

# Representation Theory for Default Logic

*V. Wiktor Marek<sup>1</sup> Jan Treur<sup>2</sup> and Mirosław Truszczyński<sup>3</sup>*

**Keywords:** default logic, extensions, normal default logic, representability

## Abstract

Default logic can be regarded as a mechanism to represent *families* of belief sets of a reasoning agent. As such, it is inherently second-order. In this paper, we study the problem of representability of a family of theories as the set of extensions of a default theory. We give a complete solution to the problem of representability by means of default theories with finite set of defaults, and by means of normal default theories. We obtain partial results on representability by arbitrary (infinite, non-normal) default theories. We construct examples of denumerable families of non-including theories that are not representable. We also study the concept of equivalence between default theories.

## 1 Introduction

In this paper we investigate the issues related to the expressibility of default logic, a knowledge representation formalism introduced by Reiter [Rei80] and extensively investigated by the researchers of logical foundations of Artificial Intelligence [Eth88, Bes89, Bre91]. A default theory  $\Delta$  describes a family (possibly empty) of belief sets of an agent reasoning with  $\Delta$ . In that, default logic is inherently second-order, but in a sense different from that used by logicians. Whereas a logical theory  $S$  describes a subset of the set of all formulas (specifically, the set  $Cn(S)$  of logical consequences of  $S$ ), a default theory  $\Delta$  describes a collection of subsets of the set of all formulas, namely the family of all extensions of  $\Delta$ ,  $ext(\Delta)$ . Hence, default theories can be viewed as encodings of families of subsets of some universe described by a propositional language. Examples of encodings describing families of colorings, kernels, and hamilton cycles in graphs are given in [CMMT95].

---

<sup>1</sup>Department of Computer Science, University of Kentucky, Lexington, KY 40506-0046, USA. E-mail: [marek@cs.engr.uky.edu](mailto:marek@cs.engr.uky.edu), Fax: +(606) 323-1971.

<sup>2</sup>Free University Amsterdam, Department of Mathematics and Computer Science, Artificial Intelligence Group, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands, [treur@cs.vu.nl](mailto:treur@cs.vu.nl).

<sup>3</sup>Department of Computer Science, University of Kentucky, Lexington, KY 40506-0046, USA, [mirek@cs.engr.uky.edu](mailto:mirek@cs.engr.uky.edu), Fax: +(606) 323-1971.

This second-order flavor of default logic makes it especially useful in knowledge representation. An important question is, then, to characterize those families of sets that can be represented as the set of extensions of a certain default theory. This is the topic of our paper.

There is a constraint on the family  $\mathcal{T}$  of extensions of a default theory  $\Delta$ . Namely that such a family must be *non-including* [Rei80]. In this paper we exhibit several classes of families of non-including theories that can be represented by default theories. We also show that there are non-representable families of non-including theories. The existential proof follows easily from a cardinality argument. There are continuum-many default theories in a given (denumerable) language, while there is more than continuum-many families of non-including theories. In the paper, we actually *construct* a family of non-including theories that can not be represented by a default theory. Moreover, our family is denumerable (the cardinality argument mentioned above does not guarantee the existence of a non-representable denumerable family of non-including theories). However, as one of our results, we prove that finite non-including families are always representable. Moreover, if each theory in such family is finitely generated over the intersection of the family, one can select a default representation with a finite set of defaults.

The family of extensions of a *normal* default theory is not only non-including, but all its members are pairwise inconsistent [Rei80]. In this paper, we fully characterize these families of theories which are of the form  $ext(\Delta)$ , for a normal default theory  $\Delta$ . In addition, we construct examples of denumerable families of pairwise inconsistent theories which are not representable by normal default theories.

Let  $\mathcal{T} = ext(\Delta)$ , for some default theory  $\Delta$ . Clearly, there are other default theories  $\Delta'$  such that  $ext(\Delta') = \mathcal{T}$ . In other words,  $\Delta$  is not uniquely determined by  $\mathcal{T}$ . Thus, it is natural to search for alternative default theories  $\Delta'$  with the same set of extensions as  $\Delta$ . Let us call  $\Delta'$  equivalent to  $\Delta$  if  $ext(\Delta) = ext(\Delta')$ . We show that for every  $\Delta$  we can effectively (without constructing extensions of  $\Delta$ ) find an equivalent theory  $\Delta'$  with all defaults in  $D'$  prerequisite-free (this result was obtained independently by Schaub [Sch92], and Bonatti and Eiter [BE95]). An important feature of our approach is that it shows that when  $\Delta$  is normal, we can construct a *normal* prerequisite-free default theory  $\Delta'$  equivalent

to  $\Delta$ .

We prove that every normal default theory  $(D, W)$ , such that  $W$  is consistent, is equivalent to a normal default theory with empty objective part. An alternative construction is provided for the case when  $W$  is finite. When applied to a finite, prerequisite-free normal default theory with consistent objective part, it yields an equivalent normal default theory with empty objective part. Moreover, the size of this theory is polynomial in the size of the original theory.

We discuss yet another (weaker) form of equivalence and prove that every normal default theory is equivalent to a theory expressing circumscription (see [McC80, Lif88]).

This paper sheds some light on the issue of expressibility of default logic and, in particular, on expressibility of normal default logic. We firmly believe that the success of default logic as a knowledge representation mechanism depends on a deeper understanding of expressibility issues.

## 2 Preliminaries

In this paper, by  $\mathcal{L}$  we denote a language of propositional logic with a denumerable set of atoms  $At$ . By a *theory* we always mean a subset of  $\mathcal{L}$  *closed under propositional provability*. Let  $B$  be a set of standard monotone inference rules. The formal system obtained by extending propositional calculus with the rules from  $B$  will be denoted by  $PC + B$ . By a *proof* in the system  $PC + B$  we mean any sequence of formulas  $\varphi_1, \dots, \varphi_n$  such that for every  $i = 1, \dots, n$ ,

1.  $\varphi_i$  belongs to  $W$ , is a tautology, or is obtained from formulas  $\varphi_j$  and  $\varphi_k$ , with  $j, k < i$ , by means of modus ponens, or
2. there is a rule  $\frac{\alpha}{\beta}$  in  $B$  such that  $\alpha = \varphi_j$ , for some  $j < i$  and  $\beta = \varphi_i$ .

