Default Logic and Specification of Nonmonotonic Reasoning

Joeri Engelfriet[1], V. Wiktor Marek[2], Jan Treur[1], and Mirosław Truszczyński[2]

[1]Free University Amsterdam
Department of Mathematics and Computer Science
Artificial Intelligence Group
De Boelelaan 1081a, 1081 HV Amsterdam
The Netherlands
{joeri, treur}@cs.vu.nl
Phone: +31 20 4447756, +31 20 4447763, Fax: +31 20 4447653.

[2]Department of Computer Science
University of Kentucky
Lexington, KY 40506-0046
USA
{marek, mirek}@cs.engr.uky.edu
Phone: +(606) 257-3961, Fax: +(606) 323-1971

Short Title: Default Logic and Specification of Reasoning

Abstract

In this paper constructions leading to the formation of belief sets by agents are studied. The focus is on the situation when possible belief sets are built incrementally in stages. An infinite sequence of theories that represents such a process is called a *reasoning trace*. A set of reasoning traces describing all possible reasoning scenarios for the agent is called a *reasoning frame*. Default logic by Reiter is not powerful enough to represent reasoning frames. In the paper a generalization of default logic of Reiter is introduced by allowing infinite sets of justifications. This formalism is called *infinitary default logic*. In the main result of the paper it is shown that every reasoning frame can be represented by an infinitary default theory. A similar representability result for antichains of theories (belief frames) is also presented.

Keywords: default logic, extensions, infinitary default logic, representability

Default Logic and Specification of Nonmonotonic Reasoning

1 Introduction

An agent that has to act in a world situation usually has incomplete knowledge about that world. Such knowledge is often not sufficient for the agent to base its actions on, if only classical deductions are used. The agent requires, and adopts, additional assumptions extending its partial understanding of the world. In general, several sets of additional assumptions may be possible or consistent with the agent's knowledge, as there may be alternative ways of interpreting the available (incomplete) information about the world. This leads to several extensions of the agent's initial knowledge. A single one of these extensions (one possible view) will be called a *belief set*, and the set of all of these possible views (given the initial knowledge) will be called a *belief frame*. Belief frames are not arbitrary collections of theories. Since agents are seeking possibly complete descriptions of the world, theories contained in other possible world views are discarded. Hence, belief frames form antichains - no belief set is a proper subset of another in the same reasoning frame.

The belief sets may not be available to the agent immediately. We assume that the agent will have to construct these belief sets by reasoning in a step by step construction process generating a *reasoning trace* that finds its limit in a belief set. A set of these reasoning traces is called a *reasoning frame*. We will require that the limits of all traces in a reasoning frame form a belief frame.

These notions will be formalized as follows. A belief set will be defined as a logical theory (a set of sentences closed under classical deduction). A belief frame will be defined as a collection of theories forming an antichain. A reasoning trace will be defined as a countable increasing (under set inclusion) sequence of theories.

The limit of a reasoning trace is its union. A reasoning frame is defined as a set (or family) of reasoning traces.

Given this conceptualization, two levels of specification of the agent can be described. The most abstract level only defines the outcomes of the reasoning and abstracts from the way the outcome was found. A specification at this level defines a belief frame, abstracting from any reasoning frame behind this belief frame. At the more detailed level of specification a reasoning frame is defined. The set of traces represents the reasoning processes of the agent, their limits — the outcomes.

The question studied in this paper is how a variant of default logic can be used in order to specify nonmonotonic reasoning at these two levels of abstraction. The problem whether a belief frame can be represented as the collection of extensions was studied in (Marek et al., 1997). Complete results in the case of representability by default theories with finite sets of defaults were obtained there. While the general problem of representability remained unresolved, it was shown in (Marek et al., 1997) that the default logic by Reiter is insufficient for specification of belief frames. Specifically, several examples of belief frames were exhibited, which cannot be represented as families of extensions of default theories. In the current paper it is shown that *infinitary default logic*, a stronger variant of default logic, allowing infinite sets of justifications, provides an adequate specification language. In particular, prerequisite-free infinitary default logic provides an adequate specification language for belief frames. Moreover, infinitary default logic in general provides an adequate specification language for reasoning frames.

In Section 2 the basic definitions and properties of infinitary default logic (IDL) are given, as a generalization of Reiter's default logic. For example, the notion of Reiter extension is generalized to the notion of an *IDL-extension*, and a fixpoint construction for IDL-extensions is given, generalizing the fixpoint construction in (Reiter, 1980). We also formally define the notions of a reasoning trace and frame there, and relate these concepts to infinitary default theories.

In Section 3 we focus on the prerequisite-free case. It is proven that the non-including belief frames are precisely the belief frames that can be obtained as the set of all IDL-extensions of a prerequisite-free infinitary default theory. This implies that, in contrast to Reiter's default logic, IDL is expressive enough to serve as an adequate specification language of belief frames.

In Section 4 we focus on reasoning frames. It is established that for any reasoning frame there is an infinitary default theory such that the fixpoint construction of its IDL-extensions precisely provides the reasoning traces in the given reasoning frame. The idea is that by using the right prerequisites any given reasoning trace can be obtained.

In Section 5 we discuss how the notions as presented depend on the initially given set of facts. The notions of a *belief set operator* and a *reasoning trace operator* are introduced to express this dependency. Conclusions and suggestions for further research are given in Section 6.

