An Argumentation-theoretic Approach to Reasoning with Specificity

Phan Minh Dung
Division of Computer Science,
Asian Institute of Technology
PO Box 2754, Bangkok 10501, Thailand
Email: dung@cs.ait.ac.th

Tran Cao Son
Department of Computer Science,
University of Texas at El Paso
El Paso, Texas 79968, USA
Email: tson@cs.utep.edu

Abstract

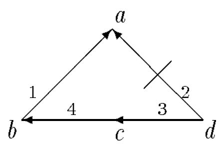
We present a new argumentation-theoretic approach to default reasoning with specificity. The new approach differs from other approaches in the way priority between defaults is handled. Here, it is context sensitive rather than context independent as in other approaches. We start by showing that any context independent handling of priorities between defaults as advocated in the literature until now is not sufficient to capture general defeasible inheritance reasoning. This motivates the introduction of an argumentation framework for default reasoning with specificity where the context sensitive priorities between defaults are captured by the attacks-relation between the arguments. We present several new and novel results. First we show that our framework subsumes the semantics of defeasible inheritance networks. We then show that the new semantics satisfies core properties of default reasoning such as the conditioning, deduction, reduction, and cumulative propositions [1]. To give a proof procedure to our approach, we present a modular transformation from default theories into logic programs where the preferred semantics of the former coincides with the answer set semantics of the later.

1 Introduction

Default reasoning has emerged as a response to the need of a mechanism that can accommodate the nonmonotonic character of commonsense reasoning [17, 18, 20, 29]. In default reasoning, default rules are used to draw new conclusions in the absence of conflicting arguments. One of the most difficult problems in default reasoning is the problem of reasoning with specificity. Though many proposals have been forwarded in the literature [1, 3, 11, 12, 26, 30] to address it, no satisfactory solution is obtained yet.

Formally a default theory T could be defined as a pair (E, K) where E is a set of evidence or facts representing what we call the concrete context of T, $K = (D, B)$ is the domain knowledge consisting of a set of default rules D, and a first order theory B representing the background knowledge. In the literature [2, 5, 11, 12] the principle of reasoning with specificity is "enforced" by first determining a set of priority orders between defaults in D using the information given by the domain knowledge K. Based on these priorities between defaults and following some sensible and intuitive criteria [1, 3, 11, 12], the semantics of T is then defined either model-theoretically by selecting a subset of the set of all models of $E \cup B$ as the set of preferred models of T or proof-theoretically by selecting certain extensions as the preferred extension [5, 30]. The problem of these approaches is that their obtained semantics is rather weak. We will show in section 2 that these approaches do not capture general defeasible inheritance reasoning. The reason is that the priorities between defaults are defined independent of the context E.

It seems that there is a consensus in the literature [5, 11, 12] recognizing that for any default reasoning system to account for general inheritance reasoning it is necessary to combine the three important ideas of minimality, conditioning and anticontrapositive. The following example illustrates this point.



Example 1.1 Let consider the default theory $T = (E, K)$ with $B = \emptyset$, $D = \{d/c, c/b, d/\neg a, b/a\}$, and

[1]To the best of our knowledge, it is the first time a general framework for default reasoning has been given which subsumes inheritance reasoning when applied to nonmonotonic inheritance networks and at the same time also satisfies the core properties of default reasoning

$E = \{d\}$ represented in the above picture. The idea of minimality restricts the set of candidates for preferred models to those in which only a minimal set of default rules are violated (or equivalently a maximal set of default rules are satisfied). In this case, the "minimal" models are:

$M_1 = \{d, c, b, \neg a\}$ with $D_1 = \{2, 3, 4\}$ being the set of rules satisfied in M_1.

$M_2 = \{d, c, b, a\}, D_2 = \{1, 3, 4\}$

$M_3 = \{d, \neg c, \neg b, \neg a\}, D_3 = \{1, 2, 4\}$

$M_4 = \{d, c, \neg b, \neg a\}, D_4 = \{1, 2, 3\}$

where in the last two models, some rules are "applied" contrapositively. The "anti-contrapositive"-idea forbids contrapositive applications of default rules. Hence, the set of candidates for preferred models in the above example would reduce to $\{M_1, M_2\}$. Finally, the idea of conditioning stating that given an evidence p in a default theory (E, K) containing a default rule p/q, q should be concluded, will pick M_1 as the preferred model.

The idea of minimality is the basic idea of circumscription [18]. To account for the other two, prioritized circumscription has been introduced with some success (see, e.g. [22]). Reiter's default logic [29] can be viewed as a first attempt to combine proof theoretically the idea of minimality and anticontrapositive. Seminormal defaults and other more recent proposals [5, 11, 12, 30] have been proposed to incorporate the idea of conditioning into default logic. Geffner and Pearl [11] showed that combining the ideas of minimality and conditioning can provide a system which can account for reasoning with specificity and a weak form of inheritance reasoning.

Argumentation has been recognized lately as an important and natural approach to nonmonotonic reasoning [4, 6, 7, 11, 21, 25, 28, 30, 37]. Dung [6], Bondarenko et all. [4] showed that most of the major nonmonotonic logics [17, 18, 20, 23, 29] are special cases of a general simple argumentation system. Geffner and Pearl [11] have used argumentation to give a proof procedure for their conditional logic. In this paper, we use argumentation to combine the three basic ideas of default reasoning where a sort of context dependent priorities between defaults is (implicitly) employed. We then demonstrate that our new framework can capture general defeasible inheritance reasoning by showing that the credulous semantics of defeasible inheritance networks [13, 14, 15, 19, 32, 33, 34] are captured in our framework. We also show that our semantics satisfies core properties of default reasoning such as the conditioning, deduction, reduction, and cumulative propositions.

To give proof procedures to our semantics we transform default theories into logic programs such that the answer set semantics of the later coincides with the argumentational semantics of the former. So, well-known proof procedures for logic programs can be used as proof-procedure for default reasoning.

The paper is organized as follows. In section 2 we will show that any default reasoning system based on a context independent priorities between defaults can not account in full for general inheritance reasoning. In section 3 we recall the argumentation framework of Dung [6]. Section 4 presents our new approach, where we show that our framework when applied to inheritance reasoning subsumes the credulous semantics of defeasible inheritance networks. We also show that the new semantics enjoy many important properties of an entailment relation discussed in [11, 16] such as conditioning, deductive, reduction, and cumulative proposition. Afterwards we present the transformation of default theories into extended logic programs where the preferred semantics of the former coincides with the answer set semantics of the later. We conclude in section 6.

2 What is Wrong with Context Independent Priorities

Let $T = (E, K)$ be a default theory where E (a first order theory) is called the context of T representing the evidence or facts, $K = (B, D)$ is called the domain knowledge of T with B (a first order theory) representing the background knowledge and D being a set of defaults. In this section, we will give an example showing that priority-based approaches to default reasoning, in which priorities between defaults are independent of the context E, do not capture general default reasoning.

In the literature [1, 3, 5, 11, 12, 24, 30], the semantics of T is defined based on certain partial orders on D which is determined solely by K. Let PO_K be the set of all these partial orders. For each partial order $\alpha \in PO_K$, where $(d, d') \in \alpha$ means that d is of lower priority than d', a partial order $<_\alpha$ between the sets of defaults in D is defined where $S <_\alpha S'$ means that S is preferred to S'. There are many ways to define $<_\alpha$ [1, 3, 5, 11, 12]. But whatever the definition of $<_\alpha$ is, $<_\alpha$ has to satisfy the following property.

Let S be a subset of D and d, d' be two defaults in D such that $(d, d') \in \alpha$. Then $S \cup d' <_\alpha S \cup d$.

$<_\alpha$ can be extended into an partial order between models of $B \cup E$ as follows:

$M <_\alpha M'$ iff $D_M <_\alpha D_{M'}$

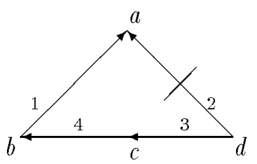
where D_M is the set of all defaults in D which are satisfiable in M and a default p/q is said to be satisfiable in M iff the material implication $p \rightarrow q$ is satisfiable in M.

A model M of $B \cup E$ is defined as a preferred model of T if there exists a partial order α in PO_K such that M is minimal with respect to $<_\alpha$. We then say that a formula β is defeasibly derived from T if β holds in each preferred model of T.

Now we want to show that any preferential semantics based on $<_\alpha$ can not account in full for general inheritance reasoning.

Let us consider again the example:

Example 2.1 Let B be empty, $E = \{d\}$, and D be:



An interpretation for the above network can be given by *d=student, c=young adult, b=adult, a=married.* The desirable semantics here is represented by the model $M = \{d, c, b, \neg a\}$. To have this semantics, all priority-based approaches in the literature [1, 3, 5, 12] assigns the default 1 a lower priority than the default 2. Note that though conditional entailment [11] does not provide the desirable semantics in this case, it could be "fixed" by restricting the set of partial orders PO_K on those in which 1 has a lower priority than 2. Further, it is important to note that the priorities between defaults are determined independent of the context E.

