Abnormality Counts!

MSc Thesis *(Afstudeerscriptie)*

written by

Aleks Knoks
(born September 2nd, 1986 in Riga, Latvia)

under the supervision of **Prof. Dr. Frank Veltman**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
September 19th, 2012	Dr. Reinhard Blutner
	Prof. Dr. Bendedikt Löwe
	Dr. Sonja Smets
	Prof. Dr. Frank Veltman

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

Defeasible inheritance networks provide a fruitful environment for modeling default reasoning. In this paper we aim to enhance their expressive power with the idea of abnormality minimisation, which plays a crucial role in circumscription. We implement it in two alternative network-based frameworks. The first is in line with the received way of modeling default reasoning by means of inheritance nets, that is to say, it is path-based or direct. The second relies on a set of non-monotonic inference rules. In both we are able to say that an object is abnormal with respect to a certain default statement, and, consequently, single out the conclusion sets that imply the least number of abnormalities. We consider some cases indicating that exactly these sets are the intuitively correct ones.

Contents

Setting the Stage

Default reasoning is best characterised as making sensible guesses in the context of incomplete information. It is hardly possible to overestimate its importance, for it is an essential part of our everyday reasoning patterns; and it is only natural that it has long since become an area of intensive research in the field of artificial intelligence. When we reason by default we proceed by means of the so-called 'default rules' (or simply 'defaults'), which are essentially statements of the form 'x-s usually are [not] y-s'. The statement 'Birds usually are able to fly' serves as the standard example. If the only information we know about some object x is that it is a bird, we would readily assume that x is able to fly on the basis of this default. However, in case later we would come to know that x is in fact an ostrich, this assumption would have to be retracted. This simple example alone makes it plain that any formalism attempting to model default reasoning has to rely on a non-monotonic consequence relation.

The first formalism of such kind — default logic — was proposed by Raymond Reiter in 1980 (see [11]), but it was soon followed by many other alternative frameworks. The more successful ones fall under three categories: (i) the approaches that model defaults as conditionals of a special kind,[1] (ii) the ones that rely on the idea of *circumscription*, and (iii) those that formalise defaults by means of the so-called 'defeasible inheritance networks'. Now, the two frameworks that will be developed in this paper combine the positive features of the approaches falling under the latter two categories. Therefore, something more has to be said about them.

Circumscription was introduced in (McCarthy, [9]), and it is, essentially, an attempt to formalise the common sense assumption that things are just as we expect them to be, unless there are some special circumstances. In [10] it is applied specifically to defaults. Although the formalism John McCarthy presents is complex, the motivating idea is fairly simple: to use circumscription in order to single out those models in which objects behave just like the defaults describe. This is carried out as 'minimisation of abnormality'. McCarthy introduces a special predicate ab, standing for *abnormality*, and suggests a peculiar formalisation for default statements. For instance, the above 'Birds usually are able to fly' would be translated as $\forall x(Bird\ x \wedge \neg ab\ aspect_f\ x \rightarrow AbleToFly\ x)$. What this formula says is that any object, which is a bird and is known not to be abnormal (= is known to be normal) in a certain respect, is also able to fly. The minimisation part means focusing only on those models (of some given set of defaults) that are *minimal* for the predicate ab. Intuitively, a model $\mathfrak{M}$ is minimal for ab if as few as possible elements belonging to the domain of $\mathfrak{M}$ satisfy ab. In other words, mini-

[1] The framework of (Delgrande, [3]), the approaches proposed by Craig Boutilier in [1] and [2], as well as that of (Veltman, [19]) fall under this category among others.

mal models are the ones containing the least number of abnormal objects. Further, the notion of a minimal model is used to define an appropriate consequence relation. Basically, it is the standard model theoretic entailment, but restricted to the minimal models. In general, McCarthy's approach gives intuitive results in the simple cases, but it has two serious drawbacks, which become evident as soon as it is applied to more complex theories. First, it is not at all pleasing computationally. Second, it requires one to state explicitly all the exceptions to every default rule in the theory at hand. Consider the following set of statements:

'Birds usually are able to fly',
'Ostriches usually are not able to fly',
'Ostriches are birds',
'Ostriches usually are fast runners', and
'Feathers is an ostrich'.

Unfortunately, it will not entail 'Feathers is not able to fly' until it is supplemented by a statement explicitly specifying that ostriches are abnormal with respect to flying. This problem becomes even graver as the number of defaults in a theory grows. However, it does not even occur in the family of approaches we consider next.

First and foremost, defeasible inheritance nets are associates with the works of David Touretzky, Richard Thomason, and John Horty (see [6], [7], [16], [17], and [18]). The forerunners of these nets — 'inheritance systems' or 'hierarchies', as they are usually called — were and still are used for representing and accessing taxonomic information. Although inheritance networks go far beyond a mere representational mechanism — they are best thought of as sets of hypotheses provided to some agent or reasoning mechanism —, much of the standard terminology used to describe them and, in fact, also the general way of thinking about them comes exactly from the work on the systems. Here (Touretzky, [17]) is the most notable source. Following the theoretical literature, we will describe inheritance networks as finite collections of nodes and positive and negative `IS-A` links, which we will write as $(x, z, +)$ and $(x, z, -)$. The interpretation of $(x, z, +)$ and $(x, z, -)$ depends on x. If it is a general term, their intended meanings are 'x-s *usually* are z-s' and 'x-s *usually* are not z-s', respectively. Thus, in this case the links are nothing else but default statements. In case x is an individual or a concrete object, however, $(x, z, +)$ should be read as 'x is a z' and $(x, z, -)$ as 'x is not a z'. A simple example will clarify this better than anything. We can rewrite the theory considered above as:

$(Bird, AbleToFly, +)$,
$(Ostrich, AbleToFly, -)$,

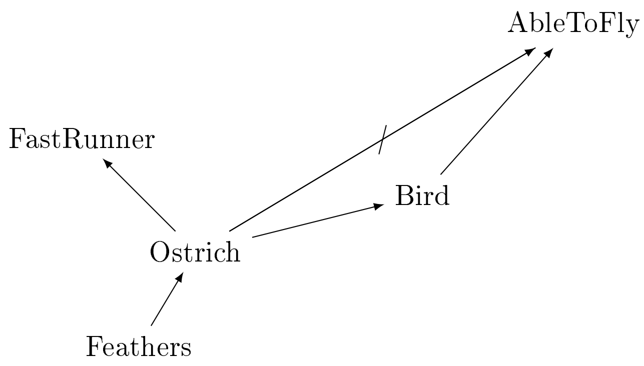

Figure 1: Γ_1

$(Ostrich, Bird, +)$,[2]
$(Ostrich, FastRunner, +)$,
$(Feathers, Ostrich, +)$,

Figure 1 contains a graphical representation of these links, and it surely seems to effectively reveal the structure of the information they encode. The picture also illustrates the graphical notation we will use throughout this paper: ordinary arrows for the positive `IS-A` links and arrows with crossbars for the negative ones.

The primary goal of any network-based account is to link inheritance networks to their corresponding *conclusion sets*. Intuitively, the conclusion set of a network is the set of all and only those statements that an ideal reasoner would arrive at on the basis of the information specified by the network. In the case of Γ_1 such a set would certainly have to contain statements that are similar to 'Feathers is not able to fly', 'Feathers is a fast runner', etc. Just like the links of a net, statements that are allowed to enter a conclusion set can be of two kinds: positive, written as $isa(x, z, +)$, and negative, written as $isa(x, z, -)$. The intended meaning of the positive ones is 'it is natural to suppose that x is a z (/x-s are z-s)'; and that of the negative ones is 'it is natural to suppose that x is not a z (/x-s are not z-s)'. In analogy to the links, we will call these '`IS-A` statements'. Now, there are different ways to get to the conclusion set corresponding to a given network. The one that is favoured by Touretzky *et al.* is roundabout. One first establishes the so-called 'expansion' (expansions) of a given network, and then constructs

[2]Strictly speaking, $(Ostrich, Bird, +)$ is equivalent to 'Ostriches usually are birds', rather than the universal statement presented above. However, it does not really matter for the present purposes. We can see universal statements as special kinds of defaults, namely, those that do not allow for exceptions. Besides, inheritance networks that allow for both default and universal statements are well-known in the literature.

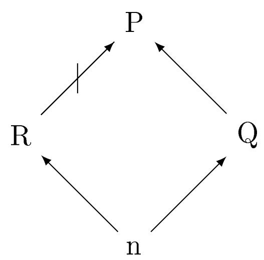

Figure 2: Γ_2, the Nixon Diamond

the conclusion set from it (them).[3] These expansions are sets of *paths*; and paths are nothing but special sequences of links. For instance, the sequence consisting of $(Feathers, Ostrich, +)$ and $(Ostrich, AbleToFly, -)$, written as $(Feathers, Ostrich, AbleToFly, -)$, is a path. Note that any link by itself is a path. (Horty, [6]) suggests viewing paths as *arguments supporting* statements that can enter a conclusion set, and we will follow him in that. Thus, we will say that the path $(Feathers, Ostrich, AbleToFly, -)$ supports the statement $isa(Feathers, Ostrich, -)$. Henceforth, any approach that relies on paths when determining conclusion sets will be referred to as *path-based*.

Every path-based approach can be distinguished by two things: (*i*) the exact notion of expansion it relies on, and (*ii*) the method it uses in order to construct expansions from a given inheritance net. A more formal treatment of this issue will have to wait until the next section. Here we will confine ourselves to an informal introduction of two representative approaches. The first one goes back to (Touretzky, [17]) and is known as the 'credulous approach'. It defines expansions as maximal consistent sets of argument paths. Any single set of this kind, i.e., any *credulous expansions*, is taken to be one possible and internally consistent 'line of reasoning' on the basis of the information contained in the inheritance net at hand. A simple, but illustrative example will make this clear. Consider the network Γ_2 depicted in Figure 2. Its interpretation is n = Nixon, R = Republican, Q = Quaker, P = pacifist. Now, this inheritance net has two credulous expansions. The first one consists of the paths $(n, R, +)$, $(n, Q, +)$, and $(n, Q, P, +)$. The second also contains $(n, R, +)$ and $(n, Q, +)$, but the place of $(n, Q, P, +)$ is taken by $(n, R, P, -)$. Thus, on this approach, given Γ_2, the reasoner is justified to infer (*i*) $isa(n, R, +)$, (*ii*) $isa(n, Q, +)$, and (*iii*) either $isa(n, P, +)$ or $isa(n, P, -)$, but certainly not both.

[3] A very different way to get to the conclusion sets goes via translating the networks into one or another of the non-monotonic logics. (Etherington & Reiter [5]) do just that, but, just as it was in McCarthy's circumscription approach, they are forced to list all the exceptions to each default explicitly. A very different way to get to the conclusion set has been introduced in (Sandewall, [13]) and taken up by Geneviève Simonet in [15]. It will be treated in detail in later sections.

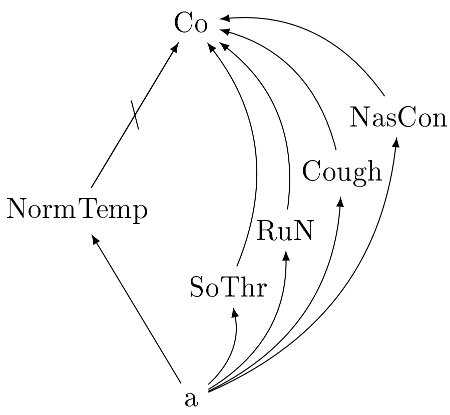

Figure 3: Γ_3

The second approach is an elaboration on the first. It was first proposed in (Simonet, [14]), and we will refer to it as the 'skeptical approach' here.[4] According to it, the *real* expansion of a network is acquired after intersecting all of its credulous expansions. The rationale behind this is that the ultimate *skeptical expansion* should contain only those argument paths which are not exposed to the peril of counter-arguments. Γ_2 would have only one skeptical expansion, and it would contain neither $(n, Q, P, +)$ nor $(n, R, P, -)$. Therefore, on this approach, the reasoner would have to remain *skeptical* with respect to Nixon's being or not being a pacifist.