2 Preliminaries

In this section, we will introduce two key concepts of the paper: reasoning trace and reasoning frame. These concepts are designed to represent the reasoning process of an agent that starts with some incomplete knowledge and, in a step-by-step process constructs a sequence of theories, each providing a more complete picture of the situation (world). We will then introduce the infinitary default logic and show that infinitary default theories can be used to encode reasoning traces and frames of an agent. In the following sections of the paper we will present a detailed study of this relationship.

In this paper, by $\mathcal{L}$ we denote a language of propositional logic with a denumerable set of atoms At. These atoms will be denoted by $p, q, r, ...$ with or without subscripts. By a *theory* we *always* mean a subset of $\mathcal{L}$ *closed under propositional provability*. We will often refer to a theory as a *belief set*. The closure of a set of formulas A under propositional provability is denoted $Cn(A)$. For formulas $\varphi_1, \varphi_2, ..., \varphi_n$ we introduce the notation $\langle \varphi_1, \varphi_2, ..., \varphi_n \rangle$ as an abbreviation of $Cn(\{\varphi_1, \varphi_2, ..., \varphi_n\})$.

When specifying reasoning agents, collections of belief sets that form antichains (no belief set is a proper subset of another) are of particular importance.

Definition 2.1 (Belief Frame) *A* belief frame *is a collection of belief sets (theories) such that no belief set is a proper subset of another.*

Throughout this paper we will use a running example to illustrate the ideas and constructions.

Example 2.2 (Running example) *Define the following theories:*

$$\begin{aligned} T_1 &= \langle p, s, t \rangle \\ T_2 &= \langle p, s, \neg u \rangle \\ T_3 &= \langle p, \neg r, \neg q, t \rangle \\ T_4 &= \langle p, \neg r, \neg q, \neg u \rangle \end{aligned}$$

It is easy to see that $\mathcal{B} = \{T_1, T_2, T_3, T_4\}$ is a belief frame.

As discussed in the introduction, belief frames capture only the outcomes of the reasoning process and abstract from the way these outcomes were found. To get a detailed specification of an agent we need to represent the process in which a belief set is constructed. In this paper, we propose to represent such a process by a sequence of theories — a reasoning trace. Collections of such reasoning traces, in turn, will form reasoning frames. Throughout the paper we will use the following notational convention. If an upper case symbol, say E, stands for a sequence of theories, then the elements of the sequence will be referred to as $E_1, E_2, \ldots$, and their union, $\bigcup_{i=1}^{\infty} E_i$ will be denoted by E^∞.

Definition 2.3 (Reasoning Trace and Reasoning Frame) *Let* $T = (T_1, T_2, \ldots)$ *be a sequence of theories from* $\mathcal{L}$.

(i) The sequence T *is a* reasoning trace *if:*

(a) $T_i \subseteq T_{i+1}$ *for* $i = 0, 1, \ldots$

(b) $T_i = T_{i+1}$ *implies* $T_i = T_j$ *for* $j > i$.

(ii) The union of a reasoning trace T *is called the* limit *of* T.

(iii) A collection $\mathcal{T}$ *of reasoning traces is called a* reasoning frame *if for every* $T, S \in \mathcal{T}$:

(a) $T_0 = S_0$.

(b) If $T^\infty \subseteq S^\infty$, *then* $T = S$.

It is easy to see that the limit of a reasoning trace is a theory, that is, it is closed under propositional provability, and that the limits of reasoning traces in a reasoning frame form a belief frame, that is, form an antichain.

Definition 2.4 (Belief Frame of a Reasoning Frame) *Let* $\mathcal{T}$ *be a reasoning frame. The belief frame* $\mathcal{B}_\mathcal{T}$ *associated with* $\mathcal{T}$ *is defined by:*

$$\mathcal{B}_\mathcal{T} = \{T^\infty : T \in \mathcal{T}\}$$

Example 2.5 *The following is an example of a reasoning frame:*

$$\begin{aligned}
\mathcal{F} \quad = \quad & \{(\langle p\rangle, \langle p, s\rangle, \langle p, s, t\rangle, \langle p, s, t\rangle, ...), \\
& (\langle p\rangle, \langle p, s\rangle, \langle p, s, \neg u\rangle, \langle p, s, \neg u\rangle, ...), \\
& (\langle p\rangle, \langle p, \neg r, \neg q\rangle, \langle p, \neg r, \neg q, t\rangle, \langle p, \neg r, \neg q, t\rangle, ...), \\
& (\langle p\rangle, \langle p, \neg r, \neg q\rangle, \langle p, \neg r, \neg q, \neg u\rangle, \langle p, \neg r, \neg q, \neg u\rangle, ...)\}
\end{aligned}$$

It is easy to show that this is indeed a reasoning frame. The reader can check that $\mathcal{B}_\mathcal{F} = \mathcal{B}$ *where* $\mathcal{B}$ *was defined in Example 2.2.*

In this paper we will show that the language of infinitary default logic can be used to describe specifications of an agent both on the level of belief frames as well as reasoning frames. Some results in this direction were obtained in (Marek et al., 1997), where the problem of encoding belief frames by (finitary) default theories was studied in detail. In addition to a number of positive results, it is proved in (Marek et al., 1997) that not every belief frame can be represented as the family of all extensions of a default theory. In this paper we will generalize default logic by allowing infinite sets of justifications. Then we will prove that infinitary default logic is powerful enough to serve as a specification language for arbitrary belief and reasoning frames.