Let consider T under a new context $E' = \{d, \neg c, b\}$. Now, since c does not hold, the default $d/\neg a$ can not be considered more specific than the default b/a. Hence, it is intuitive to expect that neither a nor $\neg a$ should be concluded in this case. This is also the result sanctioned by all semantics of defeasible inheritance networks [13, 14, 15, 19, 32]. In any priority-based system employing the same priorities between defaults with respect to E' as with respect to E, we have $M = \{\neg a, d, \neg c, b\} <_\alpha M' = \{a, d, \neg c, b\}$ since $D_M = \{4, 2\} <_\alpha D'_M = \{4, 1\}$ (due to $(1, 2) \in \alpha$). That means priority-based approaches in the literature would conclude $\neg a$ given (E', K) which is certainly not the intuitive result we expect. This leads to the idea that 1 should have lower priority than 2 only if the context $\{d, c, b\}$ holds. In other words, the priority order under the context $\{d, \neg c, b\}$ is different than the priority order under the context $\{d, c, b\}$. It turns out that this context-sensitive priority can be naturally captured in an argumentation theoretic approach.

3 Preliminaries: Argumentation Framework

We first recall the basics of the abstract theory of argumentation of Dung [6].

Definition 3.1 *An* argumentation framework *is a pair $AF = \langle AR, attacks \rangle$, where AR is a set of arguments, and $attacks \subseteq AR \times AR$.*

If $(A, B) \in attacks$ we say A attacks B or B is attacked by A. Further a set of arguments S is said to attack an argument A if some argument in S attacks A.

Definition 3.2 *A set of arguments S is said to be* conflict-free *if there exist no arguments A,B in S such that $(A, B) \in attacks$.*

The stable semantics of AF is defined as follows.

Definition 3.3 *A set of arguments S is called a* stable extension *of AF if S is conflict-free and S attacks every argument which does not belong to S.*

It is easy to see that S is stable iff S={A | A is not attacked by S}.

The stable semantics of argumentation framework captures the semantics of many other mainstream approaches to nonmonotonic reasoning such as extension of Reiter's Default Logic [29], stable expansion of Autoepicstemic Logic [17], and stable model of Logic Programming [9]. Also, the credulous semantics of an inheritance network coincides with the stable semantics of the corresponding argumentation framework [7].

Often, a more skeptical semantics is advocated in many approaches to nonmonotonic reasoning [25, 36]. This form of skeptical semantics is captured in the argumentation framework by the notion of grounded extension defined as the least fixpoint of the following operator.

$$F_{AF} : 2^{AR} \to 2^{AR}$$

$F_{AF}(S) = \{A \mid$ A is defended by S $\}$ where A is *defended by S* iff for every argument B, if B attacks A, then S attacks B.

The *grounded semantics* of an argumentation framework AF is defined as the least fixpoint of F_{AF}. It has been pointed out in [6] that both the semantics of Pollock's Inductive Defeasible Logic [25], and the well-founded semantics of Logic Programming [36] are captured by the grounded semantics of argumentation.

The preferred semantics of argumentation is defined as follows

Definition 3.4 *A set of arguments S is called a* preferred extension *of AF iff S is a conflict-free maximal fixpoint of F_{AF}.*

In general, stable extensions are preferred extensions but not vice versa. But as we will see later, in the argumentation frameworks corresponding to default theories, stable semantics and preferred semantics coincide. So it is enough for us to work with either of them.

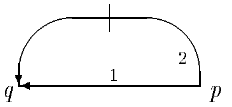
4 A New Approach

In this section we present a new approach to handle specificity in a class of default theories which comprises of a set of ground literals (evidences) E and a finite set of default rules $D = \{d_1, \ldots, d_n\}$ where each rule $d \in D$ is of the form $l_1, \ldots, l_m/h$ and $l_i's$, h are literals over an arbitrary but fixed first order language $\mathcal{L}$ such that $\{h\} \cup \{l_1, \ldots, l_m\}$ is consistent. h is called the *head* of d, denoted by $hd(d)$, and the set $\{l_1, \ldots, l_n\}$ is called the *body* of d, denoted by $bd(d)$. Two defaults d and d' are said to be *in conflict* if $\{hd(d), hd(d')\}$ is inconsistent. Throughout this paper, we will assume that the set of background knowledge in T is empty, i.e. $B = \emptyset$. The restriction that B is empty is taken to allow us to present the main ideas of our new approach in an intuitive and simple way without bothering to much with technical details. We also assume that the language $\mathcal{L}$ is finite, i.e., the sets of constant, function, and predicate symbols of $\mathcal{L}$ are finite. The set of all ground instances of the defaults in D is denoted by *ground(D)*.

Definition 4.1 1. *A default theory $T = (E, D)$ is* inconsistent *if E is inconsistent or there exists two defaults d and d' in D such that $bd(d) \equiv bd(d')$ and $\{hd(d), hd(d')\}$ is inconsistent.*

2. *A default theory is* consistent *if it is not inconsistent.*

An example of an inconsistent default theory is illustrated in the following picture.



We now define the *consequence* relation.

Definition 4.2 *Given an evidence set E and a default set D we say that a ground literal l is a* consequence of E wrt. D, *denoted by $E \vdash_D l$, iff $l \in E$ or there is a sequence of ground defaults $d_1, \ldots, d_n$ $(1 \leq n)$ in* ground(D) *such that*

1. *$l = hd(d_n)$ and*
2. *$bd(d_1) \subseteq E$, and*
3. *$bd(d_{i+1}) \subseteq E \cup \{hd(d_1), \ldots, hd(d_i)\}$.*

For a set of literals L we write $E \vdash_D L$ iff $\forall l \in L: E \vdash_D l$.

We write $E \vdash_D false$ iff there is a ground atom a such that both $E \vdash_D a$ and $E \vdash_D \neg a$ hold.

We next introduce the notion of *atom dependency graph* of a default theory $T = (E, D)$ before we define the class of acyclic default theories.

Definition 4.3 *Let $T = (E, D)$ be a default theory. The* atom dependency graph *of T is defined as a directed graph G_T where the nodes of G_T are ground atoms in the language $\mathcal{L}$ of T and there is a directed link from b to c in G_T iff there is a default $b_1, ..., b_n/h$ in ground(D) such that b is the atom of some b_i and c is the atom of h.*

Definition 4.4 *A default theory $T = (E, D)$ is* acyclic *if there is no cycle in the atom dependency graph of T.*

In following, we will restrict our study only on acyclic and consistent default theories. We believe that the class of acyclic and consistent default theories is large enough to accommodate the most of practical default reasoning. At least, most of the examples found in the literature belong to this class.

From now on, $T = (E, D)$ denotes an arbitrary but fixed acyclic and consistent default theory if not otherwise specified.

Definition 4.5 *A set of ground defaults $A \subseteq ground(D)$ is called an* argument *in T if for all $d \in A$, $E \vdash_A bd(d)$, and $E \not\vdash_A false$.*

By definition, in the example 2.1 $A_1 = \{2, 3, 4\}$, $A_2 = \{1, 4, 3\}$, $A_3 = \{2\}$ are arguments while $B = \{1, 2\}$ is not an argument since $\{d\} \not\vdash_B bd(1)$.

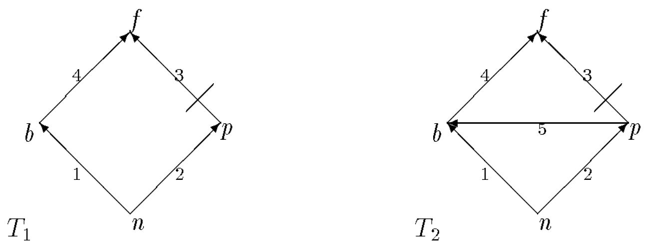
For any default theory T, the set of all arguments in T is denoted by AR_T.

Let A be an argument and w be a ground literal, we say A *supports* w iff $E \vdash_A w$. Further $Con(A)$ denotes the set of all conclusions supported by A, i.e., $Con(A) = \{w | E \vdash_A w\}$. It is easy to see that if A and B are two arguments with consistent $Con(A) \cup Con(B)$ then $A \cup B$ is an argument.

To define the semantics of T, we now determine the attack- relationship between the arguments in AR_T. First, it is easy to see that A *attacks* B if A and B support conflicting conclusions, i.e., $Con(A) \cup Con(B)$ is inconsistent.

We now consider another kind of attack.