Each of the two approach has its own advantages, but some cases indicate that both (as well as others not presented here) are substantially lacking. In order to see this consider the network Γ_3 that is shown in Figure 3. Its interpretation is $a =$ Alice, $SoThr =$ a person with a sore throat, $RuN =$ a person with a running nose, $Cough =$ a person with a cough, $NasCon =$ a person with a nasal congestion, $NormTerm =$ a person with normal body temperature, and $Co =$ a person who has a cold. Γ_3 seems to be a perfectly reasonable representation of a real-life situation. Sore throat, running nose, cough, nasal congestion, and fever are the usual symptoms of the infection known as the 'common cold'. That is to say, each of these nuisances most often is caused exactly by cold. Often enough, though, people do not develop all of the symptoms. In the particular situation we model, the individual called Alice has all the symptoms of cold but fever. Now, I argue that, despite her normal temperature, in this case it is natural to suppose that Alice has a cold.

However, let us have a look at what the above approaches say concerning Γ_3. On the credulous one, it has two equally significant expan-

[4] It should be noted though that the literature contains at least two other path-based frameworks that are also dubbed 'skeptical'.

sions. One of them contains the paths $(a, SoThr, Co, +)$, $(a, RuN, Co, +)$, $(a, Cough, Co, +)$, and $(a, NasCon, Co, +)$; and thus supports $isa(a, Co, +)$. The other contains $(a, NormTemp, Co, -)$ and supports $isa(a, Co, -)$. Of course, the first expansion is the intuitively correct one, and we must grant it that the credulous approach is able to identify it. Nevertheless, it does not provide any means for comparing different expansions, and, therefore, justifies the reasoner to infer $isa(a, Co, -)$ as much as it justifies the inference of $isa(a, Co, +)$. The skeptical approach fairs worse. For, analogously to the Nixon case, it invites the reasoner to remain skeptical about Alice's having or not having a cold. In general, on this approach, any number of argument paths can be outweighed by a single argument to the opposite. Apparently, this undercuts all of its chances to give an intuitively desirable conclusion in such cases as Γ_3.

Let us reconsider the two credulous expansions of Γ_3. Of course, the number of argument paths they contain is what makes them differ, but there is also something more. To see what it is we have to recall McCarthy's abnormality predicate. Intuitively, if the reasoner chooses the expansion that supports $isa(a, Co, +)$, it also has to conclude that Alice must be abnormal with respect to the default rule 'People with normal body temperature usually do not have a cold'. That is to say, her temperature is normal, but she still has a cold. If, on the other, the reasoner chooses the other expansion, it has to infer that Alice is abnormal with respect to a whole bunch of defaults. Now, clearly, were we to use McCarthy's notions here, we would say that the first expansion is the one that minimises (= circumscribes) abnormality. Of course, the abnormality predicate and anything alike circumscription are totally absent from the network-based accounts, but that is exactly what we want to change.

It is important to note that the approaches that rely on inheritance networks have some very nice features when compared to most of the other formalisms that model default reasoning. We mention only some. First, the network-based approaches do not require explicitly listing the exceptions to default rules from the very start. Second, they are pleasing from the perspective of complexity theory. Third, they are quite handy for articulating a variety of different intuitions about default reasoning. Fourth, the inheritance networks themselves provide a nice way of visualising complicated patterns of interaction among defaults. Given all this, augmenting some appropriate network-based account with the ideas from the circumscription approach cannot but seem a very appealing endeavour. For the result would have both all these features and a lot more expressive power. Besides, it would be based on good intuitions, just as the idea of abnormality minimisation is.

In the principal sections of this paper we develop two alternative frameworks that both supplement inheritance networks with the idea of minimising abnormality. The paper is structured as follows. Right after this introductory section we turn to developing a path-based framework along the lines of the two approaches we have already introduced. We take special care to define all the basic concepts of the system and then extend it in order to be able to deal with cases akin to Γ_3. In the two subsequent sections we devise an alternative framework that falls within the tradition pioneered by Erik Sandewall in [13]. Unlike the first approach, it operates on statements, rather than paths; and at its heart lies a set of non-monotonic inference rules. In section (2) we develop the basics: after noting that neither of the two existing approaches based on non-monotonic rules — Sandewall's and that of (Simonet, [15]) — are satisfactory, we present our own set of rules, state some results characterising the system, and show that it corresponds to the (basic) path-based approach presented in section (1). In section (3) we extend the newly developed framework so as to make it able to minimise abnormality. This all is followed by a short concluding section in which we also suggest directions for future research. The paper is accompanied by two appendices: appendix A contains the most important proofs, while in appendix B we discuss the problems with the sets of non-monotonic inference rules put forward in (Simonet, [15]).

1 A Path-Based Approach

We begin this section with a formal presentation of a fairly standard path-based approach. Contentwise it incorporates both path-based approaches informally introduced above. The only non-standard thing about it is our definition of consistency for inheritance nets. As soon as at the basic concepts are in place, we proceed to extend the system. We introduce a new kind of paths, adjust a number of concepts, and define the notion of *minimal expansion*, i.e., expansion that circumscribes abnormality. In the end of the section we return to Γ_3.

An inheritance network Γ is defined as a finite collection of positive and negative links between nodes. $(x, y, +)$ stands for a positive and $(x, y, -)$ for a negative link from some node x to another node y. s is used as a variable ranging over the signs $+$ and $-$, while $-s$ denotes the opposite sign of s. A *positive path* from x_1 to x_n through $x_2, \ldots x_{n-1}$, denoted by $(x_1, \sigma, x_n, +)$ or $(x_1, x_2, \ldots, x_{n-1}, x_n, +)$, is defined as a sequence of direct links $(x_1, x_2, +)$, $(x_2, x_3, +)$, $\ldots$, $(x_{n-1}, x_n, +)$. A *negative path* from x_1 to x_n via $x_2, \ldots x_{n-1}$, denoted by $(x_1, \sigma, x_n, -)$ or $(x_1, x_2, \ldots, x_{n-1}, x_n, -)$, is defined as a sequence of direct links $(x_1, x_2, +)$, $(x_2, x_3, +)$, $\ldots$, $(x_{n-1}, x_n, -)$. Note that a sequence consisting of only one link is a path. We use lower case Greek letters π through τ to denote paths and Φ to stand for an arbitrary set of paths. To a large extent this notation is adopted from (Simonet, [15]).

Following (Horty [6]), a pair consisting of a network Γ and a path set Φ — $\langle \Gamma, \Phi \rangle$ — will be described as an *epistemic context*. Although formally any such pairing counts as a context, we will be interested only in those where the second element is somehow acquired from the first. On the intuitive level, when we say that some reasoning agent's epistemic context is $\langle\Gamma, \Phi\rangle$, we mean that this agent has received Γ as his initial information, and that, after some reasoning about Γ, it has arrived at the set of arguments Φ.

Eventually we want to able to link any network to the total set (sets) of arguments an ideal reasoner might accept on its basis. Now, contexts can be though of as steps by means of which this linking is actually carried out. Here is a sketch of the way it works. For any given net Γ, we begin from a context representing the starting position of the reasoning agent, that is to say, from $\langle\Gamma, \Gamma\rangle$. In the first step we identify some argument π that is *acceptable* in this context, immediately change the actual context of the agent to $\langle\Gamma, \Gamma \cup \{\pi\}\rangle$, and proceed to the second step. In the second step we find another argument that is *acceptable* in $\langle\Gamma, \Gamma \cup \{\pi\}\rangle$, add it to the context, and move to the third step. This is repeated until a context $\langle\Gamma, \Phi\rangle$ is reached such that, for any argument τ that is acceptable in $\langle\Gamma, \Phi\rangle$, we already have $\tau \in \Phi$. In fact, we are most interested in exactly such Φ-s

Of course, we still have to specify when exactly a path is acceptable in a context. This is done with the help of a special relation of *inheritability.* Following a more or less established tradition, we use the symbol $\mid\!\sim$ for it. Thus, $\langle\Gamma, \Phi\rangle \mid\!\sim \pi$ would read as 'the path π is inheritable in the context $\langle\Gamma, \Phi\rangle$'. Inheritability itself will emerge as a combination of the three following notions: *constructibility*, *preclusion*, and *conflict.* Jointly they specify the conditions for an argument path to be acceptable in a given context. We introduce each one of them in turn.

Intuitively, a path is constructible in a certain epistemic context in case it can somehow be pieced together from the paths of Φ and links of Γ. The formal definition runs as follows.

Definition 1 (Constructibility)**.** *A path* (x, σ, y, z, s) *is* constructible *in the context* $\langle\Gamma, \Phi\rangle$ *iff* $(x, \sigma, y, +) \in \Phi$ *and* $(y, z, s) \in \Gamma$.

Constructibility is a necessary, but not a sufficient condition for a path to be acceptable. In some cases a context may provide all the elements that are necessary for *constructing* a new argument path, but we may still be unwilling to accept it. The two other concepts are meant to account for cases of this kind. Both will specify the negative conditions of inheritability.

The rationale behind preclusion (or *preemption*, as it is sometimes called) is this: an agent would not view an argument as persuasive, even if there are reasons for accepting it, in case his epistemic context provides a more specific (and thus better) reason for accepting an argument to the opposite. The literature contains several non-equivalent definitions of preclusion and it can be said that this notion is still a subject of debate. We will simply adopt the most wide-spread one that is known as *off-path preclusion* and is due to Erik Sandewall.

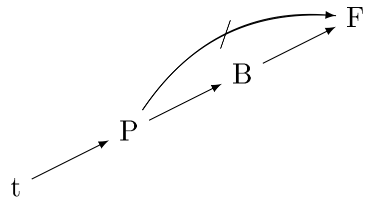

Figure 4: Γ_4, the Tweety Triangle

Definition 2 (Off-path preclusion)**.** *A path* (x, σ, y, z, s) *is* precluded *in the context* $\langle \Gamma, \Phi \rangle$ *iff there is a node* v *such that (i) either* $v = x$ *or there is a path of the form* $(x, \rho, v, \rho', y, +) \in \Phi$*, and (ii)* $(v, z, -s) \in \Gamma$*.*

Let us consider the network Γ_4 depicted in Figure 4 to see preclusion at work. Here t = Tweety, P = penguins, B = birds, and F = flying things. Suppose that some agent has already drawn the conclusion that Tweety is a bird. Thus his present epistemic context is $\langle \Gamma_4, \Phi' \rangle$, where $\Phi' = \Gamma_4 \cup \{(t, P, B, +)\}$. Note that in this context $(t, P, B, F, +)$ and $(t, P, F, -)$ are both constructible. However, since they support contradictory statements ('Tweety flies' and 'Tweety does not fly', respectively), we cannot accept both. Intuitively, though, $(t, P, F, -)$ provides a more specific argument than $(t, P, B, F, +)$ and should therefore be given preference. In fact, the above definition of preclusion achieves exactly that. Φ' contains the path $(t, P, B, +)$ and Γ_4 contains the link $(P, F, -)$, which is enough for the longer path to be precluded.

Now, let us return to the network Γ_2 presented in the introductory section and consider $\langle \Gamma_2, \Gamma_2 \rangle$ as an epistemic context. Just like in the Tweety case, here we have two constructible argument paths supporting opposite conclusions: $(n, Q, P, +)$ and $(n, R, P, -)$. However, preclusion will not help resolving the contradiction here, for neither of the two paths is more specific than the other. Instead it will be done by the concept of conflict, which will allow the reasoner to accept at most one path in this and similar cases.

Definition 3 (Conflict)**.** *A path* (x, σ, y, s) conflicts *with any path of the form* $(x, \rho, y, -s)$*. A path* π *is* conflicted *in the context* $\langle \Gamma, \Phi \rangle$ *iff* Φ *contains a path that conflicts with* π*.*

Note that the definition of conflict does not apply to the above situation. However, the situations changes after one of the argument paths gets accepted. Suppose that the agent has decided to accept $(n, R, P, -)$ and, thus, changed its epistemic context to $\langle \Gamma_2, \Phi' \rangle$, where $\Phi' = \Gamma_2 \cup \{(n, R, P, -)\}$. Consider the remaining path $(n, Q, P, +)$. It clearly is in conflict with $(n, Q, P, -)$, and, since the latter is an element of Φ', $(n, Q, P, +)$ is conflicted in $\langle \Gamma_2, \Phi' \rangle$.