An *infinitary default* (*IDL-default*, for short) is an expression d:

$$d = \frac{\alpha : \Gamma}{\beta}, \tag{1}$$

where α and β are formulas from $\mathcal{L}$, and Γ is a set, possibly infinite, of formulas from $\mathcal{L}$. The formula α is called the *prerequisite* of d ($p(d)$, in symbols) and β is called the *consequent* of d ($c(d)$, in symbols). The set of formulas Γ is called the *justification* set of d and is denoted by $j(d)$. If $p(d)$ is a tautology, d is called *prerequisite-free*. In such case, $p(d)$ is usually omitted from the notation of d. This terminology is naturally extended to a set of defaults D. Namely, the *prerequisite*, *consequent* and *justification* sets of D, in symbols $p(D)$, $c(D)$ and $j(D)$, are defined by:

$$p(D) = \bigcup_{d \in D} \{p(d)\}, \quad c(D) = \bigcup_{d \in D} \{c(d)\}, \quad j(D) = \bigcup_{d \in D} j(d).$$

A pair (D, W), where D is a set of IDL-defaults and $W \subseteq \mathcal{L}$ is a set of formulas, is called an *infinitary default theory* (or *IDT*). Rules with infinite sets of justifications were considered in (Ferry, 1991) in the context of logic programs.

We will now generalize the notion of an extension, introduced by Reiter (Reiter, 1980) for standard default theories, to the case of IDTs. To this end, we will introduce the concept of an S-trace. This notion is closely related to the fixpoint construction of extensions presented by Reiter (Reiter, 1980).

Definition 2.6 *Let* (D, W) *be an IDT. Let* $S \subseteq \mathcal{L}$ *be a theory. By the* S-trace *of* (D, W) *we mean the sequence* E *of theories defined recursively as follows:*

1. $E_0 = Cn(W)$,

2. *for every integer* $n \geq 0$:

$$E_{n+1} = Cn(E_n \cup \{c(d)\colon\ d \in D,\quad p(d) \in E_n \text{ and for all } \gamma \in j(d),\ \neg\gamma \notin S\}).$$

The notion of an S-trace allows us to introduce the notion of an IDL-extension of an IDT.

Definition 2.7 *Let* (D, W) *be an IDT. A set* $S \subseteq \mathcal{L}$ *is an* IDL-extension *of* (D, W) *if*

$$S = E^{\infty},$$

where E *is the* S*-trace for* (D, W).

Clearly, each standard (finitary) default theory (with each default having only finitely many justifications) is, in particular, an IDT. Moreover, it is easy to see that if an IDT happens to be finitary, then the notion of an IDL-extension coincides with that of extension. Therefore, throughout the paper we will refer to the IDL-extensions simply as extensions.

We will denote by $ext(D, W)$ the collections of all extensions of an IDT (D, W). The collection of all S-traces of (D, W), where $S \in ext(D, W)$ will be denoted by $tr(D, W)$.

There are several alternative characterizations of extensions of standard default theories (Marek and Truszczyński, 1993). We will now generalize one of them to the case of infinitary default theories. It can be stated in terms of the reduct of the set of defaults. 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). By the *reduct* D_S of D with respect to S we mean the set of monotone inference rules:

$$D_S = \left\{ \frac{\alpha}{\beta}\colon \text{ for some } \Gamma \subseteq \mathcal{L},\ \frac{\alpha\colon\Gamma}{\beta} \in D, \text{ and } \frac{\alpha\colon\Gamma}{\beta} \text{ is } S\text{-applicable} \right\}.$$

Each set B of standard monotone inference rules determines a formal proof system, denoted by $PC + B$, in which derivations are built by means of propositional provability and rules in B. The corresponding provability operator will be denoted by $\vdash_B$ and the consequence operator by $Cn^B(\cdot)$ (Marek and Truszczyński, 1993). In particular, each set D_S determines the provability operator $\vdash_{D_S}$ and the consequence operator $Cn^{D_S}(\cdot)$.

Proposition 2.8 *Let D be a set of IDL-defaults,and let W and S be subsets of $\mathcal{L}$. Then, S is an extension if and only if $S = Cn^{D_S}(W)$.*

Let us introduce one more useful notion. A default d is *generating* for a theory S if $p(d) \in S$ and $S \not\vdash \neg\gamma$ for every $\gamma \in j(d)$. The set of all defaults from D which are generating for S is denoted by $GD(D, S)$.

Once the reduct is computed the distinction between infinitary and standard defaults disappears. This explains why many of the properties of default logic remain true in the infinitary case. In particular, we have the following results.

Proposition 2.9 *Let (D, W) be an IDT. Then:*

1. *If S is an extension of (D, W), then S is a belief set (theory).*

2. *The operator $Cn^{D_S}(W)$ is monotone in D and W, and antimonotone in S.*

3. *The collection ext(D, W) is a belief frame. That is, if $T_1, T_2 \in$ ext(D, W) and $T_1 \subseteq T_2$, then $T_1 = T_2$.*

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

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

Parts (1) and (3) of Proposition 2.9 show that IDTs can be used to represent belief frames. The next result shows that they can also be used to represent reasoning frames.

Proposition 2.10 *Let (D, W) be an IDT.*

1. *Let S be a theory in $\mathcal{L}$. If E is the S-trace for (D, W) then E is a reasoning trace.*

2. *The collection of reasoning traces $tr(D, W)$ is a reasoning frame.*

We can now formally introduce the notions of representability of belief frames and reasoning frames by default theories.