The corresponding provability operator will be denoted by  $\vdash_B$  and the consequence operator by  $Cn^B(\cdot)$  [MT93].

A *default* is an expression  $d$  of the form  $\frac{\alpha:\Gamma}{\beta}$ , where  $\alpha$  and  $\beta$  are formulas from  $\mathcal{L}$  and  $\Gamma$  is a **finite** list of formulas from  $\mathcal{L}$ . The formula  $\alpha$  is called the *prerequisite*, formulas in  $\Gamma$  — the *justifications*, and  $\beta$  — the *consequent* of  $d$ . The prerequisite, the set of justifications

and the consequent of a default  $d$  are denoted by  $p(d)$ ,  $j(d)$  and  $c(d)$ , respectively. If  $p(d)$  is a tautology,  $d$  is called *prerequisite-free* ( $p(d)$  is then usually omitted from the notation of  $d$ ). This terminology is naturally extended to a set of defaults  $D$ .

By a *default theory* we mean a pair  $\Delta = (D, W)$ , where  $D$  is a set of defaults and  $W$  is a set of formulas. The set  $W$  is called the *objective part* of  $(D, W)$ . A default theory  $\Delta = (D, W)$  is called *finite* if both  $D$  and  $W$  are *finite*.

For a set of defaults  $D$ , define

$$Mon(D) = \left\{ \frac{p(d)}{c(d)} : d \in D \right\}.$$

A default  $d$  (a set of defaults  $D$ ) is *applicable* with respect to a theory  $S$  (is *S-applicable*) if  $S \not\vdash \neg\gamma$  for every  $\gamma \in j(d)$  ( $j(D)$ , respectively). Let  $D$  be a set of defaults. By the *reduct*  $D_S$  of  $D$  with respect to  $S$  we mean the set of monotone inference rules:

$$D_S = Mon(\{d \in D : d \text{ is } S\text{-applicable}\}).$$

When  $E$  is a set of monotonic rules then by  $PC + E$  we mean a logical system extending the rules of proof of propositional logic by rules of  $E$ . The corresponding provability relation is denoted by  $\vdash_E$ . Notice that the proofs in  $PC + E$  are finite. The difference with the usual propositional proofs is that the rules of  $E$  can be used in such proofs.

A theory  $S$  is an *extension*<sup>4</sup> of a default theory  $(D, W)$  if and only if

$$S = Cn^{D_S}(W).$$

The family of all extensions of  $(D, W)$  is denoted by  $ext(D, W)$ .

By a *quasi-proof* of a formula  $\varphi$  from a default theory  $(D, W)$  we mean any sequence of formulas  $\varphi_1, \dots, \varphi_n$  such that  $\varphi = \varphi_n$  and for every  $i = 1, \dots, n$ ,

1.  $\varphi_i$  belongs to  $W$ , is a tautology, or is obtained from formulas  $\varphi_j$  and  $\varphi_k$ , with  $j, k < i$ , by means of modus ponens, or
2. there is a default  $d = \frac{\alpha:\Gamma}{\beta}$  in  $D$  such that  $\alpha = \varphi_j$ , for some  $j < i$  and  $\beta = \varphi_i$

---

<sup>4</sup>Our definition of extension is different from but equivalent to the original definition by Reiter. See [MT93] for details.

For a quasi-proof  $\epsilon$ , by  $D_\epsilon$  we denote a set of defaults used to justify derivations of type (2) in  $\epsilon$ . (such set may not be unique). Clearly, if  $\varphi$  has a quasi-proof  $\epsilon$  from  $(D, W)$ , then  $W \cup c(D_\epsilon) \vdash \varphi$ .

A family  $\mathcal{T}$  of theories in  $\mathcal{L}$  is *non-including* if:

$\mathcal{T}$  is an antichain (that is, for every  $T, T' \in \mathcal{T}$ , if  $T \subseteq T'$  then  $T = T'$ ).

Let  $S$  be a theory. A default  $d$  is *generating* for  $S$  if  $d$  is  $S$ -applicable and  $p(d) \in S$ . The set of all defaults in  $D$  generating for  $S$  is denoted by  $GD(D, S)$ . It is well-known [MT93] that

**(P1)** If  $S$  is an extension of  $(D, W)$  then  $S = Cn(W \cup c(GD(D, S)))$ ,

**(P2)** If all defaults in  $D$  are prerequisite-free then  $S$  is an extension of  $(D, W)$  if and only if  $S = Cn(W \cup c(GD(D, S)))$ .

We will define now the key concepts of the paper.

**Definition 2.1** Default theories  $\Delta$  and  $\Delta'$  are *equivalent* if  $ext(\Delta) = ext(\Delta')$ .

**Definition 2.2** Let  $\Delta$  be a default theory over a language  $\mathcal{L}$  and let  $\Delta'$  be a default theory over a language  $\mathcal{L}'$  such that  $\mathcal{L} \subseteq \mathcal{L}'$ . The theory  $\Delta$  is *semi-equivalent* to  $\Delta'$  if  $ext(\Delta) = \{T \cap \mathcal{L} : T \in ext(\Delta')\}$ .

**Definition 2.3** Let  $\mathcal{T}$  be a family of theories contained in  $\mathcal{L}$ . The family  $\mathcal{T}$  is *representable* by a default theory  $\Delta$  if  $ext(\Delta) = \mathcal{T}$ .

### 3 Representability by general default theories

We start with the result that allows us to replace any default theory with an equivalent default theory in which all defaults are prerequisite-free. As mentioned, this result was known before. However, our argument shows that if we start with a normal default theory, its prerequisite-free equivalent replacement can also be chosen to be normal.