Now we have everything we need to define inheritability.

Definition 4 (Basic Defeasible Inheritability)**.** .

Case 1: π is a direct link. Then $\langle\Gamma, \Phi\rangle \mid\!\sim \pi$ iff $\pi \in \Gamma$.

Case 2: π is a compound path. Then $\langle\Gamma, \Phi\rangle \mid\!\sim \pi$ iff

1. *π is constructible in $\langle\Gamma, \Phi\rangle$,*
2. *π is not conflicted in $\langle\Gamma, \Phi\rangle$,*
3. *π is not precluded in $\langle\Gamma, \Phi\rangle$.*

With the inheritability relation at our disposal, we can finally state our definitions for *expansions*. We let the *credulous expansions* — that correspond to expansions of (Touretzky, [17]) — be exactly the fixed points of the inheritability relation. This effectively captures just those path sets that contain neither too few, nor too many arguments.

Definition 5 (Credulous expansion)**.** *The path set Φ is a* credulous expansion *of the net Γ iff $\Phi = \{\pi : \langle\Gamma, \Phi\rangle \mid\!\sim \pi\}$.*[5]

Notice that a net can have more than one credulous expansions. Thus the above Γ_2 has two: $\Gamma_2 \cup \{(n, Q, P, +)\}$ and $\Gamma_2 \cup \{(n, R, P, -)\}$. Much unlike this, any network will be allowed to have at most one *skeptical expansion*. This notion corresponds to the idea of expansion presented in (Simonet, [14]) and is defined as follows:

Definition 6 (Skeptical expansion)**.** *The intersection of all credulous expansions of Γ is called* skeptical expansion *of Γ.*

Recall that the ultimate aim of any network-based account is to map inheritance networks to their corresponding conclusion sets. Thus, we still have to make the last step, i.e., to specify how to translate expansions into sets of `IS-A` statements. The word 'support' has already been used a few times, but we relied on its intuitive meaning. Now it is time to define it more precisely: we say that a path (x, σ, y, s) *supports* the statement $isa(x, y, s)$. The extension of relation of support to path sets and `IS-A` statements is straightforward: a set of paths Φ is said to support a statement set (or *theory*) Δ iff Δ is the set of `IS-A` statements supported by the paths in Φ. Obviously, credulous expansions support *credulous theories*, while the single skeptical expansion supports the single *skeptical theory*.

Before we conclude the exposition of the basic system by stating some results, we still have to define some important notions. The first one is that of an *acyclic network*. Its definition makes use of an auxiliary concept of a

[5]This definition has become fairly standard since (Horty, [6]). It is also used, for instance, in (Dung and Son, [4]) and in (Simonet, [15]).

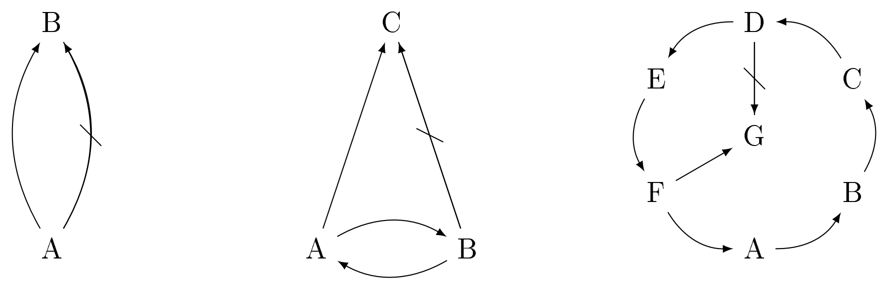

Figure 5: Examples of inconsistent networks

generalised path. The latter is a link sequence like a normal path, except that it can contain negative links anywhere. Formally, we say that a generalised path is a sequence of direct links (x_1, x_2, s), (x_2, x_3, s), $\ldots$, (x_{n-1}, x_n, s). Then a network is said to be *acyclic* if it does not contain a generalised path (x_1, σ, x_n, s) with $x_1 = x_n$. The other two notions we need are those of *network* and *expansion consistency.* Usually the definition of network consistency goes along the lines of 'a network is consistent if there are no two links of the form $(x, y, +)$ and $(x, y, -)$ in it'. Now, our definition of consistency is a generalisation of the standard one:

Definition 7 (Network consistency)**.** *A network* Γ *is* consistent *if* Γ *does not contain two links of the form* $(x_i, y, +)$ *and* $(x_j, y, -)$ *such that* (i) *either* $x_i = x_j$*, or* (ii) *there is a positive path of the form* $(x_1, x_2, \sigma, x_n, x_1, +)$ *in* Γ *with* $1 \leq i \leq j \leq n$*.*

Some illustrative examples of inconsistent networks are given in Figure 5. The main reason for adopting this more stringent definition results from the intuition about *positive cycles.* Now, a positive cycle — i.e., a positive path of the form $(x_1, x_2, \sigma, x_n, x_1, +)$ — is akin to a description of an if and only if relation between x_1, x_2, etc. (cf. (Wang *et al.*, [20])). Intuitively, if some two nodes x_i and x_j are connected by a positive cycle, all the information that is specified about x_i is just as well stated about x_j, and vice versa. Given this, it only seems natural to extend the usual prohibition on the links of the form $(x, y, +)$ and $(x, y, -)$ in the way our definition does. By the way, a number of notorious examples of cyclic nets — nets that have no expansions (see, for instance, (Horty, [6][p. 15])) — come out inconsistent on this definition.

In contrast to consistency for nets, our notion of consistency for expansions is standard. Thus we say that an expansion is consistent, if it does not contain two paths of the form (x, σ, y, s) and $(x, \rho, y, -s)$. Now, having all the concepts in place, we can state the promised results. Both are due to (Touretzky, [17]).

Proposition 1. *Every consistent acyclic network has a credulous expansion.*

Proposition 2. *An expansion* Φ *of a net* Γ *is consistent iff* Γ *is.*[6]

Recall that, ultimately, we are aiming at a system that would allow us to say explicitly that a certain object (or kind) is abnormal with respect to a certain default statement. In its present state, though, our path-based framework can hardly do that, for it can only tell what `IS-A` statements one can derive on the basis of a given network. If we want to be able to say something about abnormalities, our system has to be extended.

Let us return to the net Γ_4 (the Tweety Triangle) presented above. In principle, on its basis, we would want to be able to derive not only the standard `IS-A` statements — 'It is natural to suppose that Tweety is a Bird', etc. —, but also the following one:

Tweety is abnormal with respect to [the default] 'Birds usually Fly'.

Thus, we certainly have to allow for a new kind of statements that can enter conclusion sets. We shall write them as $abnorm(x, y, z, s)$. Here the intended meaning should be obvious: 'x is abnormal with respect to (y, z, s)'. The above statement, for instance, would be written as $abnorm(t, B, F, +)$ in this new notation.

Now, recall that in path-based approaches theories associated with nets are arrived at via expansions. Let Δ be the theory of Γ_4 and Φ its only credulous expansion.[7] Since we want Δ to contain $abnorm(t, B, F, +)$, we must have something in Φ that would support it. Since $abnorm(t, B, F, +)$ differs from the standard `IS-A`-statements, the argument supporting it should also differ from the standard paths of Φ. Of course, there is more than one way to make this work, but re-defining at least some of the basic concepts is unavoidable. I choose to modify our concept of path, by adding a brand new kind to it. As the reader will see, this will also prompt us to expand other notions. Most importantly, the relations of inheritability and support.

We have defined positive and negative paths as special sequences of links. After our next definition a path no longer needs to be such sequence.

Definition 8 (Abnormality path)**.** *An* abnormality path *(or a* path specifying abnormality*)* π *is a pairing of* (i) *the symbol* $\mathfrak{a}$ *and* (ii) *some positive or negative path* π'*, where* π' *must be compound.*

Intuitively, the abnormality path $\langle \mathfrak{a}, (x, \sigma, y, z, s) \rangle$ encapsulates the information that the object (or kind) x is abnormal with respect to the default statement (y, z, s). The symbol $\mathfrak{a}$ serves as a marker here, which we need to be able to distinguish between the paths specifying abnormality and the

[6] Although Touretzky uses a different — that is to say, standard — notion of network consistency, this result readily translates to our system. The reason is simple: ours is a more restrictive notion.

[7] Notice that the credulous and the skeptical expansions of Γ_4 coincide.

standard ones. All the other information we can read off from the second element of the pairing. We add the condition that it has to be compound to exclude nonsensical expressions. Indeed, a statement saying that x is abnormal with respect to the default statement 'x is a y' implies that 'x is a y' is no default at all. Henceforth, if I say that π is a path without other specifications, π can be a positive, a negative, or an abnormality path. If I say that π is a compound path, π is either positive or negative, but not an abnormality path.

Our next step is defining a new inheritability relation. It is a straightforward extension of the basic relation already defined:

Definition 9 (Extended Defeasible Inheritability). .

Case 1: π is a direct link. Then $\langle\Gamma,\Phi\rangle \mid\!\approx \pi$ iff $\pi \in \Gamma$.

Case 2: π is a compound path. Then $\langle\Gamma,\Phi\rangle \mid\!\approx \pi$ iff

1. *π is constructible in $\langle\Gamma,\Phi\rangle$,*
2. *π is not conflicted in $\langle\Gamma,\Phi\rangle$,*
3. *π is not precluded in $\langle\Gamma,\Phi\rangle$.*

Case 3: π is an abnormality path $\langle\mathfrak{a},\pi'\rangle$. Then $\langle\Gamma,\Phi\rangle \mid\!\approx \pi$ iff

1. *π' is constructible in $\langle\Gamma,\Phi\rangle$,*
2. *π' is either conflicted or precluded in $\langle\Gamma,\Phi\rangle$,*

Thus exactly the paths that are constructible, but still not acceptable in the given context will form the second element of an abnormality pairing. This seems to be both the most natural and the most straightforward way of identifying precisely those cases when one default statement is overrun by another. If some path (x,τ,y,z,s) is constructible, but conflicted in a given context $\langle\Gamma,\Phi\rangle$, this context must already contain an argument to the opposite, i.e., some other path $(x,\sigma,z,-s)$. In such a case we are fully justified to infer that in $\langle\Gamma,\Phi\rangle$ the object (or kind) that is represented by x has to be abnormal with respect to the default (y,z,s). Analogously, if (x,τ,y,z,s) is constructible, but precluded in $\langle\Gamma,\Phi\rangle$, we can be sure that x is abnormal with respect to (y,z,s).

Our next stop are expansions. Here only some minor adjustments are required:

Definition 10 (Extended credulous expansion). *The path set Φ is an* extended credulous expansion *of the net Γ iff $\Phi = \{\pi : \langle\Gamma,\Phi\rangle \mid\!\approx \pi\}$.*

Definition 11 (Extended Skeptical expansion). *The intersection of all extended credulous expansions of Γ is called* extended skeptical expansion *of Γ.*

Now it only remains to extend the notion of support so that it also applies to the paths specifying abnormality. Henceforth a path of the form $\langle\mathfrak{a},(x,\sigma,y,z,+)\rangle$ is said to support a statement of the form $abnorm(x,y,z,+)$; and a path of the form $\langle\mathfrak{a},(x,\sigma,y,z,-)\rangle$ is said to support $abnorm(x,y,z,-)$. To use Γ_4 as an example again, the abnormality path $\langle\mathfrak{a},(t,P,B,F,+)\rangle$ would support the statement we began with, namely, $abnorm(t,B,F,+)$.

The next two propositions relate the expanded system to its more basic counterpart. The proofs are simple enough to be left out.