Definition 2.11 *Let $\mathcal{B}$ be a family of belief sets contained in $\mathcal{L}$. The family $\mathcal{B}$ is* representable *by an IDT Δ if $ext(\Delta) = \mathcal{B}$. Similarly, if $\mathcal{T}$ is a family of reasoning traces, then it is* representable *by an IDT Δ if $tr(\Delta) = \mathcal{T}$.*

Example 2.12 *It turns out that the belief frame $\mathcal{B}$ of Example 2.2 is representable. Define the IDT (D, W) by*

$$\begin{aligned} W &= \{p\} \text{ and} \\ D &= \{\frac{p : q \vee r}{s}, \frac{p : \neg s}{\neg r}, \frac{: \neg s}{\neg q}, \frac{s \vee \neg r : u}{t}, \frac{s \vee \neg q : \neg t}{\neg u}\} \end{aligned}$$

It can be easily verified that $ext(D, W) = \mathcal{B}$. Furthermore, it can also be checked that $tr(D, W) = \mathcal{F}$, where $\mathcal{F}$ was defined in Example 2.5. This means that $\mathcal{F}$ is also representable.

The notion of representability by default theories was studied in (Marek and Truszczyński, 1993; Marek et al., 1997). A complete description of families of theories that are representable by default theories with a finite set of defaults was given there. However, the general question of representability by arbitrary default theories remained unsettled. The main difference between a standard and an infinitary default is that the latter can encode an infinite set of constraints determining its applicability (in the form of infinite sets of justifications). Our results in the next section show that the infinitary default logic is more expressive than the default logic by Reiter. In particular, we show that every family of theories satisfying the necessary condition for the representability, described in Proposition 2.9(3), is representable by an infinitary default theory.

3 Representability of Belief Frames by IDTs

We start with the result that allows us to replace any IDT with an equivalent IDT in which all defaults are prerequisite-free.

Theorem 3.1 *For every IDT Δ, there is a prerequisite-free IDT Δ' equivalent to Δ.*

Proof: Let $\Delta = (D, W)$. By a *quasi-proof* from D and W we mean any proof from W in the system $PC + D^m$, where

$$D^m = \left\{ \frac{\alpha}{\beta}\colon \text{ for some } \Gamma \subseteq \mathcal{L},\ \frac{\alpha\colon\Gamma}{\beta} \in D \right\}.$$

For every quasi-proof ϵ from D and W, let D_ϵ be the set of all defaults used in ϵ. For each such proof ϵ, define

$$d_\epsilon = \frac{\colon j(D_\epsilon)}{\bigwedge c(D_\epsilon)}$$

(observe that D_ϵ is finite and, so, d_ϵ is well-defined). Next, define

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

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

$$W \vdash_{D_S} \varphi \ \text{ iff } \ W \vdash_{Q_S} \varphi.$$

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

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

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

Hence, it is enough to show that

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

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

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

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

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

□

Example 3.2 *Let us look at the IDT (D, W) defined in Example 2.12. Every default d_ϵ defined in the proof of Theorem 3.1 is uniquely determined by the set D_ϵ of defaults used in ϵ. This means that for (D, W), which consists of five defaults, at most $2^5 = 32$ defaults d_ϵ can be defined. It turns out that the IDT defined in this way actually contains 24 defaults (the other subsets of the defaults in D can not be combined into a proof). Rather than listing all 24, we will give a number of them. First of all, the defaults with prerequisite in W are proofs, so for instance $\frac{:q \vee r}{s}$ and $\frac{:\neg s}{\neg q}$ are in Q. The second, third and fifth defaults in D give rise to the following default: $\frac{:\{\neg s, \neg t\}}{\neg r \wedge \neg q \wedge \neg u}$. But there are also defaults that contradict their own* *justification, such as:* $\frac{:\{q \vee r, u, \neg t\}}{s \wedge t \wedge \neg u}$. *These defaults are present in* Q*, but they are harmless given the other defaults in* Q.

Proposition 2.9 implies that for every infinitary default theory (D, W), its family of extensions $ext(D, W)$ is a belief frame (cf. parts (1) and (3)). To answer the question whether the converse is true as well, by Theorem 3.1 we can concentrate on prerequisite-free IDT's. It turns out that every belief frame is representable by a (prerequisite-free) IDT.

Theorem 3.3 *Let* $\mathcal{B}$ *be a family of belief sets. Then the following statements are equivalent:*

(i) $\mathcal{B}$ *is a belief frame,*

(ii) $\mathcal{B}$ *is representable by a prerequisite-free IDT.*

Proof: It suffices to prove that any belief frame is representable by a prerequisite-free IDT. To this end, let us consider a belief frame $\mathcal{B}$. If $\mathcal{B} = \emptyset$ then take any (Reiter) default theory without extensions. If $\mathcal{B} = \{T\}$, then define $D = \emptyset$. Clearly, $ext(D, T) = \mathcal{B}$.

Hence, assume that $\mathcal{B}$ contains at least two theories. Since no theory in $\mathcal{B}$ is a proper subtheory of another, it follows that all theories contained in $\mathcal{B}$ are consistent.

