] N. H. Bshouty. Exact learning via the monotone theory. In *Proceedings of the IEEE Symp. on Foundation of Computer Science*, pages 302–311, Palo Alto, CA., 1993.

[Johnson-Laird and Byrne, 1991] P. N. Johnson-Laird and R. M. J. Byrne. *Deduction*. Lawrence Erlbaum Associates, 1991.

[Johnson-Laird, 1983] P. N. Johnson-Laird. *Mental Models*. Harvard Press, 1983.

[Kautz and Selman, 1991] H. Kautz and B. Selman. Hard problems for simple default logics. *Artificial Intelligence*, 49:243–279, 1991.

[Kautz *et al.*, 1995] H. Kautz, M. Kearns, and B. Selman. Horn approximations of empirical data. *Artificial Intelligence*, 74:129–145, 1995.

[Khardon and Roth, 1994a] R. Khardon and D. Roth. Exploiting relevance through model-based reasoning. In *AAAI Fall Symposium on Relevance*, pages 109–114, 1994.

[Khardon and Roth, 1994b] R. Khardon and D. Roth. Learning to reason. In *Proceedings of the National Conference on Artificial Intelligence*, pages 682–687, 1994. Full version: TR-02-94, Aiken Computation Lab., Harvard University, January 1994.

[Khardon and Roth, 1994c] R. Khardon and D. Roth. Reasoning with models. In *Proceedings of the National Conference on Artificial Intelligence*, pages 1148–1153, 1994. Full version: TR-01-94, Aiken Computation Lab., Harvard University, January 1994.

[Khardon and Roth, 1995] R. Khardon and D. Roth. Learning to reason with a restricted view. In *Workshop on Computational Learning Theory*, July 1995.

[Khardon *et al.*, 1995] R. Khardon, H. Mannila, and D. Roth. Reasoning with examples: Propositional formulae and database dependencies. In preparation., 1995.

[Kosslyn, 1983] S. M. Kosslyn. *Image and Mind*. Harvard Press, 1983.

[Levesque, 1986] H. Levesque. Making believers out of computers. *Artificial Intelligence*, 30:81–108, 1986.

[Levesque, 1992] H. Levesque. Is reasoning too hard ? In *Proceeding of the 3rd NEC research Symposium*. 1992.

[Mannila and Raiha, 1986] H. Mannila and K. J. Raiha. Design by example: An application of Armstrong relations. *Journal of Computer and System Sciences*, 33(2):126–141, 1986.

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

[Minsky, 1975] M. Minsky. A framework for representing knowledge. In P. Winston, editor, *The Psychology of Computer Vision*. McGraw-Hill, 1975. Also in R. Brachman and H. Levesque, Readings in Knowledge Representation, 1985.

[Papadimitriou, 1991] C. H. Papadimitriou. On selecting a satisfying truth assignment. In *Proc. 32nd Ann. IEEE Symp. on Foundations of Computer Science*, pages 163–169, 1991.

[Pearl, 1988] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman, 1988.

[Reiter, 1980] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13(1,2), 1980.

[Reiter, 1987a] R. Reiter. Nonmonotonic reasoning. In *Annual Reviews of Computer Science*, pages 147–188. 1987.

[Reiter, 1987b] R. Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32(1), 1987.

[Roth, 1993] D. Roth. On the hardness of approximate reasoning. In *Proceedings of the International Joint Conference of Artificial Intelligence*, pages 613–618, August 1993. To Appear in Artificial Intelligence Journal, 1995.

[Roth, 1995] D. Roth. Learning to reason: The non-monotonic case. In *Proceedings of the International Joint Conference of Artificial Intelligence*, August 1995.

[Schuurmans and Greiner, 1994] D. Schuurmans and R. Greiner. Learning default concepts. In *Proceedings of the Tenth Canadian Conference on Artificial Intelligence (CSCSI-94)*, 1994.

[Selman and Kautz, 1990] B. Selman and H. Kautz. Model-preference default theories. *Artificial Intelligence*, 45:287–322, 1990.

[Selman and Levesque, 1990] B. Selman and H. Levesque. Abductive and default reasoning: A computational core. In *Proceedings of the National Conference on Artificial Intelligence*, pages 343–348, 1990.

[Selman, 1990] B. Selman. *Tractable Default Reasoning*. PhD thesis, Department of Computer Science, University of Toronto, 1990.

[Valiant, 1994] L. G. Valiant. Rationality. Technical Report TR-32-94, Aiken Computation Lab., Harvard University, November 1994.