**Theorem 3.1** *For every default theory  $\Delta$  there is a prerequisite-free default theory  $\Delta'$  equivalent to  $\Delta$ . Moreover, if  $\Delta$  is normal then  $\Delta'$  can be chosen to be normal, too.*

Proof: Let  $\Delta = (D, W)$ . For each quasi-proof  $\epsilon$  from  $(D, W)$ , define

$$d_\epsilon = \frac{:j(D_\epsilon)}{\wedge \text{cons}(D_\epsilon)}.$$

Next, define

$$E = \{d_\epsilon : \epsilon \text{ is a quasi-proof from } W\}.$$

Each default in  $E$  is prerequisite-free. Put  $\Delta' = (E, W)$ . We will show that  $\Delta'$  has exactly the same extensions as  $(D, W)$ . To this end, we will show that for every theory  $S$  and for every formula  $\varphi$ ,

$$W \vdash_{D_S} \varphi \text{ iff } W \vdash_{E_S} \varphi. \quad (1)$$

Assume first that  $W \vdash_{D_S} \varphi$ . Then, there is a quasi-proof  $\epsilon$  of  $\varphi$  such that all defaults in  $D_\epsilon$  are applicable with respect to  $S$ . In particular,  $W \cup c(D_\epsilon) \vdash \varphi$ . Observe that  $c(d_\epsilon) \vdash c(D_\epsilon)$ . Since  $d_\epsilon$  is prerequisite-free and  $S$ -applicable,  $W \vdash_{E_S} W \cup c(D_\epsilon)$ . Hence,  $W \vdash_{E_S} \varphi$ .

To prove the converse implication, observe that since all defaults in  $E$  are prerequisite-free,

$$\{\varphi : W \vdash_{E_S} \varphi\} = \text{Cn}(W \cup c(E_S)).$$

Hence, it is enough to show that

$$W \vdash_{D_S} W \cup c(E_S).$$

Clearly, for every  $\varphi \in W$ ,  $W \vdash_{D_S} \varphi$ . Consider then  $\varphi \in c(E_S)$ . It follows that there is a quasi-proof  $\epsilon$  such that  $d_\epsilon$  is  $S$ -applicable and  $c(d_\epsilon) = \varphi$ . Consequently, all defaults occurring in  $\epsilon$  are  $S$ -applicable. Thus, for every default  $d \in D_\epsilon$ ,

$$W \vdash_{D_S} c(d).$$

Since  $\varphi = \wedge c(D_\epsilon)$ ,

$$W \vdash_{D_S} \varphi.$$

To prove the claim for normal default theories, for every quasi-proof  $\epsilon$  define

$$d'_\epsilon = \frac{:\wedge j(D_\epsilon)}{\wedge \text{cons}(D_\epsilon)}$$

(recall that  $D_\epsilon$  denotes the set of defaults used in  $\epsilon$ ). Obviously, since we are dealing with normal defaults now,  $\bigwedge j(D_\epsilon) = \bigwedge cons(D_\epsilon)$ . Next, define

$$E' = \{d'_\epsilon : \epsilon \text{ is a quasi-proof from } W\}.$$

We will show that a theory  $S$  is an extension of  $(E', W)$  if and only if  $S$  is an extension of  $(D, W)$ . Since all defaults in  $(E', W)$  are normal and prerequisite-free, the proof will be complete. The case of inconsistent  $W$  is easy. Hence, in what follows we will assume that  $W$  is consistent.

For every theory  $S$  and every finite set of formulas  $A$ , if  $S \not\vdash \neg \bigwedge A$ , then for every  $\varphi \in A$  we have  $S \not\vdash \neg \varphi$ . It follows that if  $d'_\epsilon$  is applicable with respect to  $S$ , then  $d_\epsilon$  is applicable with respect to  $S$ . Consequently, for every theory  $S$  and every formula  $\varphi$ ,

$$W \vdash_{E'_S} \varphi \text{ implies } W \vdash_{E_S} \varphi.$$

By (1),

$$Cn^{E'_S}(W) \subseteq Cn^{D_S}(W). \quad (2)$$

Assume now that  $S$  is an extension of  $(D, W)$  (since  $W$  is consistent,  $S$  is consistent, too). Then  $S = Cn^{D_S}(W)$ . Hence,  $S \supseteq Cn^{E'_S}(W)$ . Consider  $\varphi \in S$ . Since  $S$  is an extension of  $(D, W)$ ,  $W \vdash_{D_S} \varphi$ . Let  $F$  be the set of defaults whose monotone parts are used in one such proof and let  $\epsilon$  be the corresponding quasi-proof. Clearly, for each  $d \in F$ ,  $W \vdash_{D_S} cons(d)$ . Hence,  $cons(d) \in S$ . Consequently,  $\bigwedge j(F) = \bigwedge cons(F) \in S$ . Since  $S$  is consistent,  $d'_\epsilon$  is  $S$ -applicable and  $W \vdash_{E'_S} \bigwedge cons(F)$ . It follows that  $W \vdash_{E'_S} \varphi$ . Thus,  $S \subseteq Cn^{E'_S}(W)$  and, consequently,  $S = Cn^{E'_S}(W)$ . Hence,  $S$  is an extension of  $(E', W)$ .

Next, assume that  $S$  is an extension of  $(E', W)$ . By (2),  $S \subseteq Cn^{D_S}(W)$ . To prove the converse inclusion, consider  $\varphi \in Cn^{D_S}(W)$ . Then  $W \vdash_{D_S} \varphi$ . Let  $\varphi_1, \dots, \varphi_n = \varphi$  be a proof of  $\varphi$  in  $PC + D_S$ . By induction on  $n$  we will now show that  $\varphi \in S$ . If  $\varphi$  is in  $W$ , is a tautology, or follows from  $\varphi_i$  and  $\varphi_j$ ,  $i, j < n$ , by modus ponens, then the claim is evident.

So, assume that  $\varphi_n$  is the consequent of a default  $d_0 \in D$ . Then  $S \not\vdash \neg j(d_0)$ . Let  $F$  be the set of all defaults whose monotone parts are used in the proof  $\varphi_1, \dots, \varphi_n$ , other than  $d_0$ . Clearly, for every  $d \in F$ ,  $W \vdash_{D_S} cons(d)$ . Hence, by induction,  $cons(d) \in S$ . Consequently,  $S \vdash \bigwedge cons(F) (= \bigwedge j(F))$ . Since  $S \not\vdash \neg j(d_0)$ , it follows that  $S \not\vdash \neg(\bigwedge j(F) \wedge j(d_0))$ . Observe

that the default

$$\frac{: \wedge j(F) \wedge j(d_0)}{\wedge j(F) \wedge j(d_0)}$$

is in  $E'$ . Hence,  $W \vdash_{E'_S} \wedge j(F) \wedge j(d_0)$ . Thus,  $W \vdash_{E'_S} j(d_0)$  ( $= \varphi_n = \varphi$ ) and  $\varphi \in Cn^{E'_S}(W) = S$ . It follows that  $S$  is an extension of  $(D, W)$ .  $\square$

The next result fully characterizes families of theories representable by default theories with a finite set of defaults.