For every $S, T \in \mathcal{B}$ such that $S \neq T$, define $\varphi_{S,T}$ to be any formula belonging to $S \setminus T$. For every $T \in \mathcal{B}$, define

$$D^T = \left\{ \frac{:\{\neg\varphi_{S,T} : S \in \mathcal{B},\ S \neq T\}}{\psi} : \ \psi \in T \right\}.$$

Finally, define

$$D = \bigcup_{T \in \mathcal{B}} D^T.$$

We will show that $ext(D, \emptyset) = \mathcal{B}$.

Consider $T \in \mathcal{B}$. Then $D_T = \{\frac{:}{\psi}: \psi \in T\}$. Hence, $Cn^{D_T}(\emptyset) = T$ and T is an extension of $(D, \emptyset)$.

Conversely, let T be an extension of $(D, \emptyset)$. We have just proved that $\mathcal{B} \subseteq$ *ext*$(D, \emptyset)$. Consequently, $(D, \emptyset)$ has at least two extensions. It follows that $Cn(\emptyset)$ is not an extension of $(D, \emptyset)$ (the theory $Cn(\emptyset)$ is a subset of every extension of (D, W)). In particular, $T \neq Cn(\emptyset)$. Consequently, the set D_T is not empty.

Consider a set $S \in \mathcal{B}$. Observe that all defaults in D^S have the same set of justifications. Consequently, either all of them are generating for T or none. It follows that T is the union of a nonempty (since $D_T \neq \emptyset$) family of theories in $\mathcal{B}$. If T is the union of at least two theories, than $D_T = \emptyset$, a contradiction. Hence, $T = S$, for some $S \in \mathcal{B}$. That is, $T \in \mathcal{B}$. $\square$

Example 3.4 *We already know that our example belief frame $\mathcal{B}$ is representable, and we know it is representable by a prerequisite-free IDT. In order to illustrate the construction process of the proof of Theorem 3.3, we will perform this for $\mathcal{B}$. Note, first of all, that in the definition of D^T, we need not add defaults for every $\varphi \in T$, but that it is sufficient to do this for a set of generators of T (T is generated by a set of formulas if it is the propositional closure of this set). Furthermore, when a formula $\varphi_{S,T}$ is a negation, we will eliminate the double negation in the default. We will now construct the sets D^T:*

1. *T_1: first we must choose the formulas φ_{S,T_1}. Take $\varphi_{T_2,T_1} = \neg u$, $\varphi_{T_3,T_1} = \neg r$, and $\varphi_{T_4,T_1} = \neg r$, then $D^{T_1} = \{\frac{:\{u,r\}}{p}, \frac{:\{u,r\}}{s}, \frac{:\{u,r\}}{t}\}$.*

   Note that these defaults have the same set of justifications, so instead of taking 3 defaults, we can also form one by taking the conjunction of the consequents. We will do this for the remaining theories.

2. *T_2: Let $\varphi_{T_1,T_2} = t$, $\varphi_{T_3,T_2} = t$, $\varphi_{T_4,T_2} = \neg r$, and define $D^{T_2} = \{\frac{:\{\neg t,r\}}{p \wedge s \wedge \neg u}\}$.*

3. T_3: *Let* $\varphi_{T_1,T_3} = s$, $\varphi_{T_2,T_3} = s$, $\varphi_{T_4,T_3} = \neg u$, *and define* $D^{T_3} = \{\frac{:\{\neg s, u\}}{p \wedge \neg r \wedge \neg q \wedge t}\}$.

4. T_4: *Let* $\varphi_{T_1,T_4} = s$, $\varphi_{T_2,T_4} = s$, $\varphi_{T_3,T_4} = t$, *and define* $D^{T_4} = \{\frac{:\{\neg s, \neg t\}}{p \wedge \neg r \wedge \neg q \wedge \neg u}\}$.

If we define $D = D^{T_1} \cup D^{T_2} \cup D^{T_3} \cup D^{T_4}$, *then it can be checked that indeed* $ext(D, \emptyset) = \mathcal{B}$.

Theorem 3.3 and the results in (Marek et al., 1997) imply that infinitary default logic is a more powerful knowledge representation formalism than that of default logic. In other words, allowing infinite justification sets leads to a more expressive representation formalism.

Corollary 3.5 *There are belief frames representable by an IDT but not representable by a standard default theory.*

We will give an example. Let $\{p_0, p_1, \ldots\}$ be a set of propositional atoms. Define $T_i = Cn(\{p_i\})$, $i = 0, 1, \ldots$, and $\mathcal{B} = \{T_i : i = 0, 1, \ldots\}$. It is clear that $\mathcal{B}$ consists of non-including theories, and is therefore representable by an IDT. If we define $W = \emptyset$ and $D = \{\frac{:\{\neg p_j | j \neq i\}}{p_i} \mid i \geq 0\}$, then it can be easily verified that $ext(D, W) = \mathcal{B}$. It was shown, however, in (Marek et al., 1997) (Theorem 3.5), that $\mathcal{B}$ is not representable by a (Reiter) default theory.

As another corollary, we obtain the result already proved in (Marek et al., 1997).

Proposition 3.6 *Let* $\mathcal{B}$ *be a finite belief frame. Then* $\mathcal{B}$ *is representable by a (Reiter) default theory (possibly with an infinite set of defaults).*

This is actually a special case of a more general criterion for representability by (Reiter) default theories. Let us call a family of theories $\mathcal{B}$ *finitely distinguishable* if for all $T \in \mathcal{B}$ there exists a *finite* set FD(T) such that $FD(T) \cap T = \emptyset$ and $\forall S \in \mathcal{B} : S \neq T \Rightarrow FD(T) \cap S \neq \emptyset$. We have the following result.