Example 4.1 (Motivation of Attack by Specificity)

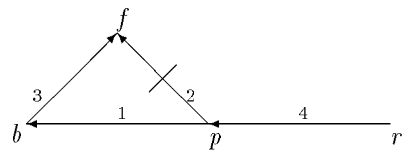


Consider the default theory T_1 in the above figure with $E = \{n\}$. Let $A_1 = \{2, 3\}$ and $A_2 = \{1, 4\}$. From the above discussion, it is clear that A_1 attacks A_2 and vice versa. Now adding the default p/b to T_1 we will have the famous Penguin-Bird-Fly example in which A_1 becomes more specific than A_2. Thus, due to the specificity principle which stipulates that specific information overrides more general but conflicting one we can say that adding p/b to T_1 creates an attack of new kind against $A_2 = \{1, 4\}$. Hence it is natural to view the argument $B = \{2, 5\}$ in the presence of the default 3 as an attack against the argument A_2. □

The above example shows that there are two types of attacks among the arguments in AR_T: attacks by conflict and attack by specificity. The definition of the first type of attacks is clear. It remains to define the second type of attacks. It is not difficult to see that in the above example, the attack by specificity is determined by the context consisting of three defaults 3,4,5 in which default 5 makes default 4 less spec ific than default 3. Such contexts are called specificity-relevant minimal conflict sets. A precise definition is given in definition 4.10. But first, we define the notion of *minimal conflict sets.*

Definition 4.6 *Let d be a default in $ground(D)$. A set of ground defaults $C \subseteq ground(D)$ is called a* minimal conflict set (abbr. **MCS**) wrt. *d, iff*

- *$d \in C$, and*
- *C is a minimal set such that $bd(d) \vdash_C false$, and*
- *there exists no $d' \in C$ such that $bd(d) \not\equiv bd(d')$ and $bd(d') \vdash_C false$*



Example 4.2 In the above picture, $C = \{1, 2, 3\}$ is a MCS wrt. 2 or 1 but $C' = \{1, 2, 3, 4\}$ is not a MCS wrt. 4 since the third condition is violated. □

The following lemma follows directly from definition of MCS.

Lemma 4.1 *Let C be an arbitrary MCS wrt. d.*

- *For all $d' \in C$ there exists $C' \subseteq C$, such that $bd(d) \vdash_{C'} bd(d')$.*
- *C contains exactly two defaults d' and d'' such that $\{hd(d'), hd(d'')\}$ is inconsistent.*

Definition 4.7 *Let C be a MCS. A default $d' \in C$ is called a* conflicted default *in C if $\exists d'' \in C$ such that $\{hd(d'), hd(d'')\}$ is inconsistent.*

To resolve the conflict caused by a MCS we will block one of its defaults from applying and to enforce the principle of specificity the least specific default should be blocked as suggested in [5]. For example, in T_2 of example 4.1 we should block 4 rather than 3.

Definition 4.8 *Let C be a MCS and $d', d'' \in C$. d'' is said to be* more specific *than d' (with respect to C) iff there exists $C' \subseteq C$ such that $bd(d'') \vdash_{C'} bd(d')$.*

Obviously, the more specific relation is a pre-order with respect to a MCS, i.e., in a MCS it is transitive and reflexive.

Definition 4.9 *Let C be a MCS and $d' \in C$. d' is said to be* most specific *in C iff for all defaults $d'' \in C$, d' is more specific than d''.*

Let C be a MCS with respect to d in an acyclic default theory. Then, it is easy to see that $d' \in C$ is most specific iff $bd(d) \equiv bd(d')$.

In example 2.1, the defaults 1 and 2 are the two conflicted defaults of the MCS of which 2 is a most specific default. In T_1 of example 4.1 we have two conflicted defaults 3 and 4 but none of them is most specific. We distinguish these two types of MCS's in the following definition.

Definition 4.10 *A MCS C is said to be* specificity relevant *iff one of the two conflicted defaults of C is most specific in C.*

By definition, the MCS of example 2.1 is specificity relevant while the MCS in T_1 of example 4.1 is not. The following easy lemma helps to characterize the specificity relevant MCS's.

Lemma 4.2 *Let C be a MCS with respect to d. Then C is specificity relevant if d is a conflicted default.*

For convenience, from now on, whenever we say that C_d is a specificity relevant MCS, we understand that C_d is a specificity relevant MCS wrt d and d is one of the conflicted defaults of C_d.

We have specified a class of MCS's which are relevant for specificity reasoning. We now give a formal definition for the attack relationship.

Definition 4.11 *Let A, B be arguments in AR_T. We say that A attacks B if one of the following conditions holds:*

1. attack by conflict: *A and B supports conflicting conclusions, i.e., $Con(A) \cup Con(B)$ is inconsistent.*
2. attacks by specificity: *there exists a specificity relevant MCS C_d whose conflicted defaults are d, d' such that*

- $E \vdash_A bd(d)$*, and* $C_d \setminus \{d', d\} \subseteq A$*, and*
- $d' \in B$.

$AF_T = (AR_T, attacks)$ *denotes the argumentation framework corresponding to* T.

In example 4.1 the default theory T_1 has following arguments: $A_0 = \emptyset, A_1 = \{2,3\}, A_2 = \{1,4\}, A_3 = \{2\}, A_4 = \{1,2,3\}, A_5 = \{1\}, A_6 = \{1,2,4\}, A_7 = \{1,2\}$, and $attacks_{T_1} = \{(A_2, A_1), (A_1, A_2), (A_2, A_4), (A_4, A_2), (A_6, A_1), (A_1, A_6), (A_6, A_4), (A_4, A_6)\}$. We then have $\{A_0, A_3, A_5, A_1, A_4, A_7\}$ and $\{A_0, A_3, A_5, A_2, A_6, A_7\}$ are the two preferred extensions which are also stable, of the argumentation framework $\langle AR_{T_1}, attacks_{T_1}\rangle$ which is the expected result. Introducing p/b into T_1 creates some more arguments $A_8 = \{2,5\}$, $A_9 = \{2,5,4\}$, $A_{10} = \{1,2,5\}$, $A_{11} = \{1,2,4,5\}$, $A_{12} = \{1,2,5,3\}$, and $A_{13} = \{2,3,5\}$ and for all arguments of AF_{T_2} we have i) if $\{2,5\} \subseteq A, 4 \in B$ then A attacks B and ii) if $3 \in A, 4 \in B$ then A attacks B and B attacks A. It is easy to see that AF_{T_2} has only one preferred extension $\{A_0, A_1, A_3, A_4, A_5, A_7, A_8, A_{10}, A_{12}, A_{13}\}$ which is also stable. □

As stated in section 3, the coincidence of preferred semantics and stable semantics in AF_T is proved in the following theorem.

Theorem 1 *Let* T *be an acyclic and consistent default theory. Then, each preferred extension of* AF_T *is also a stable extension of* AF_T *and vice versa.*

Remark 4.1 *From now on, we often refer to a preferred extension of* AF_T *simply as a preferred extension of* T.

It has been shown in [6] that preferred extension always exists for argumentation systems. So, we have

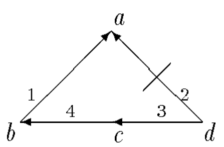
Lemma 4.3 *Each acyclic and consistent default theory has at least a preferred extension.*

Definition 4.12 *We write* $T \mid\!\sim w$ *for a ground literal* w *iff* w *is supported by at least one argument in each preferred extension of* T.

How Context-Sensitive Priorities Are Handled

The next example demonstrates that our approach deals with default priorities in a context sensitive way and hence provides the expected intuitive results in many cases where other approaches fail.

Example 4.3 (Continuation of example 2.1)



Consider $T = (E, D)$ with $D = \{1,2,3,4\}$. We consider two sets of evidence:

1. $E = \{d\}$. Then, $AR_T = \{A_0, A_1, A_2, A_3, A_4, A_5, A_6\}$ where $A_0 = \emptyset$, $A_1 = \{2\}$, $A_2 = \{3\}$, $A_3 = \{3,4\}$, $A_4 = \{1,3,4\}$, $A_5 = \{2,3,4\}$, and $A_6 = \{2,3\}$. $attacks_T = \{(A_1, A_4), (A_4, A_1), (A_4, A_5), (A_5, A_4), (A_3, A_4), (A_6, A_4), (A_4, A_6)\}$. We can say that 2 has higher priority than 1 in the context $\{d\}$ by the fact that A_4, the only argument using 1, is attacked by A_3 while A_3 does not attack the arguments containing 2. It is easy to see that AF_T has **only one preferred extension** $\{A_0, A_1, A_2, A_3, A_5, A_6\}$ which supports b, c, and $\neg a$ as in defeasible inheritance network. □
2. $E = \{d, \neg c, b\}$. Then, AF_T contains only three arguments: $A_0 = \emptyset$, $A_1 = \{2\}$, $B = \{1\}$. $attacks_T = \{(A_1, B), (B, A_1)\}$. From the symmetry in the attacks-relation between arguments A_1, B we can say that there are no priorities between defaults 1 and 2. AF_T has **two preferred extensions** $\{A_0, A_1\}$, and $\{A_0, B\}$ and so $T \not\mid\!\sim \neg a$ and $T \not\mid\!\sim a$. □

Core Properties of $\mid\!\sim$

We now prove some general propositions of $\mid\!\sim$. The next theorem shows that conditioning is satisfied in preferred semantics.

Theorem 2 *(Conditioning) Given an acyclic and consistent default theory* $T = (E, D)$ *with* $E = bd(d)$ *for an arbitrary ground default* $d \in ground(D)$*. Then,* $T \mid\!\sim hd(d)$.

Proof See appendix 2.

We will next discuss other general propositions [1, 5, 11, 16, 17] of the entailment relation $\mid\!\sim$. The deduction proposition is trivial because $\emptyset$ is an argument which supports the evidence E and $\emptyset$ belongs to every preferred extension of AF_T as there is no argument which attacks $\emptyset$. The reduction and cumulative propositions of $\mid\!\sim$ are proved in the next theorems.