**Theorem 3.2** *The following statements are equivalent:*

- (i)  $\mathcal{T}$  is representable by a default theory  $(D, W)$  with finite  $D$
- (ii)  $\mathcal{T}$  is a finite set of non-including theories, finitely generated over the intersection of  $\mathcal{T}$

Proof: Assume (i). Since every extension of  $(D, W)$  is of the form  $Cn(W \cup c(D'))$ , for some  $D' \subseteq D$ , it follows that  $ext(D, W)$  is finite. It is also well-known ([Rei80, MT93]) that  $ext(D, W)$  is non-including. Let  $U$  be the intersection of all theories in  $ext(D, W)$ . Then  $W \subseteq U$ . Consequently, each extension in  $ext(D, W)$  is of the form  $Cn(U \cup c(D'))$ . Hence, each extension is finitely generated over the intersection of  $ext(D, W)$ .

Now, assume (ii). Let  $U$  be the intersection of all theories in  $\mathcal{T}$ . It follows that there is a positive integer  $k$  and formulas  $\varphi_1, \dots, \varphi_k$  such that  $\mathcal{T} = \{T_1, \dots, T_k\}$  and each  $T_i = Cn(U \cup \{\varphi_i\})$ .

Assume first that  $k = 1$ . Then, it is evident that  $\mathcal{T}$  is the family of extensions of the default theory  $(\emptyset, T_1)$ . Hence, assume that  $k \geq 2$ . Since the theories in  $\mathcal{T}$  are non-including, for every  $j \neq i$  we have

$$U \cup \{\varphi_i\} \not\vdash \varphi_j. \tag{3}$$

In particular, each theory in  $\mathcal{T}$  is consistent and so is  $U$ . Moreover, it follows from (3) that for every  $j = 1, \dots, k$ ,

$$U \not\vdash \varphi_j. \tag{4}$$

Define

$$d_i = \frac{: \neg\varphi_1, \dots, \neg\varphi_{i-1}, \neg\varphi_{i+1}, \dots, \neg\varphi_k}{\varphi_i},$$

$i = 1, \dots, k$ . Next, define  $D = \{d_1, \dots, d_k\}$ . We will show that  $\text{ext}(D, U) = \mathcal{T}$ .

Let  $T$  be an extension of  $(D, U)$ . Then, there is a subset  $\Phi$  of  $\{\varphi_1, \dots, \varphi_k\}$  such that  $T = \text{Cn}(U \cup \Phi)$ . Assume that  $|\Phi| = 0$ . Then, by (4),  $D_T = \{\frac{\cdot}{\varphi_i} : i = 1, \dots, k\}$ . Consequently,  $U = T = \text{Cn}^{D_T}(U) = \text{Cn}(U \cup \{\varphi_1, \dots, \varphi_k\})$ . Hence, for every  $i$ ,  $U \vdash \varphi_i$ , a contradiction with (4). Hence,  $|\Phi| > 0$ . Assume that  $|\Phi| > 1$ . By the definition of  $D$ ,  $D_T = \emptyset$ . Consequently,  $T = \text{Cn}(U \cup \Phi) = \text{Cn}^{D_T}(U) = \text{Cn}(U)$ . Let  $\varphi \in \Phi$  (recall that  $\Phi \neq \emptyset$ ). Then,  $U \vdash \varphi$ , a contradiction. Hence, every extension  $T$  of  $(D, W)$  is of the form  $\text{Cn}(U \cup \{\varphi_i\})$  for some  $i$ ,  $1 \leq i \leq k$ .

To complete the proof, consider an arbitrary  $i$ ,  $1 \leq i \leq k$ . We will show that  $T_i$  is an extension of  $(D, W)$ . First, observe that, by (3),  $D_{T_i} = \{\frac{\cdot}{\varphi_i}\}$ . Consequently,  $\text{Cn}^{D_{T_i}}(U) = \text{Cn}(U \cup \{\varphi_i\}) = T_i$ . Hence,  $T_i$  is an extension of  $(D, U)$ .  $\square$

Let us observe that in the construction of the defaults  $d_i$ , we could replace justifications  $\neg\psi_j$  by any formulas  $\neg\psi_j$  such that  $\psi_j \in T_i \setminus T_j$  and the argument would remain valid. This observation allows us to show that any finite non-including family  $\mathcal{T} = \{T_1, \dots, T_k\}$  of theories is representable. Namely, for every distinct  $i, j$ , let  $\psi_{i,j} \in T_i \setminus T_j$ . Then, define

$$D_i = \left\{ \frac{\neg\varphi_{i,1}, \dots, \neg\varphi_{i,i-1}, \neg\varphi_{i,i+1}, \dots, \neg\varphi_{i,k}}{\beta} : \beta \in T_i \right\}$$

and

$$D = \bigcup_{i=1}^n D_i.$$

Following the argument of Theorem 3.2, it is easy to show that  $\{T_1, \dots, T_k\}$  is exactly the family of extensions of  $(D, \emptyset)$ . Hence, we obtain the following result.

**Proposition 3.3** *Let  $\mathcal{T}$  be a finite non-including family of theories. Then  $\mathcal{T}$  is representable by a default theory (possibly with infinite set of defaults)*

Theorem 3.2 and its argument provide the following corollary, which gives a complete characterization of families of theories representable by finite default theories, that is, theories  $(D, W)$  with both  $D$  and  $W$  finite.

**Corollary 3.4** *The following statements are equivalent:*

1.  $\mathcal{T}$  is representable by a finite default theory

2.  $\mathcal{T}$  is a finite set of finitely generated non-including theories

As pointed out in the introduction, the cardinality argument implies the existence of non-representable families of non-including theories. However, it does not imply the existence of denumerable non-representable families. We will now show two examples of such families. The first family consists of non-including finitely generated theories. The second one consists of mutually inconsistent theories.

**Theorem 3.5** *There exists a countable family of finitely generated non-including theories  $\mathcal{T}$  such that  $\mathcal{T}$  is not representable by a default theory.*

Proof: Let  $\{p_0, p_1, \dots\}$  be a set of propositional atoms. Define  $T_i = Cn(\{p_i\})$ ,  $i = 0, 1, \dots$ , and  $\mathcal{T} = \{T_i : i = 0, 1, \dots\}$ . It is clear that  $\mathcal{T}$  is countable and consists of non-including theories. We will show that  $\mathcal{T}$  is not representable by a default theory.