Proposition 3.7 *Every finitely distinguishable belief frame is representable by a Reiter default theory.*

Proof: In the proof of Theorem 3.3, we can always choose the formulas $\varphi_{S,T}$ from $FD(T)$, a finite set. Then the sets D^T contain only defaults with finite justification sets, so that the IDT defined in the proof is in fact a Reiter default theory. $\square$

It is easy to see that a finite belief frame is finitely distinguishable. Hence, Proposition 3.7 applies to finite belief frames.

4 Representability of Reasoning Frames by IDTs

In the previous section we proved that any antichain of theories (belief frame) can be represented by a prerequisite-free IDL-theory. In this section we will look not only at the outcomes of a reasoning process (the belief frame), but also at the process in which these outcomes are constructed. Note that by using prerequisites that logically depend on consequents of other defaults, it is possible to express constraints on the order in which states occur in a trace. Using this observation, we will study the question whether infinitary default logic can be used as a specification language for collections of traces — reasoning frames. In the main result of this section we will show that every reasoning frame is representable by an IDT.

Theorem 4.1 *Let $\mathcal{T}$ be a collection of reasoning traces. Then the following statements are equivalent:*

(i) $\mathcal{T}$ is a reasoning frame,

(ii) $\mathcal{T}$ is representable by an IDT.

Proof: If there is an IDT Δ such that $\mathcal{T} = tr(\Delta)$, then $\mathcal{T}$ is a reasoning frame by Proposition 2.10. To prove the converse implication, we proceed as follows. If $\mathcal{T}$ is empty, we can take Δ to be any default theory without extensions. So suppose that $\mathcal{T} \neq \emptyset$. Take any trace $T \in \mathcal{T}$, and define $W = T_0$. As $\mathcal{T}$ is a reasoning frame, we have that $W = S_0$ for all traces $S \in \mathcal{T}$.

Consider a trace $T \in \mathcal{T}$. Then T is increasing, and may become constant from a certain index on. We define this index k_T by

$$k_T = \begin{cases} min\{i : T_i = T_{i+1}\} & \text{if there exists an } i \text{ with } T_i = T_{i+1} \\ \infty & \text{otherwise} \end{cases}$$

Now for $0 < i < k_T$, define $\psi_{i,T}$ to be any formula in $T_i \setminus T_{i-1}$, and define $\psi_{0,T}$ as any formula in T_0. These formulae will serve as prerequisites for defaults that will "fire" in order to form T_{i+1}.

For the justifications of rules, we will use the same construction as used in the proof of Theorem 3.3. For any $S \in \mathcal{T}$ such that $S \neq T$, define $\varphi_{S,T}$ to be any formula belonging to $S^\infty \setminus T^\infty$. Since $\mathcal{T}$ is a reasoning frame and $S \neq T$, $S^\infty \not\subseteq T^\infty$. Hence, $\varphi_{S,T}$ can indeed be found. Now define

$$D^T = \left\{ \frac{\psi_{i,T} : \{\neg\varphi_{S,T} : S \in \mathcal{T},\ S \neq T\}}{\chi} : \chi \in T_{i+1} \setminus T_i,\ 0 \leq i < k_T \right\}.$$

Finally, define

$$D = \bigcup_{T \in \mathcal{T}} D^T.$$

We will show that $tr(D, W) = \mathcal{T}$.

Consider $T \in \mathcal{T}$. First observe that, by definition, $T_0 = W = Cn(W)$. Furthermore, the set of defaults in D which are applicable for T^∞ is exactly D^T. It follows that

$$\{c(d) : d \in D, p(d) \in T_i, d \text{ is } T^\infty\text{–applicable}\} = \{\chi : \chi \in T_{i+1} \setminus T_0\}.$$

As $T_0 \subseteq T_i$, we have that

$$T_{i+1} = Cn(T_i \cup \{c(d) : d \in D, p(d) \in T_i, d \text{ is } T^\infty\text{–applicable}\}).$$

From this we conclude that $T \in tr(D, W)$.

For the converse, suppose that $T \in tr(D, W)$. If none of the defaults in D are T^∞–applicable, then $T_i = W$ for all i. Consider an $S \in \mathcal{T}$. Then, we have $S \in tr(D, W)$. Now, since $S^\infty \supseteq W$ and extensions form an antichain, $S^\infty = W$. Hence, $S = T$ and $T \in \mathcal{T}$.

So suppose there is a T^∞–applicable default in D. Then there exists a trace $S \in \mathcal{T}$ such that all defaults in D^S are T^∞–applicable. We will show by induction that $S_i \subseteq T_i$. Indeed, if $i = 0$, then $S_0 = W = T_0$. For the induction step, observe that

$$T_{i+1} = Cn(T_i \cup \{c(d) \colon d \in D, p(d) \in T_i, d \text{ is } T^\infty\text{–applicable}\}) \supseteq$$

$$Cn(S_i \cup \{c(d) \colon d \in D, p(d) \in S_i, d \in D^S\})$$

Since $S \in tr(D, W)$ and D^S is exactly the set of defaults of D which are S^∞–applicable, the last term is equal to S_{i+1}.