Theorem 3 *(Reduction) Given T=(E,D) such that* $T \mid\!\sim a$ *and* $T' \mid\!\sim b$ *for* $T' = (E \cup \{a\}, D)$*. Then,* $T \mid\!\sim b$.

Proof See appendix 3. □

Theorem 4 *(Cumulativity) Let* $T = (E, D)$ *be a default theory such that* $T \mid\!\sim a$ *and* $T \mid\!\sim b$*. Then,* $T' \mid\!\sim b$ *for* $T' = (E \cup \{a\}, D)$.

Proof Follows directly from the lemmas 6.9 and 6.10 in appendix 4. □

Coincidence between Preferred Extensions and Credulous Extensions of Defeasible Inheritance Network

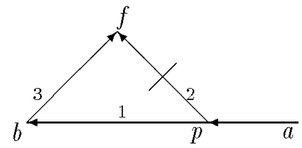
In this section we show that each defeasible inheritance network can be viewed as a default theory where the credulous semantics of the former coincides with the preferred semantics of the later. A defeasible inheritance network Γ is defined here as a finite collection of positive and negative direct links between nodes. If x, y are nodes then $x \rightarrow y$ (resp. $x \not\rightarrow y$) represents a positive (resp. negative) *direct link* from x to y. A network Γ is *consistent* if there exist no link $x \rightarrow y$ such that $x \not\rightarrow y$ also belongs to Γ. A *positive path* from x_1 to x_n through x_2, ..., x_{n-1}, denoted by $\pi(x_1, \sigma, x_n)$, is a sequence of direct links $x_1 \rightarrow x_2$, $x_2 \rightarrow x_3$, ... , $x_{n-1} \rightarrow x_n$. Similarly, a *negative path* from x_1 to x_n through x_2, ..., x_{n-1}, denoted by $\bar{\pi}(x_1, \sigma, x_n)$, is a sequence of direct links $x_1 \rightarrow x_2$, $x_2 \rightarrow x_3$, ... , $x_{n-1} \not\rightarrow x_n$. A *generalized path* is a sequence of direct links (x_1, x_2), (x_2, x_3), ... , (x_{n-1}, x_n), where (x_i, x_{i+1}) denotes a positive or negative direct link. Γ is *acyclic* if there is no generalized path (x_1, x_2), (x_2, x_3), ... , (x_{n-1}, x_n) with $x_1 = x_n$. Any path of the form $\pi(x, \sigma, y)$ *conflicts* with any path of the form $\bar{\pi}(x, \tau, y)$ and vice versa. Let Φ be a set of paths in Γ. A path σ is *conflicted in* Φ iff Φ contains a path that conflicts with σ. A positive path $\pi(x, \sigma, u) \rightarrow y$ is *constructible* in Φ iff $\pi(x, \sigma, u) \in \Phi$ and $u \rightarrow y \in \Gamma$. A negative path $\pi(x, \sigma, u) \not\rightarrow y$ is constructible in Φ iff $\pi(x, \sigma, u) \in \Phi$ and $u \not\rightarrow y \in \Gamma$. A positive path $\pi(x, \sigma, u) \rightarrow y$ is *preempted* in Φ iff there is a node v such that (i) $v \not\rightarrow y \in \Gamma$ and (ii) either $v = x$ or there is a path of the form $\pi(x, \tau_1, v, \tau_2, u) \in \Phi$. A negative path $\pi(x, \sigma, u) \not\rightarrow y$ is *preempted* in Φ iff there is a node v such that (i) $v \rightarrow y \in \Gamma$ and (ii) either $v = x$ or there is a path of the form $\pi(x, \tau_1, v, \tau_2, u) \in \Phi$. A *credulous extension* is a set Φ such that $\Phi = \{\alpha \,|\, \alpha$ is constructible, not conflicted, and not preempted in $\Phi\}$.

In the following Γ denotes a fixed but arbitrary defeasible inheritance network. A node x in Γ is called an *individual node* if there exists no direct link of the form $y \rightarrow x$ or $y \not\rightarrow x$ in Γ. Let I_Γ denote the set of individuals of Γ. An individual link is a link started from an individual node. A default link is a link which is not an individual link. The set of all default links in Γ is denoted by Γ_D. The inheritance network Γ can be represented as the default theory $T_\Gamma = (E, D)$ where

$D = \{p(x)/q(x) | p \rightarrow q \in \Gamma_D\} \cup \{p(x)/\neg q(x) | p \not\rightarrow q \in \Gamma_D\}$ and

$E = \{x(a) | a \rightarrow x \in \Gamma,\ a \in I_\Gamma\} \cup \{\neg x(a) | a \not\rightarrow x \in \Gamma,\ a \in I_\Gamma\}$

Example 4.4 Let Γ be an inheritance network represented by the picture below. It is easy to see that $T_\Gamma = (\{p(a)\}, \{p(x)/b(x),\ p(x)/\neg f(x),\ b(x)/f(x)\})$. □



A path in Γ is called *ground path* if it begins with an individual a.

Definition 4.13 *1. Let $p(a)$ be a ground literal. Then we write $\Gamma \mid\!\sim_c p(a)$ if each credulous extension of Γ contains a ground path of the form $a \rightarrow x_1 \rightarrow \ldots \rightarrow x_n \rightarrow p$*

2. Similarly we write $\Gamma \mid\!\sim_c \neg p(a)$ if each credulous extension of Γ contains a ground path of the form $a \rightarrow x_1 \rightarrow \ldots \rightarrow x_n \not\rightarrow p$.

The following theorem shows that our framework captures nonmonotonic inheritance reasoning.

Theorem 5 *Let Γ be an acyclic and consistent inheritance network and T_Γ be the corresponding default theory. Then for each ground literal w,*

$$\Gamma \mid\!\sim_c w \quad iff \quad T_\Gamma \mid\!\sim w$$

5 Transforming into Logic Programming

In this section we present a transformation from a default theory $T = (E, D)$ into an extended logic program P_T and prove the coincidence between the answer set semantics of P_T and the preferred extension semantics of T. We first recalled the notion of answer set of extended logic programs and then present the transformation from T to P_T. An *extended logic program P* [10] is a set of rules of the form

$$L_0 \leftarrow L_1, \ldots, L_m, not\ L_{m+1}, \ldots, not\ L_n,$$

where $0 \leq m \leq n$, each L_i is a literal, and *not* represents the negation-as-failure.

Let *Lit* denote the set of ground literals in the language of the program P.

1. If P does not contain *not* L (i.e. $m = n$ in every rule of P) an answer set of P is defined as the smallest set S, $S \subseteq Lit$ such that
   (a) for any ground instance $L_0 \leftarrow L_1, \ldots, L_m$ of a rule from P, if $L_1, \ldots, L_m \in S$, then $L_0 \in S$.
   (b) if S contains a pair of complementary literals, then $S = Lit$.
2. If the program P does contain *not* ($m < n$ in some rule of P), $S \subseteq Lit$ is an answer set of P if S is the answer set of the program P' obtained from the set of all ground instances of P by deleting

(a) each rule that has a formula *not* L in its body with $L \in S$, and

(b) all formulas of the form *not* L in the bodies of the remaining clauses.

Let $T = (E, D)$ be a default theory. To each default rule d_i: $l_1, \ldots, l_n / h_i$ we associate an unique abnormal predicate ab_i and write $hd_i(x)$ for $h_i(x)$, $bd_i(x)$ for $\{l_1(x), \ldots, l_n(x)\}$. Then, T is transformed into an extended logic program P_T as follows.

1. For each rule d_i, the two clauses

$$hd_i(x) \leftarrow bd_i(x),\ not\ ab_i(x)$$

$$ab_i(x) \leftarrow \neg hd_i(x)$$

belong to P_T, and

2. For each specificity relevant MCS C_{d_j} where d_j, d_t are its two conflicted defaults then the clause

$$ab_t(x) \leftarrow \bigwedge_{k \in C \setminus \{j,t\}} bd_k(x),\ not\ ab_k(x)$$

belongs to P_T, and

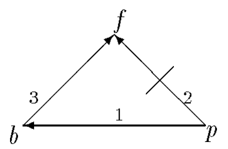
3. For each ground literal $l \in E$ the clause

$$l \leftarrow$$

is added to P_T.

We illustrate the transformation from T into P_T in the next examples.

Example 5.1 (Penguin)



P_T will contain following clauses:

$$\begin{array}{lcl} b(x) & \leftarrow & p(x), not\ ab_1(x) \\ ab_1(x) & \leftarrow & \neg b(x) \\ \\ \neg f(x) & \leftarrow & p(x), not\ ab_2(x) \\ ab_2(x) & \leftarrow & f(x) \\ \\ f(x) & \leftarrow & b(x), not\ ab_3(x) \\ ab_3(x) & \leftarrow & \neg f(x) \\ \\ ab_3(x) & \leftarrow & p(x), not\ ab_1(x) \end{array}$$

Let $E = \{p(a)\}$. Then, P_T has one more clause: $p(a) \leftarrow$. In this case we have only one answer set: $\{p(a), b(a), \neg f(a), ab_3(a)\}$. $\Box$

Definition 5.1 *Let P be an arbitrary extended logic program and w be an arbitrary ground literal. We write $P \mid\!\sim_{ans} w$ iff w is contained in each answer set of P.*

Theorem 6 *Let T be an acyclic and consistent default theory. Then for each ground literal w in the language of T,*

$$T \mid\!\sim w \quad iff \quad P_T \mid\!\sim_{ans} w$$

From the construction of P_T we can easily prove the following theorem about the modularity of the transformation from T to P_T.