Assume that  $\mathcal{T}$  is represented by a default theory  $(D, W)$ . By Theorem 3.1, we may assume that all defaults in  $D$  are prerequisite-free. We can also assume that no default in  $D$  contains a justification which is contradictory (such defaults are never used to construct extensions).

Consider a default  $d \in D$ . Since  $j(d)$  is finite, there is  $k$  such that for all  $m > k$ , all formulas in  $j(d)$  are consistent with  $T_m$ . Since  $T_m$  is an extension of  $(D, W)$ ,  $c(d) \in T_m$ , for  $m > k$ . Since

$$\bigcap_{m>k} T_m = Cn(\emptyset),$$

$c(d)$  is a tautology. Since  $d$  was arbitrary, it follows that  $(D, W)$  possesses only one extension, namely  $Cn(W)$ , a contradiction.  $\square$

**Theorem 3.6** *There exists a countable family of mutually inconsistent theories  $\mathcal{T}$  such that  $\mathcal{T}$  is not representable by a default theory. In particular  $\mathcal{T}$  is not representable by a normal default theory.*

Proof: Let  $\{p_0, p_1, \dots\}$  be a set of propositional atoms. Define

$$T_i = Cn(\{\neg p_i, p_{i+1}, \dots\}),$$

for  $i = 0, 1, \dots$ , and  $\mathcal{T} = \{T_i : i = 0, 1, \dots\}$ . It is clear that  $\mathcal{T}$  is countable and consists of pairwise inconsistent theories. Now, we apply precisely the same argument as in the proof of Theorem 3.5.  $\square$

Our counterexamples have an additional property that their infinite subsets and all supersets are also counterexamples.

## 4 Eliminating extensions

In this section, we consider the problem of representability of subfamilies of a representable family. We present a technique for constructing default theories representing some subfamilies of a family of extensions of a given default theory  $\Delta$ . Such techniques are important when we have to redesign the default theory to exclude extensions containing a specific formula and preserve all the remaining extensions unchanged.

Let  $\varphi \in \mathcal{L}$ . Define

$$d_\varphi = \frac{\varphi :}{\perp}.$$

**Theorem 4.1** *Let  $E \subseteq \mathcal{L}$  be consistent and let  $(D, W)$  be a default theory. Then,  $E$  is an extension of  $(D \cup \{d_\varphi\}, W)$  if and only if  $\varphi \notin E$  and  $E$  is an extension of  $(D, W)$ .*

*Proof:* Since  $E$  is consistent,

$$(D \cup \{d_\varphi\})_E = D_E \cup \left\{ \frac{\varphi :}{\perp} \right\}.$$

Assume that  $\varphi \notin E$  and that  $E$  is an extension of  $(D, W)$ . Then

$$E = Cn^{D_E}(W)$$

and  $\varphi \notin Cn^{D_E}(W)$ . Consequently,

$$E = Cn^{D_E}(W) = Cn^{D_E \cup \left\{ \frac{\varphi :}{\perp} \right\}}(W) = Cn^{(D \cup \{d_\varphi\})_E}(W).$$

Hence,  $E$  is an extension of  $(D \cup \{d_\varphi\}, W)$ .

Conversely, assume that  $E$  is an extension of  $(D \cup \{d_\varphi\}, W)$ . Then,

$$E = Cn^{(D \cup \{d_\varphi\})_E}(W) = Cn^{D_E \cup \left\{ \frac{\varphi :}{\perp} \right\}}(W).$$

Since  $E$  is consistent, it follows that  $\varphi \notin Cn^{D_E}(W)$ . Consequently,

$$Cn^{D_E}(W) = Cn^{D_E \cup \{\frac{\varphi}{\perp}\}}(W) = E.$$

Hence,  $\varphi \notin E$  and  $E$  is an extension of  $(D, W)$ .  $\square$

We say that a family  $\mathcal{F}$  of theories (recall that a theory is closed under propositional consequence) has a *strong system of distinct representatives* (SSDR, for short) if for every  $T \in \mathcal{F}$  there is a formula  $\varphi_T \in T$  which does not belong to any other theory in  $\mathcal{F}$ .

**Theorem 4.2** *If  $\mathcal{F}$  is representable by a default theory and has an SSDR, then every family  $\mathcal{G} \subseteq \mathcal{F}$  is representable by a default theory.*

Proof: The claim is obvious if  $\mathcal{F} = \{\mathcal{L}\}$ . So, assume that all members of  $\mathcal{F}$  are consistent (since  $\mathcal{F}$  is an antichain, there are no other possibilities). Let  $(D, W)$  be a default theory such that  $ext(D, W) = \mathcal{F}$ . Define

$$\overline{D} = D' \cup \{d_{\varphi_T} : T \in \mathcal{F} \setminus \mathcal{G}\}.$$

Since all theories in  $\mathcal{F}$  are consistent, the assertion follows from the definition of an SSDR and from the argument of Theorem 4.1.  $\square$

Let us conclude this section with two observations. First, there are families of theories closed under propositional consequence which possess SSDRs but which are not representable by a default theory (the second example presented above possesses this property). Second, in general not every subfamily of a representable family is representable. It follows by the cardinality argument from the fact that there are representable families of cardinality continuum.

## 5 Normal default theories

Our first result in this section completely describes the family of extensions of an arbitrary prerequisite-free normal default theory.

**Theorem 5.1** *Let  $W, \Psi \subseteq \mathcal{L}$ . Let  $D = \{\frac{\varphi}{\varphi} : \varphi \in \Psi\}$ . If  $W$  is inconsistent then  $ext(D, W) = \{\mathcal{L}\}$ . Otherwise,  $ext(D, W)$  is exactly the family of all theories of the form  $Cn(W \cup \Phi)$ , where  $\Phi$  is a maximal subset of  $\Psi$  such that  $W \cup \Phi$  is consistent.*

Proof: The case of inconsistent  $W$  is evident. Hence, let us assume that  $W$  is consistent. Let  $T$  be an extension of  $(D, W)$ . Since  $W$  is consistent,  $T$  is consistent, too. Let  $\Phi = \{\varphi \in \Psi : T \not\vdash \neg\varphi\}$ . Clearly,  $\Phi = c(GD(D, T))$ . By (P2),  $T = Cn(W \cup \Phi)$ . Moreover, since  $T$  is consistent,  $W \cup \Phi$  is consistent. We will show that  $\Phi$  is a maximal subset of  $\Psi$  with this property. Let  $\Phi'$  be such that  $\Phi \subseteq \Phi' \subseteq \Psi$ . Assume that  $W \cup \Phi'$  is consistent. Then,  $T \cup \Phi'$  is consistent. Hence,  $\Phi' \subseteq \Phi$  and, consequently,  $\Phi = \Phi'$ .