Now we have that $S^\infty \subseteq T^\infty$. Moreover, since both S^∞ and T^∞ are extensions of (D, W), it follows that $S^\infty = T^\infty$. But then a default is S^∞–applicable if and only if it is T^∞–applicable, so that $S_i = T_i$ for all i, or $S = T$. We conclude that $T \in \mathcal{T}$. □

Example 4.2 *We will give a reasoning frame such that its associated belief frame is $\mathcal{B}$ (different than the one we gave earlier), and construct an IDT that represents it. The reasoning frame consists of the following traces:*

$$\begin{aligned}
T^1 &: (\langle p\rangle, \langle p, s \vee t\rangle, \langle p, s, t\rangle ...) \\
T^2 &: (\langle p\rangle, \langle p, s, \neg u\rangle ...) \\
T^3 &: (\langle p\rangle, \langle p, t \to \neg r \wedge \neg q\rangle, \langle p, t, \neg r, \neg q\rangle ...) \\
T^4 &: (\langle p\rangle, \langle p, \neg r\rangle, \langle p, \neg r, \neg q\rangle, \langle p, \neg r, \neg q, \neg u\rangle ....)
\end{aligned}$$

It can again easily be seen that this is a reasoning frame. We define $W = Cn(\{p\})$. We will take the $\varphi_{S,T}$ the same as in Example 3.4. Then we have to define the formulas $\psi_{i,T}$. We will not enumerate all of these explicitly, but just give an example. Consider ψ_{1,T^4}. This has to be a formula in $(T^4)_1 \setminus (T^4)_0 = \langle p, \neg r\rangle \setminus \langle p\rangle$, so we *could take* $\psi_{1,T^4} = \neg r$. *We now give the sets of defaults:*

$$
\begin{array}{lcl}
D^{T^1} & = & \{\dfrac{p\colon \{u,r\}}{s \vee t}, \dfrac{s \vee t\colon \{u,r\}}{s \wedge t}\} \\
D^{T^2} & = & \{\dfrac{p\colon \{\neg t,r\}}{s \wedge \neg u}\} \\
D^{T^3} & = & \{\dfrac{p\colon \{\neg s,u\}}{t \rightarrow \neg r \wedge \neg q}, \dfrac{t \rightarrow \neg r \wedge \neg q\colon \{\neg s,u\}}{t}\} \\
D^{T^4} & = & \{\dfrac{p\colon \{\neg s,\neg t\}}{\neg r}, \dfrac{\neg r\colon \{\neg s,\neg t\}}{\neg q}, \dfrac{\neg q\colon \{\neg s,\neg t\}}{\neg u}\}
\end{array}
$$

The reader can check that, setting $D = D^{T^1} \cup D^{T^2} \cup D^{T^3} \cup D^{T^4}$, *indeed* $tr(D,W) = \{T^1, T^2, T^3, T^4\}$.

As was the case in the construction of an IDL-theory in the previous section, we again have considerable freedom in choosing the formulae $\varphi_{S,T}$. A second source of freedom comes from the choice of the prerequisites in the above construction. Thus, in general there are many different theories which all specify the same reasoning frame.

From the construction in the proof it is clear that, analogous to the case of belief frames, reasoning frames with a *finite* number of reasoning traces can be represented by a (Reiter) default theory.

One could ask if finitary representability of the belief frame of a reasoning frame implies that the reasoning frame itself has a finitary representation. A cardinality argument shows that this is not the case. Specifically, let us consider the belief frame $\mathcal{B}$ consisting of all complete theories over the set of atoms $\{p_1, p_2, \ldots\}$. This belief frame has a finitary representation (see (Marek et al., 1997), Corollary 5.5). It is easy to see that there are more than continuum reasoning frames with belief frame $\mathcal{B}$. On the other hand, there is only continuum many finitary default theories.

Until now, we have looked at the specification of belief sets and reasoning frames which represent the reasoning process of an agent from a given set of initial facts. In the next section we will take a broader perspective and look at the different belief sets and reasoning frames an agent may have when varying the set of initial facts.

5 Varying the Initial Facts

In the preceding sections we have seen that infinitary default logic can be used for the specification of belief frames and reasoning frames. These two notions are a formalization (at two levels of abstraction) of the reasoning process of an agent from a *fixed* initial situation. This initial situation is described by (part of) the intersection of theories in a belief frame, or the common first point in the traces of a reasoning frame. Thus, a belief frame or reasoning frame gives no information about the reasoning process of the same agent from different sets of initial facts. In order to take into account these different initial situations, we consider belief set operators and reasoning trace operators. (See (Engelfriet et al., 1995; Engelfriet et al., 1996). See also (Makinson, 1994).)

Definition 5.1

1. *A* belief set operator *is a function which assigns to each $X \subseteq \mathcal{L}$ a collection of belief sets that include X.*

2. *A* reasoning trace operator *is a function which assigns to each $X \subseteq \mathcal{L}$ a collection of reasoning traces that start at X.*

We would like to specify these operators using (infinitary) default logic, and an obvious way of doing this is using families of sets of defaults, indexed by sets of formulae.