Theorem 7 *Given a $T = (E, D)$ and $T' = (E', D')$ with $E \subseteq E'$ and $D \subseteq D'$. Then, $P_T \subseteq P_{T'}$.*

It is easy to see that if we translate the extended logic program P_T into a general logic program P'_T by: i) introducing a new literal $a'(x)$ representing the negative literal $\neg a(x)$ for each atom $a(x)$ in the language of P_T, and ii) replacing each negative literals $\neg a(x)$ of P_T by the newly introduced literal $a'(x)$ (see [10] for more detail about this transformation) then the stable semantics of the normal logic program P'_T coincides with the answer set semantics of P_T.

Theorem 8 *1. Let M be a stable model of P'_T. Then the set of literals obtained by replacing each atom of the form $a'(t)$ in M by $\neg a(t)$ is an answer set of P_T.*

2. Let M be an answer set of P_T. Then the set of atoms obtained by replacing each literals of the form $\neg a(t)$ in M by $a'(t)$ is a stable model of P'_T.

6 Conclusions

A new approach for reasoning with specificity in acyclic and consistent default theories has been proposed where the three basic ideas - "maximality", "anti-contrapositive", and "specific overriding" - have been combined to solve the specificity problem naturally. Here it is worthwhile to mention again that none of the general approaches to default reasoning with specificity proposed in the literatures up to now captures nonmonotonic inheritance reasoning and at the same time satisfies core properties of default reasoning. We also provide a transformation from acyclic and consistent default theories into general logic programs where answer set semantics coincides with our newly defined semantics. Thus, the new approach points out that logic program can be used to implement reasoning systems which can deal with specific information correctly. It is worth noting that the transformation is modular. Thus, it can be applied in building large scale knowledge base involving reasoning with specificity. But, many works are still needed, for example, the problem of reasoning with specificity in default theories where the background knowledge is not empty, i.e., $B \neq \emptyset$ and the set of evidences contains disjunctive evidences. These are also the topics of our next papers.

Acknowledgment

The first author is partially supported by EEC "Keep in Touch" activity KIT011-LPKRR. We would like to thank Michael Gelfond and the anonymous referees for their many constructive critisms, which helped us to improve the paper in many ways.

References

[1] Brewka, G.: Adding Priorities and Specificity to Default Logic, Proc. JELIA 94, LNAI 838, Springer Verlag, 247–260.

[2] Brewka, G.: Cumulative default logic: In defense of nonmonotonic inference rules, AI'50, 183–205.

[3] Baader, F. and Hollunder, B.: How to prefer more specific defaults in terminological default logic, IJCAI-93

[4] Bondarenko A., Dung P.M., Kowalski R.A., Toni F.: An abstract, argumentation-theoretic approach to default reasoning. To appear in AI Journal.

[5] Delgrande , J.P., and Schaub, T.H.: A General Approach to Specificity in Default Reasoning, Proceeding of KR'94, p. 147–158.

[6] Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning and logic programming and N-person game. AI Vol. **77** 2 (1995) 321–357

[7] Dung, P.M. and Son, T.C.: Nonmonotonic Inheritance , Argumentation, and Logic Programming, Proceeding of LPNMR, 1995, 316–329.

[8] Etherington, D., and Reiter, R.: On Inheritance Hierarchies with Exceptions in Proc. of AAAI-83, Pg. 104–108, 1983

[9] Gelfond, M., Lifschitz, V.: The stable model semantics for logic programs, Proceeding of the 5th ICLP, MIT Press, (1988), 1070 – 1079

[10] Gelfond, M., Lifschitz, V.: Logic Programs with Classical Negation, Proceeding of the 7th ICLP, MIT Press, (1990), 579 –597

[11] Geffner, H., and Pearl J.: Conditional entailment: bridging two approaches to default reasoning, AI'53, 209–244, 1992

[12] Geert, P. and Veimert, D.: A nonmonotonic reasoning formalism using implicit specificity information, LPNRM Proceeding, 1994, p. 380–396

[13] Horty, J.F.: Some direct Theories of Nonmonotonic Inheritance in Handbook of Logic and Artificial Intelligence and Logic Programming, D.Gabbay and C. Hogger, Oxford Uni., 1991.

[14] Thomason, R.H., and Horty, J.F: Logics for Inheritance Theory, 2nd Workshop on nonmonotonic reasoning

[15] Horty, J.F., Thomason, R.H., and Touretzky, D.S.: A skeptical theory of inheritance in nonmonotonic semantic networks, AI Vol. 42, pp. 311–348, 1987.

[16] Kraus, S., Lehmann, D., and Magidor, M.: Nonmonotonic Reasoning, Preferential Models and Cumulative Logics. AI 44 (1990) 167–207

[17] Moore, R.C.: Semantical Considerations on Nonmonotonic Logics, AI'25, 75–94

[18] McCarthy, J.: Circumscription - a form of nonmonotonic reasoning , AI'13, 27–39

[19] Matkinson, D., Schlechta K.: Floating conclusion and zombie paths: two deep difficulties in the 'directly skeptical' approach to defeasible inheritance nets, AI Vol. 48, pp 99–209

[20] McDermott, D. and Doyle, J.: Nonmonotonic Logic I, AI'13, 41–72

[21] Lin, F.: A study of nonmonotonic reasoning, Ph.D. Dissertation, Stanford University. (1991)

[22] Lifschitz, V.: Circumscription, in Handbook of Logic and Artificial Intelligence and Logic Programming, D.Gabbay and C. Hogger, Oxford University, 1994, pp 297–351.

[23] Lloyd, J.W.: Foundations of Logic Programming, Springer Verlag, 1987.

[24] Pearl, J.: System Z: A natural ordering of defaults with tractable applications to nonmonotonic reasoning: A survey; Proc. KR-89, Toronto, Morgan Kaufman, 505–516

[25] Pollock, J.L.: Defeasible reasoning. Cognitive Science **17** (1987) 481–518

[26] Poole, D.: On the comparison of theories: Preferring the most specific explanation, Proceeding IJCAI-85,pp.144–147

[27] Poole, D: A logical framework for default reasoning, AI 36, 1988

[28] Prakken, H., and Sartor, G.: A semantics for argument-based systems with weak and strong negation and with explicit priorities. Draft, July 1995. An extended abstract can be found in Proceedings of the 5th International Conf. on AI and Laws, ACM Press, 1995, 1-9.

[29] Reiter R.: A Logic for Default Reasoning, in Readings in Nonmonotonic Reasoning, Edited by M. L. Ginsberg, Morgan Kaufmann Publishers, Inc., Los Altos, California (1987) 68–93

[30] Simary, G.R., and Loui, R.P.: A mathematical treatment of defeasible reasoning and its implementation, Elsevier, AI, pp. 125–257.

[31] Simonet, G.: RS theory: a really skeptical theory of inheritance with exceptions, Proceeding of ECAI (1990) 615–626.

[32] Stein, L.A.: Resolving ambiguity in nonmonotonic inheritance hierarchies, AI'55, 259–310.

[33] Touretzky, D. S., Horty, J.F., and Thomason, R.H.: A clash of Intuition: The current state of Non-monotonic Multiple Inheritance Systems, IJCAI, 1987, pp.476–482

[34] Touretzky, D. S., Horty, J.F., and Thomason, R.H.: A skeptic's Menagerie: Conflictors, Preemptors, Reinstaters, and zombies in nonmonotonic Inheritance, IJCAI, 1991, pp.478–483

[35] Touretzky, D. S.: The mathematics of Inheritance Systems, Morgan Kaufmann Pub., Inc., Los Altos, CA, 1986.

[36] Van Gelder A., Ross, K., Schlipf J.S.: Unfounded sets and well-founded semantics for general logic programs. Proceeding of PODS 1988.

[37] Vreeswijk, G.A.W.: Studies in Defeasible Argumentation, Ph.D Thesis, Vrije University, Holland

Appendix 1: Proof of Theorem 1

We now prove the theorem 1. But first we need a couple of new notions and lemmas.

Let a be a ground atom. Then we define the *degree of* a in T, denoted by $deg_T(a)$, as the maximum of the lengths of the directed paths $a_0 \rightarrow \ldots \rightarrow a_n$ in the atom dependency graph of T such that $a = a_n$ and a_0 is the atom of some literal in E. The degree of a literal l is defined by $deg_T(l)$ (resp. $deg_T(\neg l)$) if l is a positive (resp. negative) literal. The *degree of an argument* is defined as the maximum of the degrees of the conclusions of its defaults. Similarly the *degree of a conjunction* is defined as the maximum of the degrees of its conjuncts.