Assume next that  $T = Cn(W \cup \Phi)$ , where  $\Phi$  is a maximal subset of  $\Psi$  such that  $W \cup \Phi$  is consistent. Then, it is easy to see that

$$GD(D, T) = \left\{ \frac{\varphi}{\varphi} : \varphi \in \Phi \right\}.$$

Hence,  $\Phi = c(GD(D, T))$  and  $T = Cn(W \cup c(GD(D, T)))$ . Since all defaults in  $D$  are prerequisite-free, it follows by the property (P2) that  $T$  is an extension of  $(D, W)$ .  $\square$

As a corollary, we obtain a full characterization of families of theories that are representable by normal default theories. A similar but stronger result is given as a corollary to Theorem 5.4 at the end of this section.

**Corollary 5.2** *A family  $\mathcal{T}$  of theories in  $\mathcal{L}$  is representable by a normal default theory if and only if  $\mathcal{T} = \{\mathcal{L}\}$  or there is a consistent set of formulas  $W$  and a set of formulas  $\Psi$  such that  $\mathcal{T} = \{Cn(W \cup \Phi) : \Phi \subseteq \Psi \text{ and } \Phi \text{ is maximal so that } W \cup \Phi \text{ is consistent}\}$ .*

Proof: By Theorem 3.1,  $\mathcal{T}$  is representable by a normal default theory if and only if it is representable by a normal default theory with all defaults prerequisite-free. Hence, the assertion follows from Theorem 5.1.  $\square$

We will next study the issue of equivalence between normal default theories. We have already seen that we can replace every normal default theory with an equivalent normal prerequisite-free one (Theorem 3.1). The problem of interest now will be to establish when a normal default theory can be replaced by an equivalent normal default theory with empty objective part.

First, consider a normal default theory  $(D, W)$  such that  $W$  is inconsistent. Then  $ext(D, W) = \{\mathcal{L}\}$ . On the other hand, for every set of normal defaults  $D'$ ,  $ext(D', \emptyset)$

contains only consistent extensions. Hence,  $(D, W)$  is not equivalent to any normal default theory with empty objective part. From now on we will focus on normal default theories  $(D, W)$  for which  $W$  is consistent and we will show that each such theory is equivalent to a default theory with empty objective part.

To prove our results, we need to introduce some notation. Given a formula  $\varphi$ , define:

$$\epsilon\varphi = \begin{cases} \varphi & \text{if } \epsilon = 1 \\ \neg\varphi & \text{if } \epsilon = 0. \end{cases}$$

Now let  $\Psi$  be a set of formulas,  $\Psi = \{\varphi_1, \varphi_2, \dots\}$  (if  $\Psi$  is finite, the last formula of its finite enumeration is repeated infinitely many times) and let  $\Phi \subseteq \Psi$ . Define

$$\epsilon_i^\Phi = \begin{cases} 1 & \text{if } \varphi_i \in \Phi \\ 0 & \text{if } \varphi_i \notin \Phi. \end{cases}$$

Finally, for  $\Phi \subseteq \Psi$  define  $\vec{\Phi}$  as

$$\vec{\Phi} = \{\epsilon_1^\Phi \varphi_1 \wedge \dots \wedge \epsilon_n^\Phi \varphi_n : n = 1, 2, \dots\}$$

Given sets of formulas  $\Psi$  and  $W$ , define a collection  $\mathcal{M}$  as follows:

$$\mathcal{M} = \{\Phi \subseteq \Psi : \Phi \text{ is maximal such that } \Phi \cup W \text{ is consistent}\}$$

We first state and prove a lemma describing maximal consistent subsets of  $\bigcup_{\Phi \in \mathcal{M}} \vec{\Phi} \cup W$ .

**Lemma 5.3** *Maximal consistent subsets of  $\bigcup_{\Phi \in \mathcal{M}} \vec{\Phi} \cup W$  are exactly the sets of the form  $\vec{\Phi} \cup W$ , where  $\Phi$  is a maximal subset of  $\Psi$  such that  $\Phi \cup W$  is consistent (that is,  $\Phi \in \mathcal{M}$ ).*

Proof: We shall make a number of observations which together entail the lemma.

1. Let  $\Phi \in \mathcal{M}$ . Then  $Cn(\Phi \cup W) = Cn(\vec{\Phi} \cup W)$ .

Indeed, if  $\varphi \in (\Psi \setminus \Phi)$  then, by maximality of  $\Phi$ ,  $\Phi \cup W \vdash \neg\varphi$ . This implies that  $\vec{\Phi} \cup W \vdash \neg\varphi$ . The converse property,  $\vec{\Phi} \cup W \vdash \Phi \cup W$ , is obvious, as  $\vec{\Phi} \vdash \Phi$ .

2. If  $\Phi \in \mathcal{M}$  then  $\vec{\Phi} \cup W$  is consistent.

This observation follows from (1) and the fact that  $\Phi \cup W$  is consistent.

3. If  $\Phi \in \mathcal{M}$  then  $\vec{\Phi} \cup W$  is a maximal consistent subset of  $\bigcup_{\Phi \in \mathcal{M}} \vec{\Phi} \cup W$ .

Consider  $\psi = \epsilon_1\varphi_1 \wedge \dots \wedge \epsilon_k\varphi_k \notin \vec{\Phi}$ . Then there is  $i$ ,  $1 \leq i \leq k$ , such that  $\epsilon_i \neq \epsilon_i^\Phi$ . Hence  $\vec{\Phi} \vdash \neg\psi$  or, in other words,  $\vec{\Phi} \cup \{\psi\}$  is inconsistent. Thus  $\vec{\Phi} \cup W \cup \{\psi\}$  is inconsistent.

4. Finally, let  $\Sigma$  be a maximal consistent subset of  $\bigcup_{\Phi \in \mathcal{M}} \vec{\Phi} \cup W$ . Then there is  $\Phi \in \mathcal{M}$  such that  $\Sigma = \vec{\Phi} \cup W$ .