Proposition 3. *To any credulous expansion* Φ *of a net* Γ *corresponds an extended credulous expansion* Φ' *such that* $\Phi \subseteq \Phi'$*, and vice versa: to any extended credulous expansion* Φ' *of* Γ *corresponds a credulous expansion* Φ *of* Γ *such that* $\Phi \subseteq \Gamma$.[8]

Proposition 4. *Let* Φ *and* Φ' *be the skeptical expansion and the extended skeptical expansion of a given network* Γ*. Then* $\Phi \subseteq \Phi'$*.*

Now, when our path-based approach has been supplemented with abnormality paths together with all the appropriate adjustments, we can single out the expansions that minimise abnormality. That is to say, the expansions that contain as few abnormality paths as possible. To this end we first define a function:

$$\#abnorm(\Phi) = |\{\pi \in \Phi : \pi = \langle\mathfrak{a},\pi'\rangle\}|$$

The output of $\#abnorm$ is the number of abnormality paths a given path set Φ contains. In fact, this is all we need to define the notion of a *minimal expansion*, i.e., the expansion that actually minimises abnormality.

Definition 12 (Minimal expansion)**.** *The minimal expansion of a network* Γ *is an extended credulous expansion* Φ *of* Γ *such that, for no extended credulous expansion* Φ'*, we have* $\#abnorm(\Phi') < \#abnorm(\Phi)$*.*

Notice that, on this definition, any network will have a minimal expansion, unless, of course, it has no expansions at all. Besides, some networks will have more than one minimal expansion. The Nixon Diamond may serve as an example. Both of its extended credulous expansions qualify as minimal. Now, let us return to the inheritance net Γ_3 we considered in the introduction to see whether the notion does the job it should. On the present approach, Γ_3 has two (extended credulous) expansions Φ and Φ'. Here is how they look:

$$\Phi = \Gamma_3 \cup \{(a, SoThr, Co, +), (a, RuN, Co, +), (a, Cough, Co, +),$$
$$(a, NasCon, Co, +), \langle\mathfrak{a}, (a, NormTemp, Co, -)\rangle\}$$

[8] Notice that via this proposition the two results characterising the basic system readily translate to the extended one.

$$\Phi' = \Gamma_3 \cup \{(a, NormTemp, Co, -), \langle \mathfrak{a}, (a, SoThr, Co, +)\rangle,$$
$$\langle \mathfrak{a}, (a, RuN, Co, +)\rangle, \langle \mathfrak{a}, (a, Cough, Co, +)\rangle, \langle \mathfrak{a}, (a, NasCon, Co, +)\rangle\}$$

Since Φ and Φ' are the only expansions of Γ_3 and $\#abnorm(\Phi) < \#abnorm(\Phi')$, Φ satisfies the above definition. Note that it is the credulous expansion we identified as the intuitively correct one for Γ_3, with the only difference that it also contains an abnormality path. The two non-trivial statements that enter the theory Φ supports — $isa(a, Co, +)$ and $abnorm(a, NormTemp, Co, -)$ — are also just what we wanted to be able to derive on the basis of Γ_3. Thus, at least in this case, minimal expansion gives the correct result. Our conjecture (and a working hypothesis) is this: if Δ is a theory supported by a minimal expansion of some given net Γ, then Δ is exactly its consequence set.

Now, in spite of the fact that our path-based approach seems to produce the desired results, we leave it here. Our reasons for choosing to develop yet another alternative theory, instead of being content with this one alone, have to do with the set-up of path-based frameworks as such. First, path-based approaches are very remote from ordinary logic, which makes them quite inconvenient when it comes to extensions. Second, path-based approaches do poorly as soon as we have to deal with nets that contain cycles, which are indispensible for representing certain situations (see (Wang *et al.*, [20][pp. 156-7])). For instance, even the simplest cyclic net consisting of two links $(A, B, +)$ and $(B, A, +)$ has an infinite expansion. Given this, it seems only reasonable to avoid being overly focused on path-based theories.

2 Non-Monotonic Inference Rules: the Basic System

In this section we develop an approach to default reasoning using the framework that was pioneered by Erik Sandewall in [13]. It is a network-, but not a path-based approach. Still it has most of the advantages of the path-based theories, and also some of its own. In particular, it is much better for dealing with nets that contain cycles and also easier and more natural to extend in the way we want, i.e., to make it able to express facts about abnormality. We begin by introducing Sandewall's language. Then we outline his framework and argue that it is not satisfactory as it stands. After we develop a new notion of preclusion and extend the original language. Then we present a set of non-monotonic inference rules — the core of the approach —, comment on it, and discuss an illustrative example. We conclude the section by stating some results that characterise the system and relate it to the path-based approach developed above.

The heart of Sandewall's approach is a set of non-monotonic inference rules. These rules, however, do not operate on paths and links, but rather on sets of statements. Sandewall allows for statements that are either atoms or negations of atoms, and every atom has to be of one of the following forms:

$isax(x, y, s)$,

$isa(x, y, s)$,

$precl(x, y, z, s)$.[9]

We have already encountered the `IS-A` statements. Therefore we know that the intended meaning of $isa(x, y, s)$ is 'it is natural to suppose that x is [not] a y (/x-s are [not] y-s)'. Now, statements of the *isa*-form play roughly the same role in Sandewall's framework as paths play in path-based approaches. The *isax*-statements, on the other hand, are more like links. In fact, any statement of the form $isax(x, y, s)$ has the same meaning as (x, y, s). Thus, if x is an individual object, $isax(x, y, s)$ is a first-order statement of the form 'x is [not] a y'; if x is a kind, it is a default statement. Finally, the meaning of the *precl*-statements is close to the notion of *preclusion* discussed above, just as the name suggests. We will read $precl(x, y, z, s)$ as 'the default (y, z, s) is precluded for x'.

The purpose of the non-monotonic inference rules is to determine the *extension* (or *extensions*)[10] of a given inheritance network. Extensions are special sets of statements that are analogous to credulous expansions in our path-based approach. Sandewall's rules have the following general form: *If D_1 is in the set and D_2 is not, then infer D_3.* D_1, D_2, and D_3 stand for sets of statements. Note also that the part of the rule that is related to D_2 may be left out. Extensions of a given inheritance net are determined by constructing a sequence of increasing sets of statements, $E_0, E_1, \ldots, E_i, \ldots$, where:

E_0 is the initial set of statements $isax(x, y, s)$ standing for the links in the net, and

each E_{i+1} is acquired from E_i by instantiating one of the inference rules, D_1 has to be a subset of E_i, D_2 disjoint from E_i, and then E_{i+1} is set to be $E_i \cup D_3$.

This process is continued to its (possibly infinite) limit. Such a limit E is an extension, if it satisfies two conditions: (i) E is consistent, i.e., it does not contain two statements of the form p and $\neg p$ for some p, and (ii) E is a fixed point for the given set of rules, i.e., for any rule applicable to E, D_3 is a subset of E. We should add that applying the rules in different order may

[9] In fact, Sandewall also allows for statements of the form $cntr(x, y, z, s)$, but we will not use them. The purpose they served can just as well be served by the *isa*-statements.

[10] My notion of *extension* corresponds to Sandewall's notion of *consistent extension*.

result in several different extensions. Notice that our notion of consistency for extensions differs from the notions of consistency for expansions. Thus, an extension could, at least in principle, contain two statements of the form $isax(x, y, +)$ and $isax(x, y, -)$ (or $isa(x, y, +)$ and $isa(x, y, -)$). However, this cannot happen as long as one deals with consistent inheritance nets, and, as a matter of fact, we restrict our attention to nets of such kind in what follows.

Unfortunately, the set of six inference rules originally proposed in (Sandewall, [13]) turned out not to be flawless. Most importantly, it gives counter-intuitive results on some fairly simple examples that are well-known in the literature. (Simonet, [15]) — the only serious follow-up to Sandewall's framework — contains a detailed discussion of the problematic cases along with a proposal for revision of the rules. However, Geneviève Simonet goes beyond a mere revision in this paper. She also presents an alternative set (or rather sets) of non-monotonic inference rules drawing their motivation directly from some path-based approaches. Her basic idea is to take a certain credulous path-based theory (which is very much alike the approach we presented in the previous section), carry it over to Sandewall's setting, and give a correspondence proof between the two.[11] Unfortunately, the way she implemented this idea is not satisfactory. First, just like the Sandewall rules, those of Simonet produce counter-intuitive results. Second and even more importantly, in some concrete cases the results given by the rules differ from those suggested by the underlying path-based approach, which cannot but cast doubts on Simonet's correspondence proof. The discussion of the issues with [15] is left for the B appendix of the present paper. Here I at once turn to developing my own set of non-monotonic inference rules.

Let us first step back and compare the language of the Sandewall framework — henceforth, the isa-framework — and that of our path-based approach. The latter, just as any other path-based approach, is very sensitive with respect to distinguishing paths. Thus, despite the fact that $(a, \sigma, B, \sigma', D, +)$ and $(a, \sigma, C, \sigma', D, +)$ differ in one node only, they already count as different paths. The language of the isa-framework, on the other hand, is much less sensitive. $(x, \sigma, y, +)$ and $isa(x, y, +)$ may both carry the information that x is a y, but there is a big difference. The path $(x, \sigma, y, +)$ specifies what links (defaults) ascribing y to x depends on, while $isa(x, y, +)$ gives no such information. In fact, in the isa-framework we are simply not able to say anything like 'x inherits from y via z'. In order to see why we may want to be able to say such things, we have to recall our definition of preclusion.

Definition 2 (Off-path preclusion). *A path* (x, σ, y, z, s) *is* precluded *in the context* $\langle \Gamma, \Phi \rangle$ *iff there is a node* v *such that* (*i*) *either* $v = x$ *or there is a*

[11] In fact, it would be more precise to talk of three different credulous theories here, for Simonet works with three non-equivalent definitions of preclusion at the same time.

path of the form $(x, \rho, v, \rho', y, +) \in \Phi$*, and (ii)* $(v, z, -s) \in \Gamma$.

The cases when a path of the form $(x, \rho, v, \rho', y, +)$ is in Φ are the ones we are most interested in. Notice that here we have to have a path in Φ that *runs through a certain node.* Now, if we want to be able to transfer this notion of preclusion to the *isa*-framework, we better find way of capturing this in its language. We will do just that, but in a roundabout way. We will first define a different notion of preclusion that is easier to implement in the *isa*-framework and prove that it is equivalent to the standard off-path preclusion.

The new definition of preclusion will rely on the notion of an *intermediary.*[12] Of course, we use the terminology of the path-based accounts here:

Definition 13 (Intermediary). *The node* v *is an* intermediary *to the path* $(x_1, \ldots, x_n, s)$ *in the path set* Φ *if* $v = x_i$ *for some* $1 \leq i < n$*, or else* Φ *contains a path of the form* $(x_1, \ldots, x_k, y_1, \ldots, y_l, x_m, +)$ *where* $1 \leq k < m < n$ *and* $v = y_j$ *for some* $1 \leq j \leq l$*. See Figure 6.*

With the intermediaries in place we can easily define preclusion. I have chosen to call it i-preclusion, 'i' standing for intermediaries.

Definition 14 (i-preclusion). *A path* $(x_1, \sigma, x_{n-1}, x_n, s)$ *is precluded in the context* $\langle \Gamma, \Phi \rangle$ *iff there is a node* v *such that (i)* v *is an intermediary to the path* $(x_1, \sigma, x_{n-1}, +)$ *in the path set* Φ*, and (ii)* $(v, x_n, -s) \in \Gamma$.

Our next lemma shows that the two definitions of preclusion are equivalent.

Lemma 1 (Preclusion lemma). *A path* π *is off-path precluded iff it is* i-*precluded.*

Proof. $\Rightarrow$ Suppose that some path $\pi = (x_1, \sigma, x_{n-1}, x_n, s)$ is off-path precluded in the context $\langle \Gamma, \Phi \rangle$. Thus there must be a path of the form $(x_1, \rho, v, \rho', x_{n-1}, +)$ in Φ and a link of the form $(v, x_n, -s)$ in Γ. A minute reflection reveals that v satisfies the condition of being an intermediary to π in the path set Φ. Since we also have $(v, x_n, -s) \in \Gamma$, π is i-precluded.