Lemma 6.1 *Let a, b be two atoms such that there is a nonempty path from a to b in the atom dependency graph of T. Then $deg_T(a) < deg_T(b)$.*

Lemma 6.2 *Let $T = (E, D)$ be a default theory and A be a minimal (wrt set inclusion) argument supporting a, i.e., $E \vdash_A a$ and $\forall B \subset A, E \not\vdash_B a$. Then, there is a path from $atom(hd(d))$ to $atom(a)$ in G_T for every $d \in A$.* □

Lemma 6.3 *Let T be an acyclic and consistent default theory and A, B be two arguments in T such that A attacks B by specificity. Then, there is an argument $A' \subseteq A$ such that A' attacks B by specificity and $deg_T(A') < deg_T(B)$.*

Proof A attacks B by specificity means that there exists a specificity relevant C_d with d' being the other conflicted default in C_d such that $d' \in B$, and $C_0 = C_d \setminus \{d', d\} \subseteq A$ and A supports bd(d). From the definition of MCS, it follows immediately that $C_0 \cup \{d'\}$ is a minimal argument for $hd(d')$ wrt $bd(d)$. Therefore, following lemmas 6.2 and 6.1 we can conclude that for each default $x \in C_0$, $deg_T(hd(x)) < deg_T(hd(d'))$. Further since A supports $bd(d)$, there is a minimal set $C_1 \subseteq A$ such that $E \vdash_{C_1} bd(d)$. Hence from the lemmas 6.2 and 6.1, it follows that each default $x \in C_1$ $deg_T(hd(x)) \leq deg_T(bd(d)) < deg_T(hd(d)) = deg_T(hd(d'))$.

Consider $A' = C_0 \cup C_1$. It is clear that A' attacks B by specificity. Since for each $x \in C_0 \cup C_1$, $deg_T(hd(x)) < deg_T(hd(d'))$ and $d' \in B$ it is clear that $deg_T(A') < deg_T(B)$. The lemma is proved. □

Proof of Theorem 1 Let W be a preferred extension of AF_T. To prove that W is stable, we prove the following proposition by induction:

For each natural number i, the set $R_i = \{A \in AR_T \mid A \notin W$ and W does not attack A and $deg_T(A) = i\}$ is empty.

Let $i = 0$. Since the only argument A with $deg_T(A) = 0$ is the empty argument which belongs to W, it follows that $R_i = \emptyset$.

Assume that we have proved that $R_k = \emptyset$. We now prove that $R_{k+1} = \emptyset$.

Assume that there exists $A \in R_{k+1}$ (i.e. $R_{k+1} \neq \emptyset$). Consider $W' = W \cup \{A\}$. Since W is preferred, W' is not admissible, i.e. there is an argument B such that B attacks A and W' does not attack B. Thus, B attacks A by specificity. Because of lemma 6.3 we know that there exists some argument B' such that $B' \subseteq B$ and B' attacks A by specificity and $deg_T(B') < deg_T(A)$. Since W does not attack A, B' does not belong to W. Since $B' \subseteq B$ and W' does not attack B, W does not attack B' either. By inductive hypothesis, such B' does not exist. Hence A does not exist either. That means that $R_{k+1} = \emptyset$.

□

Appendix 2: Proof of the Conditioning Property (theorem 2)

The proof is based on the following lemma.

Lemma 6.4 *Let $T = (E, D)$ be an acyclic and consistent default theory with $E = bd(d')$ for some $d' \in ground(D)$. Then, the argument $B = \{d'\}$ is not attacked by specificity.*

Proof Assume the contrary that there is an argument A that attacks B by specificity. Then, by definition

4.11 there is a specificity relevant MCS C_d such that d, d' are the conflicted defaults in C_d such that

- $E \vdash_A bd(d)$, and $C_d \setminus \{d', d\} \subseteq A$, and

Since $E = bd(d')$, we have $bd(d') \vdash_A bd(d)$.

On the other side, from $d' \in C_d$ and lemma 4.1 we have that there exists $S \subseteq D$ such that $bd(d) \vdash_S bd(d')$. Because T is acyclic we have $bd(d) \equiv bd(d')$. Hence T is inconsistent. Contradiction. !!! $\Box$

Proof of Theorem 2 To prove the theorem we prove that every preferred extension of T contains the argument $B = \{d\}$. Assume the contrary, there exists a preferred extension S of T which does not contain B (i.e. $B \notin S$). By theorem 1, S attacks B. Consider an argument A in S that attacks B. By lemma 6.4, A and B are in conflict. Since $Con(B) = E \cup \{hd(d)\}$ and $bd(d) \cup \{hd(d)\}$ is consistent, $bd(d) \vdash_A \neg hd(d)$. So, $A \neq \emptyset$. Since $bd(d) \vdash_A \neg hd(d)$ and T is consistent there exists a minimal argument $K \subseteq A$, $K \neq \emptyset$ such that $bd(d) \vdash_K \neg hd(d)$ and $bd(d) \not\vdash_K false$. It follows that K contains a default x with $hd(x) = \neg hd(d)$. It follows that $C = K \cup \{d\}$ is a MCS wrt. d with d and x are its two conflicted defaults. Because for all $y \in K$, $bd(d) \vdash_K y$ d is the most specific default of C. Hence, $C = K \cup \{d\}$ is a specificity relevant MCS wrt. d. Thus, by definition of attack by specificity, $K' = K \setminus \{x\}$ attacks K by specificity. Because S is a preferred extension and $A \in S$, $K' \subseteq K \subseteq S$, $K, K' \in S$. Therefore, S is conflict. Contradiction because S is a preferred extension of T. $\Box$

Appendix 3: Proof of the Reduction Property (theorem 3)

Let AR_T and $AR_{T'}$ be the set of arguments of T and T' respectively. Further, let

$DA_1 = AR_T \setminus AR_{T'}$ and $DA_2 = AR_{T'} \setminus AR_T$

It is easy to see that

$DA_1 = \{A \in AR_T \mid s.t.\ \neg a \in Con(A)\}$ and

$DA_2 = \{A \in AR_{T'}\ s.t.\ \exists d \in A,\ E \not\vdash_A bd(d)\}$.

Lemma 6.5 *let S be an arbitrary stable extension of T and*

$S' = S \cup \{A' \in DA_2\ s.t.\ \exists A:\ A$ supports a and $A \cup A' \in S\}$.

Then S' is a stable extension of T'.

Proof Since $T \mid\!\sim a$ there exists an argument $A \in S$ supporting a and there are no arguments in S supporting $\neg a$. Hence, $S \subseteq AR_{T'}$.

We first prove that S' is conflict free. Assume the contrary that there are P and Q in S' such that P attacks Q. There are two cases:

1. $P \in S$ and $Q \in DA_2$. Then, there is some $A \in S$ such that $A \cup Q \in S$ and P attacks $A \cup Q$ (wrt T). Contradictory !

2. $P \in DA_2$ and $Q \in DA_2$. Then, there are $A, B \in S$ such that $A \cup P \in S$ and $B \cup Q \in S$ and $A \cup P$ attacks $B \cup Q$ (wrt T). Contradictory !

So, we have proved that S' is conflict free (i).

Now we prove that S' is stable, i.e. S' attacks every argument that does not belong to S'.

Let $P \notin S'$. Then either $P \in AR_T$ or $P \in DA_2$.

If $P \in AR_T$ then S attacks P since S is a stable extension of T (ii).

If $P \in DA_2$ then it follows directly from the definition of S' that for each $A \in S$ that supports a, either $A \cup P$ is not an argument of T or $A \cup P$ is an argument of T but $A \cup P \notin S$.

If $A \cup P$ is not an argument of T then A, P are in conflict wrt T. So A attacks P by conflict wrt T'.

If $A \cup P$ is an argument in T but $A \cup P \notin S$ then S attacks $A \cup P$ since S is a stable extension of T. Hence, S' attacks P wrt T'.

So we have that S' is a stable extension of T'. $\Box$

Proof of Theorem 3

Let $T = (E, D)$ such that $T \mid\!\sim a$ and $T' \mid\!\sim b$ for $T' = (E \cup \{a\}, D)$. To prove that $T \mid\!\sim b$, we prove that each stable extension of T containing an argument supporting b. Let S be a stable extension of T. From the above lemma, we have that $S' = S \cup \{A' \in DA_2\ s.t.\ \exists A:\ A$ supports a and $A \cup A' \in S\}$ is also a stable extension of T'. Therefore there exists an argument $B \in S'$ which supports b. If $B \in S$ then the theorem is proved. If $B \in DA_2$ then there is $A \in S$ such that A supports a and $B \cup A \in S$. Since B supports b wrt T', it is clear that $B \cup A$ supports b. $\Box$.

Appendix 4: Proofs of the Cumulative Property (theorem 4)

We call an argument B a subargument of an argument A if $B \subseteq A$.

In following, by $T_a = (E, D')$ we denote the default theory which is obtained from a default theory $T = (E, D)$ by removing all the default $d \in D$ such that there is no path from $atom(hd(d))$ to $atom(a)$ in G_T. Obviously, an argument in T_a is an argument in T but not vice versa.