Notice that if two formulas  $\epsilon_1 \varphi_1 \wedge \dots \wedge \epsilon_k \varphi_k$  and  $\epsilon'_1 \varphi_1 \wedge \dots \wedge \epsilon'_{k'} \varphi_{k'}$  belong to  $\Sigma$  then either  $\langle \epsilon_1, \dots, \epsilon_k \rangle$  is a suffix of  $\langle \epsilon'_1, \dots, \epsilon'_{k'} \rangle$  or conversely. Therefore, there is  $\Phi \in \mathcal{M}$  such that

$$\Sigma \cap \bigcup_{\Phi \in \mathcal{M}} \vec{\Phi} \subseteq \vec{\Phi}.$$

But then we have  $\Sigma \subseteq \vec{\Phi} \cup W$ . Since both  $\Sigma$  and  $\vec{\Phi} \cup W$  are maximal consistent subsets of  $\bigcup_{\Phi \in \mathcal{M}} \vec{\Phi} \cup W$ , it follows that  $\Sigma = \vec{\Phi} \cup W$ .  $\square$

Define now

$$D_W = \left\{ \frac{\cdot \psi}{\psi} : \psi \in W \right\}$$

and, for every  $\Phi \in \mathcal{M}$ ,

$$D_\Phi = \left\{ \frac{\cdot \epsilon_1^\Phi \phi_1 \wedge \dots \wedge \epsilon_n^\Phi \varphi_n}{\epsilon_1^\Phi \phi_1 \wedge \dots \wedge \epsilon_n^\Phi \varphi_n} : n = 1, 2, \dots \right\}$$

**Theorem 5.4** *For every normal default theory  $(D, W)$  with  $W$  consistent there exists a prerequisite-free normal default theory  $(D', \emptyset)$  equivalent to  $(D, W)$ .*

Proof: By Theorem 3.1, without loss of generality we can assume that each default in  $D$  is prerequisite-free. Set  $\Psi = c(D)$ . Let

$$D' = D_W \cup \bigcup_{\Phi \in \mathcal{M}} D_\Phi$$

By Theorem 5.1,  $E$  is an extension of  $(D', \emptyset)$  if and only if  $E = Cn(\Sigma)$  for some maximal consistent subset of  $c(D')$ . Since  $c(D') = \bigcup_{\Phi \in \mathcal{M}} \vec{\Phi} \cup W$ , therefore, by Lemma 5.3,  $E$  is an extension of  $(D', \emptyset)$  if and only if  $E = Cn(\vec{\Phi} \cup W)$  for some  $\Phi \in \mathcal{M}$ . Thus by Theorem 5.1 again,  $E$  is an extension of  $(D', \emptyset)$  if and only if  $E$  is an extension of  $(D, W)$ .  $\square$

In the case when  $W$  is finite, we have an alternative construction which can be used to replace a normal default theory  $(D, W)$  with an equivalent normal default theory  $(D', \emptyset)$ .

By Theorem 3.1, without loss of generality we can assume that each default in  $D$  is prerequisite-free. Define  $\omega = \bigwedge W$ .

First, assume that the justification of every default in  $D$  is inconsistent with  $\omega$ . Then,  $ext(D, W) = \{Cn(W)\}$ . Let  $D' = \{\frac{\cdot \omega}{\omega}\}$ . Clearly,  $ext(D', \emptyset) = ext(D, W)$ .

Hence, assume that there are defaults in  $D$  whose justifications are consistent with  $\omega$ . For every default  $d = \frac{i\beta}{\beta}$  in  $D$ , define  $d_\omega = \frac{i\beta \wedge \omega}{\beta \wedge \omega}$ . Finally, define  $D' = \{d_\omega : d \in D\}$ . Using Theorem 5.1, one can show that  $(D', \emptyset)$  is equivalent to  $(D, W)$ .

In the case when  $D$  consists of prerequisite-free defaults only, the size of the theory  $(D', \emptyset)$  is polynomial in the size of  $(D, W)$ . This is not, in general, true for the construction described earlier.

We conclude this section with a corollary to Theorem 5.4 which strengthens the characterization of families of theories representable by normal default theories given in Corollary 5.2.

**Corollary 5.5** *A family  $\mathcal{T}$  of theories in  $\mathcal{L}$  is representable by a normal default theory if and only if  $\mathcal{T} = \{\mathcal{L}\}$  or there is a set of formulas  $\Psi$  such that  $\mathcal{T} = \{Cn(\Phi) : \Phi \subseteq \Psi \text{ is maximal so that } \Phi \text{ is consistent}\}$ .*

## 6 Representability with circumscription

Next, we explore the connections of normal default logic with circumscription [McC80, Lif88] and the Closed World Assumption [Rei78]. Consider a set of atoms  $P$ . Define the set of defaults

$$D^{CWA^P} = \left\{ \frac{i\neg p}{\neg p} : p \in P \right\}.$$

Informally, a default  $\frac{i\neg p}{\neg p}$  allows us to derive  $\neg p$  if  $p$  is not derivable. This has the flavor of the Closed World Assumption. At the same time, such defaults can be viewed as minimizing  $P$ . The exact connection with  $CWA$  and circumscription is given by the following results [MT93]:

1. If  $P = At$  then  $W$  is  $CWA$ -consistent if and only if  $(D^{CWA^P}, W)$  possesses a unique consistent extension ([MT93], Corollary 4.25).
2. If  $P = At$  and  $W$  is consistent, then extensions of  $(D^{CWA^P}, W)$  are complete and consistent theories containing  $W$ , and minimal with respect to  $P$ . Hence, they are in one-to-one correspondence to minimal models of  $W$  ([MT93], Corollary 4.22).

The result of this section shows that all normal default theories can be represented, at a cost of adding new constants, by means of default theories of the form  $(D^{CWA^P}, W)$ . In

particular, it follows that circumscription can be used to express families of extensions of normal default theories.