$\Leftarrow$ Suppose that $\pi = (x_1, \sigma, x_{n-1}, x_n, s)$ is i-precluded in the context $\langle \Gamma, \Phi \rangle$. Thus there is a node v such that v is an intermediary to $(x_1, \sigma, x_{n-1}, +)$ and $(v, x_n, -s) \in \Gamma$. Now either v is a node of $(x_1, \sigma, x_{n-1}, +)$ itself or Φ contains a path of the form $(x_1, \rho, v, \rho', x_{n-1}, +)$. In both cases π is off-path precluded. $\square$

[12] In (Touretzky, [17]) preclusion is also defined using the notion of an intermediary. I have borrowed the name from there, but my definition is different. Consequently, also the notion is, and in two respects. First, it is somewhat weaker than that of Touretzky. Second, I allow for any paths to have intermediaries, while Touretzky has them only for the positive ones.

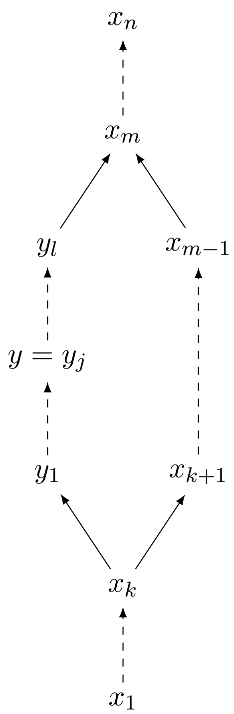

Figure 6: y is an intermediary to the path $(x_1,\ldots,x_k,y_1,\ldots,y_l,x_m,\ldots, x_n, +)$ in Φ, for Φ contains $(x_1,\ldots,x_k,y_1,\ldots,y_l, x_m,+)$ and $y = y_i$ for some j, $1 \leq j \leq l$, $1 \leq k < m < n$.

Now we extend Sandewall's original language by allowing for atomic propositions of the following form: $interm(x, y, z, s)$. One could think of $interm(x, y, z, s)$ as saying that node y is an intermediary to any path of polarity s that begins with x and ends with z. However, since there are no paths in the *isa*-framework, this cannot be entirely precise. A correct though somewhat cumbersome reading of $interm(x, y, z, s)$ would go along the lines of 'In case y and z justify inferring opposite statements about x, y must be given priority'. Be it as it may, *interm*-statements will tell us effectively when defaults turn out precluded.

It is finally time to state the inference rules.

Set R of inference rules

1. If $isax(x, y, s)$ is in E

   then add $isa(x, y, s)$ and $interm(x, x, y, s)$ to E.

2. If $isa(x, y, +)$, $isax(y, z, s)$, $interm(x, v, y, +)$, and $isax(v, z, -s)$ are in E,

   then add $precl(x, y, z, s)$ to E.

3. If $isa(x, y, +)$ and $isax(y, z, s)$ are in E,

   and $precl(x, y, z, s)$ and $isa(x, z, -s)$ are not in E,

   then add $isa(x, z, s)$, $\neg precl(x, y, z, s)$, and $interm(x, y, z, s)$ to E.

4. If $interm(x, y, z, +)$ and $interm(x, z, w, s)$ are in E,

   then add $interm(x, y, w, s)$ to E.

Now each rule requires a little comment. The main aim of the first rule is to translate the *isax*-statements into the corresponding statements of the *isa*-form. Intuitively, if we know $isax(x, y, s)$, $isa(x, y, s)$ is the least we should be able to derive. This rule is carried over from (Sandewall, [13]) with the only difference that we add the statement $interm(x, x, y, s)$ to E as well. We need it (just as any other *interm*-statement) to be able to deal with preclusion, in case it might occur: if x and y will justify inferring opposite statements about x — for instance, via two defaults $(x, w, +)$ and $(y, w, -)$ —, $interm(x, x, y, s)$ will tell us that x should be given priority.

The second rule is the one that takes care of preclusion. Since we do it by means of the *interm*-statements, it is quite unlike the rules that account for preclusion in both (Sandewall, [13]) and (Simonet, [15]). Still it has the advantage of being simpler. Let us consider a concrete situation when this rule can be applied: suppose we have $isa(a, B, +)$, $isax(B, C, +)$, $interm(a, D, B, +)$, and $isax(D, C, -)$ in our set. Notice what kind of situation these statements describe. We know that it is natural to suppose that a is a B and that 'B-s are usually C-s'. However, we also know that a is a D and that D-s have priority over B-s in case both justify inference of opposite statements about a, which is just what $interm(a, D, B, +)$ tells us. Moreover, it turns out we are dealing with a case of exactly such kind, for we have 'D-s usually are not C-s' as well. Now, in order to actually prioritise $(D, C, -)$ over $(B, C, +)$ *for* a, we preclude $(B, C, +)$ for it. That is to say, we add $precl(a, B, C, +)$ to the statement set.

The third rule accounts for adding new statements of the isa-form to the would-be extension. I would say that it even reads well (at least, when we instantiate s with either $+$ or -): If it is natural to suppose that x is a y and there is a default $(y, z, +)$, and neither is $(y, z, +)$ precluded, nor is it natural to suppose that x is not a z, then we conclude that 'x is a z' is a natural supposition to make. Now for the other two statements that are added to E here. In light of what has been said above, addition of $interm(x, y, z, s)$ should feel only natural. Clearly, we conclude $isa(x, z, s)$ via y here, and want to indicate that, as long as x is concerned, y should be given priority over z. The inclusion of $\neg precl(x, y, z, s)$ is a delicate matter. In short, it is our (and also Sandewall's) way of enforcing a certain order on rule application. By adding $\neg precl(x, y, z, s)$ here, we arrange invalidation of those would-be extensions in which the rules from R are instantiated in an incorrect order. That is to say, when first $isa(x, z, s)$ is derived via (y, z, s), and only then $precl(z, y, z, s)$ is inferred. The way it actually works will become clear as soon as we will have considered a concrete example. To a great extent, the third rule is a rule of Simonet. The only adjustment I have made is that we add $interm(x, y, z, s)$ instead of $\neg isa(x, z, -s)$.

Finally, the fourth rule is meant to add those $interm$-statements that cannot be added to the statement set by either rule (1) or rule (3), but should nevertheless be there. It is important to see in what kind of situations it is applied. The second positive condition for its instantiation — $interm(x, z, w, s)$ — will always be added to the set by means of rule (3). For the sake of simplicity, suppose that it happened in the previous step, i.e., that $isa(x, w, s)$ and $interm(x, z, w, s)$ have just been added to the set on the basis of $isa(x, z, +)$ and $isax(z, w, s)$. Notice that $interm(x, z, w, s)$ tells us that z must be given priority over w, and that, at this point, there is nothing else we know about w's relation to other nodes. However, the derivation of $isa(x, z, +)$ (and, thus, also of $isa(x, w, s)$) must have relied on other defaults — $(x, x_1, +)$, $(x_1, x_2, +)$, $\ldots$, $(x_n, z, +)$ —,[13] and in this situation x, x_1, x_2, $\ldots$, x_n should all be given priority over w. Now, in fact, (4) does exactly that for us. Intuitively, what it says is this: whichever node has priority over z should have priority over w as well. Thus, in a sense, rule (4) enforces transitivity on the relation described by $interm$-statements.

We now turn to an example in order to clarify the functions of our four rules. Let us consider Γ_4 — the Tweety Triangle — again. Clearly, E_0 here

[13]One at the very least.

must be $\{isax(t,P,+)$, $isax(P,B,+)$, $isax(B,F,+)$, $isax(P,F,-)\}$. This is how we construct its extension:

- rule (1) on $isax(t,P,+)$:
  $E_1 = E_0 \cup \{isa(t,P,+), interm(t,t,P,+)\}$
- rule (3) on $isa(t,P,+)$ and $isax(P,B,+)$: $E_2 =$
  $E_1 \cup \{isa(t,B,+),\ \neg precl(t,P,B,+),\ interm(t,P,B,+)\}$
- rule (4) on $interm(t,t,P,+)$ and $interm(t,P,B,+)$:
  $E_3 = E_2 \cup \{interm(t,t,B,+)\}$
- rule (2) on $isa(t,B,+)$, $isax(B,F,+)$, $interm(t,P,B,+)$, $isax(P,F,-)$:
  $E_4 = E_3 \cup \{precl(t,B,F,-)\}$

Notice that now our would-be extension contains $precl(t,B,F,-)$, which says that the default $(B,F,+)$ is precluded for t. As soons as we recall the negative conditions for instantiating rule (3), we see that the presence of this statement effectively blocks its application on $isa(t,B,+)$ and $isa(B,F,-)$. We need another five steps, however, to arrive at the full extension of Γ_4:

$$E = E_4 \cup \{isa(t,F,-), \neg precl(t,B,F,-),\\ interm(t,P,F,-),\ interm(t,t,F,-),\ isa(P,F,-),\ interm(P,P,F,-),\\ isa(P,B,+),\ interm(P,P,B,+),\ precl(P,B,F,+)\}$$

It is plain that E is consistent and it is not difficult to verify that it is a fixed point. We can still apply inference rules on it, but none of them will make it grow larger. Hence, E is an extension of Γ_4.

At this point the reader may wonder what would have happened if in the fourth step, instead of applying rule (2), we would first have applied rule (3). In short, this would result in a statement set that will not and that we do not want to grow into an extension of Γ_4. But let us see what actually happens. After instantiating (3) we would have acquired a set of statements $E_4' = E_3 \cup \{isa(t,F,+), \neg precl(t,B,F,+), interm(t,B,F,+)\}$. Now, no matter in which order we apply the rules afterwards, at one point we still have to apply the second one (preclusion) on $isa(t,B,+)$, $isax(B,F,+)$, $interm(t,P,B,+)$, and $isax(P,F,-)$. For these statements are present in the set and rule (2) has no negative conditions. Instantiating (2) would make us add $precl(t,B,F,+)$ to E_4' (or any of its possible supersets) and immediately render it inconsistent. Notice that inconsistency results due to the presence of $\neg precl(t,B,F,+)$, which was added to the set along with $isa(t,F,+)$ when applying rule (3). Thus, $\neg precl(t,B,F,+)$ functions as a 'mine' here, and its purpose is an eventual invalidation of E_4'.

Recall that the ultimate purpose of any network-based account is to map any net Γ to its corresponding conclusion set (or sets). In our path-based account we first constructed expansions (sets of paths) from nets, and then mapped them into sets of `IS-A` statements by means of the relation of support. Now, since extensions are sets of statements already, there is no need for anything akin to support. We can get the conclusion sets by simply taking subsets of extensions. That is to say, if E is an extension of Γ, then $\{\varphi \in E : \varphi = isa(x, y, s)\}$ must be its conclusion set.

Now, after all the constituent parts of our *isa*-system have been presented, we can state some results characterising it. We begin by relating it to the (basic) path-based account developed in the previous section. As it turns out, any extension E of a given net Γ has a corresponding credulous expansion Φ of Γ, and vice versa. We say that an extension E corresponds to an expansion Φ if the conclusion set acquired from E is exactly the set of `IS-A` statements supported by Φ. The first direction — extensions to expansions — will emerge as a corollary of our first (more general) theorem. We will, however, also need the following lemma.

Lemma 2 (Ordering lemma). *Any extensions E of Γ can be obtained with the following decreasing order of priority on the choice of instantiation of rules from R: (1) the first rule; (2) the fourth rule; (3) the second rule; (4) the third rule.*

Proof. See the appendix. □

Now comes the theorem:

Theorem 1. *For any consistent statement set E that is obtained by the set R of inference rules from a consistent inheritance net Γ (with priority ordering as specified in the Ordering lemma) there is a corresponding consistent path set Φ such that:*

(*i*) *for any $\pi \in \Phi$, $\langle \Gamma, \Phi \rangle \mid\!\sim \pi$, and*

(*ii*) *for any isa-statement, $isa(x, z, s) \in E$ iff for some path $\sigma = (x, \sigma', z, s)$, $\sigma \in \Phi$.*

Proof. See the appendix. □

Together the theorem and the lemma give us the desired result:

Corollary 1. *For any extension E of a consistent inheritance network Γ there is a corresponding credulous expansion Φ of Γ such that $isa(x, z, s) \in E$ iff for some path $\sigma = (x, \sigma', z, s)$, $\sigma \in \Phi$, for any isa-statement.*

Our next theorem shows that the other direction — expansions to extensions — holds just as well.