Lemma 6.6 *Let $T = (E, D)$ be a default theory and a be a ground literal in the language $\mathcal{L}$ and A be an argument in T_a. Then, each argument $B \in AR_T$ which attacks A contains a subargument $B' \subseteq B$ such that $B' \in AR_{T_a}$ and B' attacks A wrt T_a.*

Proof

Let B' be a minimal (wrt set inclusion) subargument of B that attacks A wrt T.

If B' is in conflict with A then there exists a literal l such that $l \in Con(A)$ and $\neg l \in Con(B')$. Since $E \not\vdash_A false$ and $E \not\vdash_{B'} false$, neither l nor $\neg l$ belongs to E. Therefore both A and B' are not empty and there exist $d \in A$ and $d' \in B'$ such that $hd(d) = l$ and $hd(d') = \neg l$. Since B' be a minimal (wrt set inclusion) subargument of B that attacks A, it is clear that B' is a minimal argument supporting $\neg l$. Therefore following lemma 6.2, there is a path from the atom of $hd(x)$ to $atom(l)$ for each $x \in B'$. Since $atom(hd(d)) = atom(hd(d'))$ and $d \in A$ and $A \in AR_{T_a}$, it is clear that there is a path from $hd(l)$ to $atom(a)$ in G_T. From the definition of T_a, it follows that B' is an argument of T_a.

If B' attacks A by specificity then there is a specificity relevant MCS C_d with d' being the other conflicted default and $d' \in A$ and $C_d \setminus \{d', d\} \subseteq B'$ and $bd(d) \subseteq Con(B')$. From the minimality of B' and lemma 4.1 it is easy to see that for any default $x \in B'$ there is some path from $atom(hd(x))$ to $atom(hd(d'))$ from which there is a path to $atom(a)$. Thus, B' is an argument in T_a. □

The next lemma follows from lemma 6.6.

Lemma 6.7 *Let T be a default theory and a be a literal in $\mathcal{L}$. Then,*

1. *If W is a stable extension of T then $W_a = W \cap AR_{T_a}$ is a stable extension of T_a.*
2. *If W_a is a stable extension of T_a then there exists a stable extension W of T such that $W_a \subseteq W$.*

Proof

1. Let W be a stable extension of T and $W_a = W \cap AR_{T_a}$. Further let $A \in AR_{T_a}$ such that $A \notin W_a$. Therefore $A \notin W$. Since W is a stable extension of T, there is an argument $B \in W$ such that B attacks A wrt T.
   Lemma 6.6 implies that there is a subargument B' of B such that B' attacks A wrt T_a and $B' \in AF_{T_a}$.
   Since W is a stable extension of T, we have that $B' \in W$. Thus, $B' \in W_a$. Hence, W_a attacks A. That means that W_a is stable.
2. Now, let W_a be a stable extension of T_a. We will prove that W_a is an admissible set of arguments in T. Let $B \in AR_T$ be an argument which attacks W_a. Then, by lemma 6.6 there is a subargument B' of B such that $B' \in AF_{T_a}$ and B' attacks W_a. By definition of stable extension we have that W_a attacks B' wrt T_a and therefore W_a attacks B wrt T. Thus, W_a is an admissible set of AF_T. It follows that there is a preferred extension W of T such that $W_a \subseteq W$. By theorem 1 we have that W is also stable extension of T. The lemma is proved. □

Lemma 6.8 *Let $T = (E, D)$ be a default theory such that $T \mid\!\sim a$. Then following assertions hold.*

1. *$T_a \mid\!\sim a$.*
2. *Let $T' = (E \cup \{a\}, D)$ and S be a stable extension of T'_a. Then S is also a stable extension of T_a.*

Proof

1. Let S be a stable extension of T_a. Then there is a stable extension S' of T such that $S \subseteq S'$. Since $S \subseteq (S' \cap AR_{T_a})$ and $S' \cap AR_{T_a}$ is a stable extension of T_a, it follows $S = (S' \cap AR_{T_a})$. It is easy to see that every minimal argument in AR_T which supports a belongs to AR_{T_a}. Since $T \mid\!\sim a$ there is a minimal argument $A \in S'$ such that A supports a. Therefore $A \in S$. Part (1) of the lemma is proved.
2. We need to prove first the following proposition.
   Proposition For each argument $K \in AR_{T'_a}$, if S attacks K wrt T'_a then S attacks K wrt T_a.

   Proof: Let $H \in S$ such that H attacks K wrt T'_a.
   If H attacks K (wrt T'_a) by conflict then there is a ground literal l such that $E \cup \{a\} \vdash_H l$ and $E \cup \{a\} \vdash_K \neg l$. Since both H and K are arguments of T'_a, it is clear that $atom(l) \neq atom(a)$. From the definition of T'_a, it should be clear that $E \vdash_H l$ and $E \vdash_K \neg l$. Hence H attacks K wrt T_a by conflict.
   Now let H attack K wrt T'_a by specificity. Hence there exists a specificity relevant MCS C_{d_1} whose conflicted defaults are d_1, d_2 such that
   $E \cup \{a\} \vdash_H bd(d_1)$, and $C_d \setminus \{d_1, d_2\} \subseteq H$, and $d_2 \in K$.
   From the definition of T'_a, it follows again directly that $E \vdash_H bd(d_1)$. That means that H attacks K by specificity wrt T_a. □
   We continue again with the proof of the lemma 6.8. Since $S \subseteq AR_{T'_a}$ and $AR_{T'_a} \subseteq AR_{T_a}$ we have $S \subseteq AR_{T_a}$. We want to prove now that S is an admissible set of arguments in T_a. Clearly, S is conflict free. Thus, we only need to prove that for each $A \in AR_{T_a}$, if A attacks S wrt T_a then S attacks A wrt T_a.
   There are two cases:
   (a) A attacks S (wrt T_a) by conflict. Then it is clear that S attacks A (wrt T_a) by conflict too.
   (b) A attacks S (wrt T_a) by specificity.
   Hence there exists an argument $B \in S$ and a specificity relevant MCS C_d whose conflicted defaults are d, d' such that

- $E \vdash_A bd(d)$, and $C_d \setminus \{d', d\} \subseteq A$, and
- $d' \in B$.

Let C_0 be a minimal subargument of A such that $E \vdash_{C_0} bd(d)$ and $C_1 = C_d \setminus \{d, d'\}$. Let $C = C_0 \cup C_1$. Then from the definition of T_a, C_0, C_1 together with the fact that $A \in AR_{T_a}$ and T is acyclic, it follows directly that C is an argument in AR_{T_a} and $E \not\vdash_C \neg a$. Hence $C \in AR_{T'_a}$ and C attacks B wrt T'_a. Therefore S attacks C wrt T'_a. From the above proposition, it follows that S attacks C wrt T_a. Hence S attacks A wrt T_a, too. So we have proved that S is admissible wrt T_a.

Let W be a preferred extension of T_a such that $S \subseteq W$. From part (1) of this lemma, it follows that $W \mid\!\sim a$. Hence W is also a stable extension of T'_a. Since S is a stable extension of T'_a, it follows that W and S coincide. That means that S is a stable extension of T_a. □

Lemma 6.9 *Let $T = (E, D)$ be a default theory such that $T \mid\!\sim a$ and $E \not\vdash_\emptyset a$. Then, each stable extension of $T' = (E \cup \{a\}, D)$ contains a nonempty argument $A \in AR_T$ supporting a (wrt T).*

Proof Consider the default theory T'_a. It is easy to see that $AR_{T'_a} \subseteq AR_{T_a}$ and $AR_{T_a} \setminus AR_{T'_a} = \{A \ s.t.\ A \in AR_{T_a},\ A$ supports $\neg a\}$. Assuming now the contrary that there exists a stable extension S of T' which contains no argument A such that $A \neq \emptyset$ and A supports a wrt T.

Let $S_a = S \cap AR_{T'_a}$. From the lemma 6.7 and part 2 of the lemma 6.8, it follows that S_a is a stable extension of T_a. Since every argument in AR_{T_a} supporting a is also an argument in $AR_{T'_a}$, S_a attacks every argument which supports a. So S_a contains no argument supporting a. Contradictory because $T_a \mid\!\sim a$ (part (1) lemma 6.8). Thus, our assumption that S contains no argument A such that $A \neq \emptyset$ and A supports a wrt T is incorrect. Hence, the lemma is proved. □

Lemma 6.10 *Let $T = (E, D)$ be a default theory such that $T \mid\!\sim a$. Further let $T' = (E \cup \{a\}, D)$ and W' be a stable extension of T'. Then $W = W' \cap AR_T$ is a stable extension of T*

Proof There are two cases:

1. $a \in E$. Then the theorem is trivial because $T = T'$.
2. $a \notin E$, then $E \not\vdash_\emptyset a$. Let $B \in AR_T \setminus W$. If $E \not\vdash_B \neg a$ then B is also an argument of T'. Therefore B is attacked by some $A' \in W'$. From lemma 6.9, there is $A \in W$ such that A supports a. Hence it is clear that B is attacked by $A \cup A'$ wrt T.

   If $E \vdash_B \neg a$ then A attacks B by conflict.