Definition 5.2

1. *Let $\mathcal{B}$ be a belief set operator. The operator $\mathcal{B}$ is* representable *by an indexed family of sets of defaults $(D_X)_{X \subseteq \mathcal{L}}$ if for all $X \subseteq \mathcal{L}$: $\mathcal{B}(X) = ext(D_X, X)$.*

2. *Let $\mathcal{F}$ be a reasoning trace operator. The operator $\mathcal{F}$ is* representable *by an indexed family of sets of defaults $(D_X)_{X \subseteq \mathcal{L}}$ if for all $X \subseteq \mathcal{L}$: $\mathcal{F}(X) = tr(D_X, X)$.*

Given the results in the previous sections, the following is easy to see:

Proposition 5.3

1. *A belief set operator $\mathcal{B}$ is representable by an indexed family of sets of (prerequisite-free) defaults iff $\mathcal{B}(X)$ is a belief frame for all $X \subseteq \mathcal{L}$.*

2. *A reasoning trace operator $\mathcal{F}$ is representable by an indexed family of sets of defaults iff $\mathcal{F}(X)$ is a reasoning frame for all $X \subseteq \mathcal{L}$.*

In principle, this is a valid way of specifying belief set operators and reasoning trace operators. However, it is intuitively not very likely that an agent should use a (completely) different set of defaults in every situation. Instead, it seems more plausible that the agent has *one* set of defaults which (s)he uses regardless of the initial facts (meaning that $D_X = D_Y$ for all X, Y). This leads to a different representability question: given a belief set operator $\mathcal{B}$, does there exist a set of (prerequisite-free) defaults D, such that for all $X \subseteq \mathcal{L}$ we have $\mathcal{B}(X) = ext(D, X)$ (and similarly for reasoning trace operators). It seems that this is a non-trivial question; we will leave this for future research.

6 Conclusions

In (Marek et al., 1997) the usefulness of Reiter's Default Logic for specifying multiple belief sets of an agent was investigated. It was established that every finite non-including family of belief sets is representable by a default theory. However, examples of denumerably infinite non-including families were constructed that are not representable by a default theory. In the current paper these results are extended in two directions. Firstly, a new variant of default logic was introduced, Infinitary Default Logic, that allows to represent every non-including family of belief sets, independent of its cardinality.

Secondly, not only the representability of families of belief sets as an outcome of default reasoning processes was investigated, but also the representability of default reasoning traces constructing these belief sets. Here a positive answer was also obtained for infinitary default logic, whereas Reiter's Default Logic fails for the non-finite case.

Thus specification of default reasoning is made possible at two levels of abstraction. For specification at the level of families of belief sets that are the outcomes of default reasoning processes (abstracting from the reasoning traces constructing them), prerequisite-free infinitary default theories are adequate means. Using them no commitment is made to any particular traces to construct the belief sets. For specification at the level of reasoning traces general infinitary default theories are adequate means. They specify both the families of belief sets that are the outcomes and the traces constructing them.

It is interesting to note that from a representation viewpoint, the only role played by the prerequisites lies in guiding the construction process. Of course, even when specifying only belief sets, it may be the case that an IDL-theory *with* prerequisites exists which is more concise than a prerequisite-free theory. However, this would also give a specification at a lower level of abstraction since it not only specifies the outcomes of the reasoning but also the way outcomes are generated. One can then choose to commit to this particular specification of the traces, but one could also consider the specification as meant only to specify the outcomes and give a different specification for the traces. One way of changing the specification for the traces is by introducing so-called lemma default rules are (see e.g. (Schaub, 1992)). This causes conclusions to be added earlier in a trace.

A specification language for belief set operators and reasoning trace operators based on temporal logic was introduced in (Engelfriet and Treur, 1996). An embedding of default logic into this specification language was given there.

Further issues for research include representability of belief set operators and reasoning trace operators using default logic (as mentioned in Section 5) and the general question of representability using finitary default logic (with infinite sets of defaults).

References

Engelfriet, J., Herre, H., and Treur, J. (1995). Nonmonotonic belief state frames and reasoning frames (extended abstract). In Froidevaux, C. and Kohlas, J., editors, *Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Proceedings ECSQARU'95*, volume 946 of *Lecture Notes in Artificial Intelligence*, pages 189–196. Springer-Verlag.

Engelfriet, J., Herre, H., and Treur, J. (1996). Nonmonotonic reasoning with multiple belief sets. In Gabbay, D. and Ohlbach, H., editors, *Practical Reasoning, International Conference on Formal and Applied Practical Reasoning, FAPR'96*, volume 1085 of *Lecture Notes in Artificial Intelligence*, pages 331–344. Springer-Verlag.

Engelfriet, J. and Treur, J. (1996). Specification of nonmonotonic reasoning. In Gabbay, D. and Ohlbach, H., editors, *Practical Reasoning, International Conference on Formal and Applied Practical Reasoning, FAPR'96*, volume 1085 of *Lecture Notes in Artificial Intelligence*, pages 111–125. Springer-Verlag.

Ferry, A. (1991). Enriched nonmonotonic rule system. Master's thesis, University of Kentucky.

Makinson, D. (1994). General patterns in nonmonotonic reasoning. In Gabbay, D., Hogger, C., and Robinson, J., editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3, pages 35–110. Clarendon Press, Oxford.

Marek, V., Treur, J., and Truszczyński, M. (1997). Representation theory for default logic. *Annals of Mathematics and Artificial Intelligence*, 21:343–358.

Marek, V. and Truszczyński, M. (1993). *Nonmonotonic Logic: Context-dependent Reasoning*. Springer-Verlag.

Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence*, 13(1-2):81–132.

Schaub, T. (1992). On constrained default theories. In Neumann, B., editor, *Proceedings of the 11th European Conference on Artificial Intelligence*, pages 304–308. John Wiley & Sons.