Theorem 2. *For any expansion Φ of a consistent network Γ there exists a corresponding extension E such that, for any path $\pi = (x, \pi', z, s)$, $\pi \in \Phi$ iff $isa(x, z, s) \in E$.*

Proof. See the appendix. $\square$

In [12] Erik Sandewall proved some general results for systems of non-monotonic inference rules. These results are directly applicable to our *isa*-framework. Thus, we have the two following propositions:

Proposition 5. *Let E and E' be two extensions of a network Γ, for which $E \subseteq E'$. Then $E = E'$*

Proposition 6. *Every union of distinct extensions of the same network Γ is inconsistent.*

This concludes exposition of the basic *isa*-framework.

3 Non-Monotonic Inference Rules: Adding Abnormality

The main aim of this section is to extend the *isa*-approach presented above so that we could circumscribe abnormality in it. We begin by extending the *isa*-language. Immediately afterwards we reformulate the inference rules of R; each change is accompanied by a detailed commentary. After we introduce the problem of *decoupling* and show how to deal with it in our approach. Then we define the notion of a *minimal extension* — extension that minimises abnormality —, relate the extended *isa*-system to its basic counterpart, and consider some important examples. In the end of the section we discuss some of the ideas introduced in (Touretzky *et al.*, [18]), and examine our system from their perspective.

Our first step is to enrich the *isa*-language. Eventually we want to be able to compare distinct extensions of the same net by the number of abnormalities each of them implies. Thus, a change that would allow extensions to contain statements expressing abnormalities is only natural here. We have already introduced statements of the *abnorm*-form when extending our path-based approach in the first section, and they will do just as well for our *isa*-framework. Recall that the intended meaning of $abnorm(x, y, z, s)$ is 'x is abnormal with respect to (y, z, s)'.

Having the *abnorm*-statements at our disposal, we can reformulate the non-monotonic inference rules. Our changes are of two kinds: (i) we make some slight adjustments to R, and (ii) we add new inference rules. Here is how our four rules look like after the adjustments (the changes are underlined):

Set R' of inference rules

1. If $isax(x,y,s)$ is in E

   then add $isa(x,y,s)$ and $interm(x,x,y,s)$ to E.

2. If $isa(x,y,+)$, $isax(y,z,s)$, $interm(x,v,y,+)$, and $isax(v,z,-s)$ are in E,

   then add $precl(x,y,z,s)$ $\underline{\text{and } abnorm(x,y,z,s)}$ to E.

3. If $isa(x,y,+)$ and $isax(y,z,s)$ are in E,

   and $precl(x,y,z,s)$ and $isa(x,z,-s)$ are not in E,

   then add $isa(x,z,s)$, $\underline{\neg abnorm(x,y,z,s)}$, and $interm(x,y,z,s)$ to E

4. If $interm(x,y,z,+)$ and $interm(x,z,w,s)$ are in E,

   then add $interm(x,y,w,s)$ to E.

Notice that rules (1) and (2) are left unchanged. A minute reflection should convince the reader that there was nothing to change about them. For one simply translates the *isax*-statements into those of the *isa*-form, while the sole concern of the other are intermediaries. The single change made in rule (2) is that now, after it is applied, we do not add a *precl*-statement only, but also one of the *abnorm*-form. The motivation here should be apparent: if we conclude that a certain default is precluded for some object, this object has to be abnormal with respect to it.

The adjustment made to the third rule is this. Previously, after applying it, a negation of a *precl*-formula was added to the would-be extensions; now we add a negation of an *abnorm*-formula instead. Recall that in the basic system the only statements that could get negated were the *precl*-ones and also that they served a concrete aim, i.e., to invalidate certain unwanted would-be extensions. Now, in the extended system negations will be put to the same use. However, there will also be more types of extensions we will want to invalidate, and for each type we will use a different negation (i.e., a negation of a different kind of atom). This will help us effectively distinguish between the various types of unwanted extensions. Negations of the *abnorm*-statements will do just what $\neg precl$-formulae did in our basic system, while the latter will henceforth have a slightly different function. Independently of any purposes negations may serve, having $\neg abnorm(x,y,z,s)$ added to the would-be extensions after an instantiation of rule (3) is only natural. For x has to be normal (= not abnormal) with respect to (y,z,s) if we are to be justified to infer $isa(x,z,s)$ on its basis.

This much for the changes in the existing rules. We still need a new rule, if we want the extensions to contain all the *abnorm*-formulae they should. The second rule adds an *abnorm*-statement in cases of preclusion, but those that have a Nixon Diamond-like shape still have to be taken care of. Our fifth rule serves exactly this purpose.

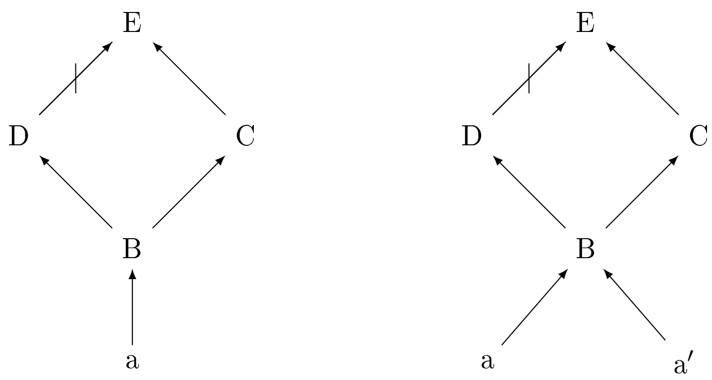

Figure 7: Decoupling, Γ_5 and Γ_6

5. If $isa(x, y, +)$, $isax(y, z, s)$, and $isa(x, z, -s)$ are in E,
   and $precl(x, y, z, s)$ is not in E,
   then add $abnorm(x, y, z, s)$, $\neg precl(x, y, z, s)$, and $\neg isa(x, z, s)$ to E.

Notice that the conditions for instantiating this rule are nothing but a description of a Nixon Diamond-like situation. Let us suppose that s stands for $+$ for the sake of simplicity. Now, we have a statement saying that it is natural to suppose that x is a y and a default (y, z, s) which is known not to be precluded for x. However, there also is another statement specifying that, in fact, it is natural to suppose that x is not a z. Clearly, in such a situation x must be abnormal with respect to (y, z, s). The other two negated statements — $\neg precl(x, y, z, s)$ and $\neg isa(x, z, s)$ — are added for technical purposes. The aim of the first is to invalidate those would-be extensions where this new rule is instantiated before a statement of the *precl*-form is derived by rule (2). The reason for adding $\neg isa(x, z, s)$ as well can only be made clear after we discuss *decoupling*.

The problem of decoupling is well-known in the literature on inheritance networks and is best made clear by means of a concrete example. Consider the net Γ_5 depicted in Figure 7. In total, it has four extensions, but two of them are somewhat problematic. In particular, the extension E' contains, among others, the following two statements: $isa(a, E, +)$ and $isa(B, E, -)$. Apparently, the shape of the network itself suggests us that everything that we can conclude about a must depend on the only piece of factual information we are given, i.e., that a is a B. The problem here is this: in spite of this fact, one of the conclusion sets associated with Γ_5 contains both 'B-s are usually not E-s' and 'a is an E'. It definitely seems that in case all we know about a is that it is a B and 'B-s are usually not E-s', we would not conclude that a is an E. The other net depicted in Figure 7 — Γ_6 — illustrates a slightly different, but a closely related problem. One of the extensions of Γ_6, let us call it E'', contains both the statement $isa(a, E, +)$ and the statement $isa(a', E, -)$. What is odd here is that the net gives us exactly the same factual information about a and a', but we are still allowed to conclude that one is an E, while the other is not.

The name of the problem — 'decoupling' — comes from the path-based approaches. There it is even easy to define in a formal fashion. Thus, we could say that a path (x, y, σ, z, s) is *decoupled* in a path set Φ in case $(y, \sigma, z, s) \notin \Phi$ (cf. (Horty, [6][pp. 42–43])). Now, it may seem to be easy to transfer this definition from paths to the isa-statements. For instance, along some such lines as: an isa-statement $isa(x, z, s)$ is decoupled in an extension E if E contains $isax(x, y, +)$, but not $isa(y, z, s)$. We should not forget, however, that the isa-language is somewhat more limited than that of paths. In particular, there is no way to specify that the inference of $isa(x, z, s)$ (in the definition) has to depend on $isa(x, y, +)$.[14] It might still be possible to devise a proper definition of a decoupled isa-statement, but my intention is to deal with the problem without it. Returning to rule (5), I can say that one of the reasons for adding a formula of the $\neg isa(x, z, s)$-form to the would-be extension is to tag the Nixon Diamond-like situations, which are the only ones subject to decoupling. The other reason for having them is to invalidate those extensions in which decoupling actually occurs. The invalidation itself, though, is executed via the following sixth rule of inference:

6. If $isax(x, y, +)$, $isa(x, z, s)$, and $\neg isa(y, z, s)$ are in E,
   then add $\neg isax(x, y, +)$ to E.

Notice that this rule is purely destructive: if it is applied, the would-be extension is immediately rendered inconsistent. Here it is important to see what shape do these would-be extensions have: $\neg isa(y, z, s)$ does not only indicate a Nixon Diamond, but also implies the presence of $isa(y, z, -s)$ (recall rule (5)). Thus, what we have here are the three statements $isax(x, y, +)$, $isa(x, z, s)$, and $isa(y, z, -s)$ in a Nixon Diamond-like situation, and I claim that this can only be a case of decoupling. Now, as soon as we add rule (6) to R', the counter-intuitive extensions of the two networks discussed above — Γ_5 and Γ_6 — simply disappear. We leave out the exposition of the way it actually happens.

Now, when all the inference rules are in their place, it is time we show that nothing has been lost in the process of expanding the system, i.e., that the extended system is as powerful as the basic one. Note that we will not be able to link extensions acquired by R to extensions obtained by R', because the basis system allows for decoupling, while the extended system does not.

[14]In fact, the definition I have given fails in very simple cases. It suggests, for instance, that the single extension of the Tweety Triangle contains a decoupled statement.

Still we can show that all the important formulae derivable by R are also derivable by R'.

Proposition 7. *For any atomic formula p, if there is an extension E acquired by the rule set R from an inheritance network Γ such that $p \in E$, then there is an extension E' acquired by the rule set R' from Γ such that $p \in E'$.*

Proof. A straightforward induction on the construction of E. $\square$

Henceforth, we will be concerned only with the expanded system. If I talk of extensions with no further specifications, it is the extensions of the expanded system I have in mind. Likewise, 'the set of rules' will mean R' rather than R in the remainder of the paper, unless explicitly specified otherwise.

Recall our aim: it is to devise a system that is able to distinguish between different extensions by means of the number of abnormalities they imply. Now, the rule set R' works in such a way that every extension already contains statements expressing abnormalities. Thus, for any given net Γ, we only have to count the number of *abnorm*-statements in its extensions and single out the one for which this number is the least. Formally, it is done exactly as we did it for our path-based framework. For each extension E, let the function $\#abnorm(E)$ be defined thus:

$$\#abnorm(E) = |\{\varphi \in E : \varphi = abnorm(x, y, z, s)\}|$$

Now with the help of $\#abnorm$ we can at once define the notion of a *minimal extension*, i.e., the extension in which abnormality is minimised.

Definition 15 (Minimal extension)**.** *The* minimal extension *of a network Γ is an extension E of Γ such that, for no extension E' of Γ, we have $\#abnorm(E') < \#abnorm(E)$.*

This definition is analogous to that of a minimal expansion. Thus, every net mustl have a minimal extension, and there are nets with several. Here let us turn to two illustrative examples to see minimal extensions at work. We are already familiar with the first one.