□

Appendix 5

Let E be a credulous extension of Γ. Then, the set of ground paths in E, denoted by E_G, is called a *ground credulous extension* of Γ.

For a ground path $\alpha = \pi(a, x_1, ..., x_n)$ in Γ, A_α denotes the set consisting of ground defaults $x_i(a)/x_{i+1}(a)$, $1 \leq i \leq n-1$.

For a set of ground paths Φ, let $A_\Phi = \bigcup_{\alpha \in \Phi} A_\alpha$.

Given an argument B in T_Γ we define $\Phi_B = \{\alpha \mid \alpha$ is a ground path in Γ, s.t. $A_\alpha \subseteq B\}$ and for a set W of arguments we write $\Phi_W = \bigcup_{B \in W} \Phi_B$.

It is easy to see that if Γ is acyclic and consistent then T_Γ is also acyclic and consistent. And, if α be a ground path in Γ then A_α is an argument in T_Γ. Also, it is clear that if a ground literal l is supported by an argument A then there exists a ground path $\alpha \in \Phi_A$ such that $c(\alpha) = l$. In the next sequel a path $\alpha = \pi(x, \sigma, u)$ is called a *prefix* of a path $\beta = \pi(x, \sigma, u, \tau, v)$ in Γ.

A ground path $\alpha = \pi(a, .., x)$ is said to *support the conclusion* $x(a)$ (resp. ¬x(a)). $x(a)$ (resp. $\neg x(a)$) is called the conclusion of α, denoted by $c(\alpha)$.

For a set Φ of ground paths, $c(\Phi) = \{c(\alpha) \mid \alpha \in \Phi\}$.

Lemma 6.11 *Let Γ be an acyclic and consistent inheritance network and Φ be a credulous extension of Γ. Then, $W = \{A \mid A$ is an argument, $A \subseteq A_\Phi\}$ is a preferred extension of T_Γ.*

Proof It is easy to see that A_Φ is an attack-free argument. We first prove that A_Φ attacks every argument B that attacks A_Φ. Consider two cases: 1. B attacks A_Φ by conflict, then A_Φ attacks B by definition. 2. B attacks A_Φ by specificity. It means there is a path $\alpha \in \Phi$ s.t. α is preempted in Φ_B. Without the lose of generality we assume that $\alpha = \pi(x, \ldots, u, y)$. α is preempted in Φ_B means there exists $\beta = \pi(x, \ldots, v, \ldots, u) \in \Phi_B$ and $u \not\to y \in \Gamma$. Because $\alpha \in \Phi$ we have $\beta \notin \Phi$. Hence, there exists a prefix δ of β which is conflicted or preempted in Φ. It follows then A_Φ attacks B by conflict or by specificity. Thus, A_Φ attacks every argument B that attacks A_Φ. Hence, for all $A \in W$, A is defendable in W (i). Now, let B be an argument and $B \notin W$. Trivially, there exists some path $\alpha \notin \Phi$ s.t. $A_\alpha \subseteq B$, $A_\alpha \setminus A_\Phi \neq \emptyset$ and therefore A_Φ attacks A_α. Thus, A_Φ attacks B (ii). From (i)-(ii) we can conclude that W is a preferred extension of T_Γ. □

Lemma 6.12 *Let W be a preferred extension of T_Γ. Then, Φ_W is a ground credulous extension of Γ.*

Proof Let σ be a ground path in Γ. We prove that (i) if $\sigma \in \Phi_W$ then σ is defeasibly inheritable in Φ_W and (ii) if $\sigma \notin \Phi_W$ then σ is not defeasibly inheritable in Φ_W. We first prove (i). It follows from $A_\sigma \subseteq W$

that $A_\delta \subseteq W$ if δ is a prefix of σ. Hence, all prefixes of σ belong to Φ_W. So, σ is constructible in Φ_W. Furthermore, W is a preferred extension implies that σ is not conflicted or preempted in Φ_W. Thus, $\Phi_W \mid\!\sim \sigma$ for $\sigma \in \Phi_W$ (1). Now, consider $\sigma \notin \Phi_W$. Without the lose of generality we assume that all prefixes of σ are in Φ_W. It follows that A_σ is attacked by W. If A_σ is attacked by conflict then σ is conflicted in Φ_W. If A_σ is attacked by specificity then it is preempted in Φ_W. Thus, $\Phi_W \not\mid\!\sim \sigma$ for $\sigma \notin \Phi_W$ (2). From (1) and (2) we can conclude that Φ_W is a ground credulous extension of Γ. The lemma is proved. □

Theorem 5 follows from lemma 6.11 and lemma 6.12 .

Appendix 6

Let I_T be the set of constants in T and AB_T denotes the set of all ground positive literals of abnormal-predicates in P_T. Further, for a set W of arguments in T we write $Con(W) = \bigcup_{A \in W} Con(A)$. For a preferred extension W of T let $Ab_W = \{ab_j(a) | \ a \in I_T, d_j \in D,$ $bd(d_j) \subseteq Con(W),\ hd(d_j) \notin Con(W)\}$ We prove that

Lemma 6.13 *Let W be a preferred extension of T. Then, $S = Con(W) \cup Ab_W$ is an answer set of P_T.*

Proof S is an answer set of P_T [10] if S is the answer set of $P_T(S)$ which is obtained by (i) deleting all clauses containing one literal *not* $ab_j(a)$ if $ab_j(a) \in S$ and (ii) deleting all the abnormal-literals in the remaining clauses. In following let $\alpha(P_T(S))$ be the answer set of $P_T(S)$. By definition of $Con(W)$ and Ab_E, we prove the lemma by two cases:

- Case: $q(a) \in S$
  iff there is a sequence σ of $d_1, \ldots, d_n$ in W s.t. $hd(d_n) = q(a)$ and $E \vdash_\sigma bd(d_i)$ for $1 \leq i \leq n$
  iff $ab_j(a) \in (AB_T \setminus Ab_W)$ for all d_j
  iff $ab_j(a) \notin S$ for all defaults j belongs to σ
  iff $P_T(S)$ consists of $hd_j(a) \leftarrow bd_j(a)$ for all d_j belongs to σ
  iff $q(a) \in \alpha(P_\Gamma(S))$.
- Case: $ab_j(a) \in S$
  iff $ab_j(a) \in Ab_W$
  iff $\exists d_j \in D$, $bd_j(a) \subseteq Con(W)$ but $hd_j(a) \notin W$
  iff there is an argument $A \in W$ but $B = A \cup \{d_j\} \notin W$
  iff B is attacked by W
  iff $ab_j(a) \leftarrow \neg hd_j(a)$ is a clause of $P_T(S)$ and $\neg hd_j(a) \in S$
  or $ab_j(a) \leftarrow \bigcup_{k \in C_i} hd_k(a)$ is a clause of $P_T(S)$ and $\forall k \in C_i :\ hd_k(a) \in S$
  iff $ab_j(a) \in \alpha(P_T(S))$.

For a set of ground literal M of the program P_T we define $A_M = \{A \ |A$ is an argument in T s.t. $Con(A) \subseteq M\}$, $P_M = \{A \ |A \in A_M,\ A$ is attacked by A_M by specificity$\}$. We show that

Lemma 6.14 *Let M be an answer set of P_T. Then, $A_M \setminus P_M$ is a preferred extension of T.*

Proof Let $W = A_M \setminus P_M$. It is easy to see that there exists no argument $A \in A_M$ such that A attacks A_M by conflict because M is consistent. Hence, W is attack-free. Consider an argument A in T, if $A \in W$ then we can prove that W attacks B if B attacks A. Otherwise, $A \notin W$. It means that there exists some default $d_i \in A \setminus W$. So, $ab_i \in M$. There are two cases: (i) $\neg h_i$ is supported in M or (ii) there is a specific relevant MCS C such that d_i is one of the conflicted default but not the most specific default. It is easy to see that in (i) W attacks A by conflict and in (ii) W attacks A by specificity. Thus, W attacks A if $A \notin W$. Hence, the lemma is proved. □

From lemma 6.13, 6.14 we have Theorem 6.

Proof of Theorem 8 According to [10] we need to prove that there is no consistent stable model S of P'_T and a be an arbitrary ground atom in the language of P_T such that $\{a, a'\} \subset S$. Assume the con trary. Since there is no negative abnormal literal in P_T we can conclude that a is a ordinary literal in the language $\mathcal{L}$ of T. Further, because T is consistent $\{a, a'\} \subset S$ if and only if there exists a constant k and two ground instances of P'_T such that $hd_i(k) \leftarrow bd_i(k), ab_i(k)$ and $hd_j(k) \leftarrow bd_j(k), ab_i(k)$ such that $a = hd_i(k)$, $a' = hd_j(k)$, and $bd_i(k) \subset S$, $bd_j(k) \subset S$, $\{ab_i(k), ab_j(k)\} \cap S = \emptyset$ (*). But, by definition of P_T we have that $ab_i(k) \leftarrow hd_j(k)$ and $ab_j(k) \leftarrow hd_i(k)$ are ground instances of P'_T. Hence, S must include $\{ab_i(k), ab_j(k)\}$ if S is a stable model. Contradictory to (*). !!! The theorem is proved. □