**Theorem 6.1** *For every normal default theory  $(D, W)$  in  $\mathcal{L}$  there exists a language  $\mathcal{L}' \supseteq \mathcal{L}$ , a set of atoms  $P$  in  $\mathcal{L}'$ , and  $W' \subseteq \mathcal{L}'$  such that  $(D, W)$  is semi-equivalent to a default theory  $(D^{CWA^P}, W')$ .*

Proof: By Theorem 3.1 we can assume that all defaults in  $D$  are prerequisite-free. Let  $\Psi$  be the set of consequents of defaults in  $D$ . For each  $\psi \in \Psi$  select a new atom not belonging to  $At$  (recall that  $At$  is the set of atoms in  $\mathcal{L}$ ). This atom is denoted by  $p_\psi$  and the set  $P$  is defined as  $\{p_\psi : \psi \in \Psi\}$ . Define now  $\mathcal{L}'$  to be the language generated by the set of atoms  $At' = At \cup P$ . Next, define  $V$  as this set of formulas:

$$\{\neg p_\psi \Leftrightarrow \psi : \psi \in \Psi\}.$$

We notice the following fact:

(F1) Let  $\Phi \subseteq \Psi$ . Then  $W \cup \Phi$  is consistent if and only if  $W \cup V \cup \{\neg p_\psi : \psi \in \Phi\}$  is consistent.

Indeed, for a model  $v$  of  $W \cup \Phi$ , define  $v'$  as follows:

$$v'(p) = \begin{cases} v(p) & \text{if } p \in At \\ 1 - v(\psi) & \text{if } p = p_\psi \end{cases}$$

It is clear that  $v'$  is a model of  $W \cup V \cup \{\neg p_\psi : \psi \in \Phi\}$ . Conversely, when a valuation  $v'$  of  $At'$  is a model of  $W \cup V \cup \{\neg p_\psi : \psi \in \Phi\}$  then  $v = v'|_{At}$  is a model of  $W \cup \{\psi : \psi \in \Phi\}$ . Hence, (F1) follows.

Observation (F1) implies that  $\Phi$  is a maximal subset of  $\Psi$  consistent with  $W$  if and only if  $\{\neg p_\psi : \psi \in \Phi\}$  is a maximal subset of  $\{\neg p_\psi : \psi \in \Psi\}$  which is consistent with  $W \cup V$ .

Next, observe that if  $\Phi$  is a maximal set of formulas contained in  $\Psi$  and consistent with  $W$  then for all  $\theta \in \Psi \setminus \Phi$

$$W \cup V \cup \{\neg p_\psi : \psi \in \Phi\} \vdash p_\theta.$$

We are now ready to construct the desired default theory. We put  $W' = W \cup V$  and  $D' = \{\frac{\neg p_\psi}{\neg p_\psi} : \psi \in \Psi\}$ . Clearly,  $D' = CWA^P$ . Using the observations given above, one can now show that the theory  $(D', W')$  is semi-equivalent to  $ext(D, W)$ .  $\square$

Theorems 5.4 and 6.1 can be viewed as “normal form” results for normal default logic. The main difference is the class of default theories used for representation. Theorem 5.4 allows us to replace normal default theories (with consistent  $W$ ) by normal prerequisite-free default theories with empty objective part. Theorem 6.1 shows that every normal default theory can be replaced by a normal default theory with especially simple defaults of the form  $\frac{\neg p}{\neg p}$ .

## 7 Conclusions

The concepts studied in this paper, representability and equivalence, are of key importance for default logic and its applications. Representability provides insights into the expressive power of default logic, while equivalence provides normal form results for default logic, allowing the user to find simpler representations for his/her default theories.

In this paper we characterized those families of theories that can be represented by default theories with a finite set of defaults (Theorem 3.2 and Corollary 3.4). We also presented some sufficient conditions for representability for infinite families of theories. We constructed several countable families that are not representable. However, we have not found a complete characterization of families of theories that are representable by default theories with an infinite set of defaults. This problem seems to be much more difficult and remains open. Topological methods used in [Fer94] may lead to the solution to this problem.

We also studied representability by means of normal default theories. Here, our results are complete. Corollary 5.2 provides a full description of families of theories that are collections of extensions of normal default theories.

Another notion studied in the paper was equivalence of default theories. We showed (Theorem 3.1) that for every normal default theory there exists a normal default theory consisting of prerequisite-free defaults and having exactly the same extensions as the original one. Moreover, we can find an equivalent normal, prerequisite free default theory with empty objective part.

## Acknowledgments

The first and third authors were partially supported by the NSF grant IRI-9400568. The work of the second author was partially supported by the ESPRIT III Basic Research Project 6156 DRUMS II on Defeasible Reasoning and Uncertainty Management Systems. We are grateful to Joeri Engelfriet for his thoughtful comments on an earlier draft of this paper.

## References

- [BE95] P.A. Bonatti and T. Eiter. Querying disjunctive databases through nonmonotonic logics. In *Proceedings of the 5th International Conference on Database Theory — ICDT 95*. Berlin: Springer-Verlag, 1995. Lecture Notes in Computer Science 893.
- [Bes89] P. Besnard. *An introduction to default logic*. Springer-Verlag, Berlin, 1989.
- [Bre91] G. Brewka. *Nonmonotonic reasoning: logical foundations of commonsense*. Cambridge University Press, Cambridge, UK, 1991.
- [CMMT95] P. Cholewiński, W. Marek, A. Mikitiuk, and M. Truszczyński. Experimenting with nonmonotonic reasoning. In *Proceedings of the 12th International Conference on Logic Programming*, Cambridge, MA, 1995. MIT Press.
- [Eth88] D. W. Etherington. *Reasoning with incomplete information*. Pitman, London, 1988.
- [Fer94] A. Ferry. Topological characterizations for logic programming semantics. Doctoral Dissertation, University of Michigan, 1994.
- [Lif88] V. Lifschitz. Circumscriptive theories: a logic-based framework for knowledge representation. *Journal of Philosophical Logic*, 17:391–441, 1988.
- [McC80] J. McCarthy. Circumscription — a form of non-monotonic reasoning. *Artificial Intelligence*, 13:27–39, 1980.
- [MT93] W. Marek and M. Truszczyński. *Nonmonotonic logics; context-dependent reasoning*. Berlin: Springer-Verlag, 1993.
- [Rei78] R. Reiter. On closed world data bases. In H. Gallaire and J. Minker, editors, *Logic and data bases*, pages 55–76. Plenum Press, 1978.
- [Rei80] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.
- [Sch92] T. Schaub. On constrained default theories. In *Proceedings of the 11th European Conference on Artificial Intelligence, ECAI'92*, pages 304–308. Wiley and Sons, 1992.