Recall the inheritance network Γ_3 from the introduction. On the present approach, it has two extensions, E and E'. If we leave out all the statements serving technical purposes — negations and the *interm*-statements — and also the trivial ones — i.e., $isax(x, z, s)$ and $isa(x, z, s)$ such that (x, z, s) is a link of Γ_3 —, the two extensions look as follows:

$$E = \{isa(a, Co, -), abnorm(a, SoThr, Co, +), abnorm(a, RuN, Co, +), abnorm(a, Cough, Co, +), abnorm(a, NasCon, Co, +)\};$$

$$E' = \{isa(a, Co, +), abnorm(a, NormTemp, Co, -)\}.$$

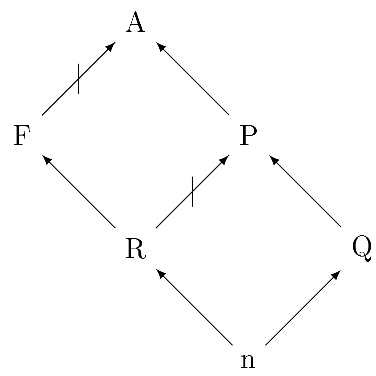

Figure 8: Γ_7, the Zombie Path

Clearly, $\#abnorm(E) = 4$ and $\#abnorm(E') = 1$. Given that E and E' are the only extensions of Γ_3, we see immediately that E' satisfies the condition for being its minimal extension. Again, the minimal extensions also corresponds to the intuitively correct one.

Our second example is provided by network Γ_7 that is depicted in Figure 8. Its standard interpretation is n = Nixon, R = Republican, Q = Quaker, P = pacifist, F = football fan, A = anti-military. It is well known in the literature as an example of a network that contains a 'zombie'. The notion of zombie has been first suggested by Makinson and Schlechta in [8]. In their parlance a zombie is a path that is dead, but can still kill other paths. Using our path-based approach, we could explicate it as follows. Suppose we are interested in the *skeptical* expansion of Γ_7. Recall that it is acquired after intersecting all the credulous expansions of the given net. Now, in spite of the conflict between $(n, R, P, -)$ and $(n, Q, P, +)$ in the lower diamond, one of the credulous expansions, call it Φ, must contain $(n, Q, P, A, +)$. However, exactly because of this conflict $(n, Q, P, A, +)$ is 'dead', for there is no way it could be part of the skeptical expansion. Still it prevents us from adding $(n, R, F, A, -)$ to Φ and, thus, 'kills' $(n, R, F, A, -)$'s chances of being included in the skeptical expansion of Γ_7.

Be it as it may, in our *isa*-framework Γ_7 has three extensions. Again, I focus only on the statements that are relevant:

$$E = \{isa(R, A, -),\ isa(Q, A, +),\ isa(n, F, +),\ isa(n, P, +),$$
$$isa(n, A, +),\ abnorm(n, R, P, -),\ abnorm(n, F, A, -)\}$$

$$E' = \{isa(R, A, -),\ isa(Q, A, +),\ isa(n, F, +),\ isa(n, P, +),$$
$$isa(n, A, -),\ abnorm(n, R, P, -),\ abnorm(n, P, A, +)\}$$

$$E'' = \{isa(R, A, -),\ isa(Q, A, +),\ isa(n, F, +),\ isa(n, P, -),$$

$isa(n, A, -)$, $abnorm(n, Q, P, +)\}$

Since $\#abnorm(E'') < \#abnorm(E) = \#abnorm(E')$, and E, E', E'' are the only extensions of Γ_7, we can conclude that E'' is the minimal one. Here evaluating the intuitiveness of extensions is somewhat more difficult, but E'' still seems to represent a more appealing line of reasoning on the basis of Γ_7 than either E or E'. For resolving the opposition between $isa(n, P, +)$ and $isa(n, P, -)$ in the lower diamond in favour of the latter (as it is done in E'') effectively solves the problem in the upper diamond as well. Besides, in such a way that no additional assumptions about abnormality of n are needed. Thus, exactly E'' seems to the preferable extension.

In the remainder of the section I will discuss some interesting features of minimal extensions. As it turns out, when a net has a unique minimal extension, it can exhibit effects similar to those of some *defeater-defeaters* discussed in [18] by Touretzky, Thomason, and Horty. The reader not interested in this subject matter may at once proceed to the conclusion, for the material presented in the next few pages is supplementary. It does, though, provide some more examples.

The discussion in [18] aims to classify various kinds of argument (= argument path) defeat, and special attention is devoted to what is called *defeat of defeaters*, i.e., situations in which an argument that defeats another is defeated itself. The authors recognise two kinds of defeat: defeat by preclusion and defeat by conflict. Although both are, as their names suggest, closely related to the notions of *preclusion* and *conflict*, respectively, the idea of defeat is somewhat broader and more informal. For instance, in our path-based approach a path can be *conflicted* only with respect to an epistemic context, but we can talk of paths being *defeated by conflict* in expansions, extensions, and inheritance networks as such. The intuitive meaning of the two kinds of defeat is just what one would expect: a path is defeated by preclusion in case it is precluded in the expansion of a given net; a path is defeated by conflict in case it is absent from the expansion due to being in conflict with some other path.

Of course, the notion of defeat is relational. Thus, whenever there is a defeated path, there must just as well be another one — its 'defeater'. If some path π is defeated by preclusion, the path σ responsible for precluding it will be called the *precludor* of π.[15] Similarly, if a path π is defeated by conflict, the path defeating it will be called the *conflictor* of π. Just like the authors of [18], we are most interested in *defeater-defeaters*, or second order defeat. The intuition behind it is this: it may be possible that the ability of a path to act as a defeater is itself defeated by some other path. Given the two types of defeaters we recognise — precludors and conflictors —, four kinds of defeater-defeaters are in principle possible:

[15] In [18] it is called 'a preemptor'.

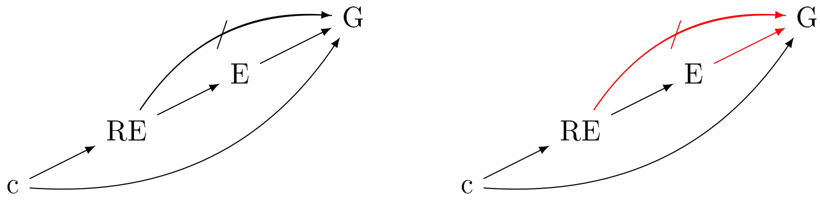

Figure 9: Γ_8 and a graphic depiction of $E(\Gamma_8)$

- precludor-precludor;
- precludor-conflictor;
- conflictor-precludor;
- conflictor-conflictor.[16]

In [18] these four combinations are evaluated from the perspective of a certain skeptical approach initially introduced in (Horty *et al.*, [7]). The first two are rejected on intuitive grounds, at least some cases of the fourth are accepted, while the third is only mentioned. My discussion of the defeater-defeaters will be based on concrete examples. That is to say, I will present an inheritance network that could, in principle, allow for a certain kind of defeater-defeat, introduce the statement set that our *isa*-framework predicts for the given net, and then judge whether the effect of second order defeat is present.

Let us begin with precludor-precludors. Consider the network depicted in Figure 9, on the left. Its interpretation is this: c = Clyde, RE = royal elephant, E = elephant, and G = *grey*. Notice that, on our definition, the path $(c, RE, E, G, +)$ must be precluded by $(c, RE, G, -)$. However, the path $(c, RE, G, -)$ is itself precluded by the direct link $(c, G, +)$. Now, $(c, G, +)$ would be a precludor-precludor in case it would cancel the preclusion of the path $(c, RE, E, G, +)$. But let us have a look at what our *isa*-approach predicts for Γ_8. Despite the fact that it operates on sets of statements, the network is simple enough for us to be able to present a certain visual approximation of its minimal (and also sole) extension E. We focus only on two kinds of statements: the *isa*- and the *abnorm*- ones. Besides, only on those which begin with c. Now, the *isa*-statements are depicted as standard paths, while the *abnorm*- ones are depicted as paths that end in red links. The approximation itself is presented in Figure 9 to the right of Γ_8. The reason we introduce it is the fact that it allows us to talk of effect of essentially path-based notions in the context of our approach that is hardly path-

[16]Touretzky *et al.* list one more defeater-defeater in addition to these, but I have chosen to leave it out so as not to make the discussion overtly complicated.

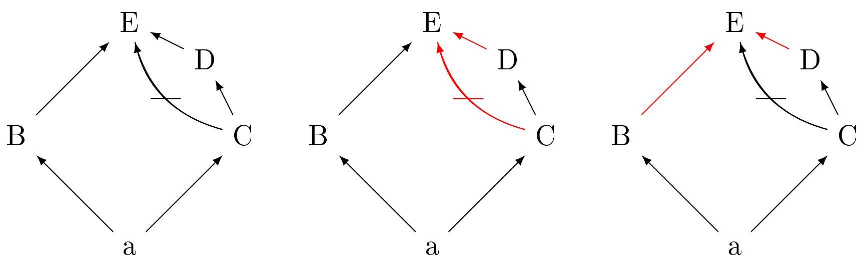

Figure 10: Γ_9, $E(\Gamma_9)$, and $E'(\Gamma_9)$

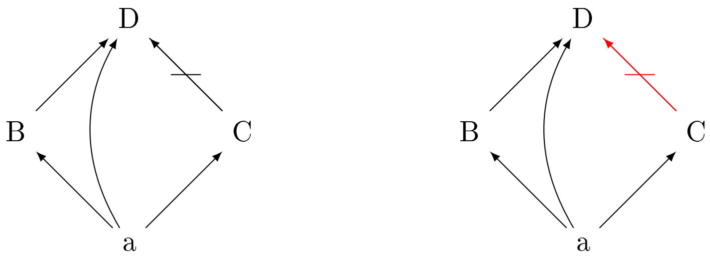

Figure 11: Γ_{10} and $E(\Gamma_{10})$

based. Anyway, note that c is abnormal with respect to both $(c, RE, G, -)$ and $(c, RE, E, G, +)$. This can only mean that, in spite of being precluded, $(c, RE, E, G, +)$ does not loose its ability to preclude. Therefore, we conclude that our *isa*-approach does not admit precludor-precludors.

Now we proceed to precludor-conflictors and network Γ_9 (Figure 10, on the left). The only precludor here is $(a, C, E, -)$. Note that it is in conflict with the path $(a, B, E, +)$. We would say that the system allows for precludor-conflictors in case $(a, B, E, +)$ would be somehow able to cancel the preclusion of $(a, C, D, E, +)$. It turns out that Γ_9 has two extensions — E and E' —, and both qualify as minimal. The visualisation makes it plain that the default $(D, E, +)$ is precluded for a in both E and E', independently of what happens to $(C, E, -)$. This shows that the effect of precludor-conflictors is also absent from our *isa*-framework.

Any kind of defeat of precludors is referred to as *reinstatement* in [18]. Its authors also argue that it is counter-intuitive, and it seems that they are absolutely right. To use Γ_8 as an example, painting an albino elephant grey hardly makes it a normal elephant. More generally, exceptions to exceptions do not produce normalities. Thus, we should be pleased that our system does not allow for defeat of precludors.

Our next step are conflictor-defeaters. Consider the net Γ_{10} depicted in Figure 11. We say that the path $(a, C, D, -)$ is the conflictor of $(a, B, D, +)$.

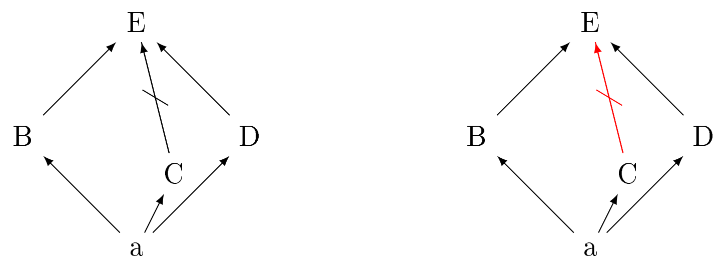

Figure 12: Γ_{11} and $E(\Gamma_{11})$

It is also the case that $(a, B, D, +)$ is the conflictor of $(a, C, D, -)$, but it does not really matter for the present purposes. Notice that the direct link $(a, D, +)$ precludes $(a, C, D, -)$, which makes it a good candidate for a conflictor-precludor. Along with Γ_{10}, Figure 11 also contains a visualisation of its minimal (and only) extension. The latter shows that a is rendered abnormal with respect to 'C-s usually are not D-s', and, thus, the conflictor of $(a, B, D, +)$, i.e., $(a, C, D, -)$, is, in fact, defeated by preclusion. This indicates the presence of the effect of conflictor preclusion in our *isa*-approach.

At this point only conflictor-conflictors remain. Let us turn to Γ_{11} (Figure 12) and focus on the path $(a, C, E, -)$. It clearly is a conflictor of $(a, B, E, +)$, but it has another conflictor itself, namely, $(a, D, E, +)$. Γ_{11} has two extensions. One of them contains $isa(a, E, -)$ along with two statements of the *abnorm*-form: $abnorm(a, B, E, +)$ and $abnorm(a, D, E, +)$. The other extension contains $isa(a, E, +)$ and only one abnormality statement, namely, $abnorm(a, C, E, -)$. Now, the latter clearly is the minimal extension of Γ_{11} and it is the one visualised in Figure 12. It is given preference only due to the fact that it contains less statements of the *abnorm*-form, but it exhibits the same effect conflictor-conflictors would have. Were $(a, B, E, +)$ and $(a, C, E, -)$ the only paths leading to E, the latter would act as a defeater of the former. But, with an additional positive path, $(a, C, E, -)$ is overwhelmed. Hence, the conflictor-conflictor effect is also present in our framework.

Conclusion

Our main aim was to develop an approach to default reasoning that would augment the network-based frameworks with the idea of minimising abnormality. Instead of one approach, though, we have devised two.

The one presented in the first section is in line with the received way of modeling defaults by means of defeasible inheritance nets. That is to say, it is path-based. We began with a fairly standard system, but extend it soon afterwards. Most importantly, we introduced what we called 'abnormality paths' and adjusted the system in such a way that they would enter credulous expansions — path sets representing possible lines of reasoning on the basis of the given net. As a result, we were able to determine the amount of abnormalities each expansion implies and single out the minimal ones, i.e., those that imply as few abnormalites as possible. We also argued that the conclusion sets associated with networks are exactly the statement sets supported by minimal expansions.

Then we turned to devising of an alternative formalism using the framework of (Sandewall, [13]). Unlike the first approach, it operates on statements, rather than paths, and its very core is a set of non-monotonic inference rules. We began the second section with an introduction into the framework: the *isa*-language, the core concepts, and the fact that neither Sandewall's rules, nor those of (Simonet, [15]) are satisfactory. Afterwards we devised our own set of non-monotonic inference rules R, which required, among other things, extending the original language and developing an alternative notion of preclusion. In the end of the section we presented some results characterising the newly developed *isa*-framework. In particular, we established a correspondence between it and the basic path-based approach developed in section (1). That it is to say, we showed that any *extension* E of a given net Γ — i.e., a total set of statements inferable from Γ by means of R — has a corresponding credulous expansion Φ of Γ, and vice versa. Intuitively, the two are said to correspond to each other, in case both encode the same information.

In the third section the *isa*-framework was extended so as to be able to circumscribe abnormality. We expanded the language by allowing for formulae of the *abnorm*-form, i.e., statements that say explicitly that an object is abnormal with respect to a certain default. Then we presented the set of inference rules R' comprised of rules from R (with some adjustments) and two new ones. The last rule, that is, rule (6), was added in order to deal with the problem of decoupling. Having the *abnorm*-statements at hand, we were able to count the number of abnormalities each extension implies and define the notion of a minimal extension, i.e., extension that minimises (= circumscribes) abnormality. At the end we discussed some concrete examples of networks for which exactly the minimal extensions are the intuitively correct ones.

It is too early to draw any far-reaching conclusions about the two approaches though, for there is still a lot of work to be done. In particular:

- The approaches should be tested on many more different examples. Only afterwards would we be justified to draw conclusions about the intuitiveness of the extensions (expansions) that circumscribe abnormality. Here implementations could prove to be of great help.

- Although we have established a correspondence between the two approaches in their basic form, we still have to determine the relation between minimal extensions and minimal expansions. We will not be able to establish a one-to-one relation, for our path-based approach allows for decoupling, while the *isa*-framework does not. Still I am certain that minimal expansions that contain no decoupled paths correspond to minimal extensions, and vice versa. I envisage no serious difficulties in extending the proofs of the two theorems.

- Cyclic networks form another area that would allow us to judge the adequacy of our *isa*-framework and also estimate its full power. In the received path-based approaches cycles have usually been disregarded, and one of the compelling reasons seems to have been the fact that they are simply difficult to handle when one is working with paths and links. The *isa*-framework proceeds differently and seems to provide a more natural environment for dealing with cyclic nets. Besides, we have also made a slight change in the standard definition of network consistency, having exactly cycles in mind. With all this cyclic networks seem to provide a fruitful area for future research. It would be especially pleasing if we were able to generalise proposition 1, that is to say, prove that every consistent network (with or without cycles) has an extension.

- We should try extending the *isa*-framework in different ways, which would allow us to judge its flexibility. There are two directions worth considering. First, recall how we dealt with the problem of decoupling. In order to get the desired result we only had to add a new inference rule. Now, it would be interesting to see whether other constraints could also be enforced by adding rules of inference. Second, the expressibility of the `IS-A` links (and, thus, networks) is quite limited. Although hardly sufficient for purposes of realistic knowledge representation, addition of *strict links* — universal statements — is the most straightforward way to enhance the language of nets. This, though, would certainly require adjusting and rethinking the rule set R', but would also make us see how flexible the approach is.

- Both the path-based approach and the *isa*-system are proof-theoretic in their nature. Finding a corresponding semantics certainly seems to be an interesting and one of the most important directions for future research.

Acknowledgement

First and foremost, I would like to thank Frank Veltman. Not only for all the valuable comments, suggestions, and discussions we had during the months this thesis was being written, but also for the fact that he made the whole process of working on it highly enjoyable. I am also greatful to Reinhard Blutner and Sonja Smets for agreeing to be on my thesis committee, and to my fellow student Riccardo Pinosion for the discussions we had while writing a joint paper on defaults and abnormalities for Frank's course *Philosophical Logic*.

Appendix A: Proofs of Theorems and the Ordering Lemma

Lemma 2 (Ordering lemma). *Any extensions E of Γ can be obtained with the following decreasing order of priority on the choice of instantiation of rules from R: (1) the first rule; (2) the fourth rule; (3) the second rule; (4) the third rule.*

Proof. The result will follow immediately, as soon as we prove the following:

Claim 1. *For any sequence of increasing statement sets E_0, E_1, $\ldots$ that is obtained by rules from R, there is a corresponding sequence E_0, E'_1, $\ldots$ which* (i) *has the same limit and* (ii) *complies with the above order of priority on rule instantiation.*

(Proof of claim) The proof is by induction on the construction of a sequence.

Base case. Trivial.

Inductive step. We suppose that the claim holds for all sequences constructed in $n-1$ steps and show that it holds for sequences constructed in n steps as well. Since there are four inference rules, we have four cases to consider.

(1) E_n is acquired by an application of the first rule. Here $E_n = E_{n-1} \cup \{isa(x,y,s), interm(x,x,y,s)\}$. For the ease of discussion we refer to $\{isa(x,y,s), interm(x,x,y,s)\}$ as Δ. Now, by the inductive hypothesis, we know that E_0, E_1, $\ldots$, E_{n-1} has a corresponding sequence E_0, E'_1, $\ldots$, E'_{n-1} that obeys the priority ordering. We set E_0, $E_0 \cup \Delta$, $E'_1 \cup \Delta$, $\ldots$, $E'_{n-1} \cup \Delta$ to be the sequence corresponding to E_0, E_1, $\ldots$, E_{n-1}, E_n. It is evident that the new sequence complies with the ordering.

(2) E_n is obtained by an instantiation of rule (2). Here the sequence we are dealing with looks as follows: E_0, E_1, $\ldots$, E_{n-1}, $E_{n-1} \cup \{precl(x,y,z,s)\}$. Again, we will refer to $\{precl(x,y,z,s)\}$ as Δ. By the inductive hypothesis, $E_0, E_1, \ldots, E_{n-1}$ must have an order-obeying counterpart. Now, let E'_i be the first statement set in it that contains all the conditions necessary for inferring $precl(x,y,z,s)$. Let E'_j

$(i \leq j \leq n-1)$ be the first statement set after E'_j that is acquired by rules other than (1) or (4). Thus, the sequence at issue has to look as follows:

E_0, E'_1, ..., E'_i, ..., E'_{j-1}, E'_j, ..., E_{n-1}.

Now we set E_0, E_1, ..., E_n to correspond to:

E_0, E'_1, ..., E'_i, ..., E'_{j-1}, $E'_{j-1} \cup \Delta$, $E'_j \cup \Delta$, ..., $E_{n-1} \cup \Delta$.

It is not difficult to check that this sequence complies with the ordering. Suppose it did not. Then there would have to be some $E'_k \cup \Delta$ with $j < k \leq n-1$ acquired by either (1) or (4) such that the formulae added to $E'_{k-1} \cup \Delta$ could have been inferred at step j already. We consider one case only. Suppose $E'_k \cup \Delta$ was acquired by rule (4). Then $E'_{j-1} \cup \Delta$ must contain formulae $interm(x', y', z', +)$ and $interm(x', z', w, s')$, which can only mean that these formulae are in E'_{j-1}. This implies, however, that the original sequence E_0, E'_1, ..., E'_i, ..., E'_j, ..., E'_k, ..., E'_{n-1} did not comply with the above order. For at step j either rule (2) or (3) was applied, when (4) could have been instantiated just as well.

(3) E_n is obtained by applying rule (3). Here our sequence is E_0, E_1, ..., E_{n-1}, $E_{n-1} \cup \{isa(x, z, s), \neg precl(x, y, z, s), interm(x, y, z, s)\}$. This is the easy case. If E_0, E'_1, ..., E'_{n-1} is the sequence corresponding to E_0, E_1, ..., E_{n-1}, let E_0, E'_1, ..., E'_{n-1}, $E'_{n-1} \cup \{isa(x, z, s), \neg precl(x, y, z, s), interm(x, y, z, s)\}$ be the one corresponding to E_0, E_1, ..., E_n.

(4) E_n is a result of an instantiation of the fourth rule. Here the sequence is E_0, E_1, ..., E_{n-1}, $E_{n-1} \cup \{interm(x, y, w, s)\}$. Let Δ stand for $\{interm(x, y, w, s)\}$ for the ease of discussion. E_0, E_1, ..., E_{n-1} has a corresponding sequence that complies with the above order. Let E'_i be the first statement set that contains the formulae that are needed to infer $interm(x, y, w, s)$. Let E'_j (with $i \leq j \leq n-1$) be the first set after E'_i that is not acquired by rule (1). The sequence we are dealing with here must look as follows:

E_0, E'_1, ..., E'_i, ..., E'_{j-1}, E'_j, ..., E_{n-1}.

We set E_0, E_1, ..., E_n to correspond to:

E_0, E'_1, ..., E'_i, ..., E'_{j-1}, $E'_{j-1} \cup \Delta$, $E'_j \cup \Delta$, ..., $E_{n-1} \cup \Delta$.

It is straightforward to see that this sequence complies with the order. For suppose it did not. Then there would have to be a set $E'_{k-1} \cup \Delta \cup \{isa(x', y', s'), interm(x', x', y', s)\}$, with $j < k \leq n-1$, acquired by rule (1) such that $isa(x', y', s')$ and $interm(x', x', y', s)$ might have been derived already at step j. Note that the only condition for inferring these statements — $isax(x', y', s')$ — is present already in E_0, which implies that these statements could have been derived at step 1 of the construction. Now, given that E'_j was obtained by some other rule than (1) and that $k > j$, we can conclude that the original sequence E_0, E'_1, $\ldots$, E_{n-1}, $\ldots$ was not order obeying. Thus, we have a contradiction.

This concludes the proof of our claim.

The rest is trivial. Consider an extension E of some arbitrarly network Γ. Clearly, E must be a limit of an increasing sequence of statement sets E_0, E_1, E_2, $\ldots$. By the claim, there is another sequence E_0, E'_1, E'_2, $\ldots$ that complies with the above order and has the same limit as E_0, E_1, E_2, $\ldots$. Hence, E can be obtained with the specified order of priority on rule instantiation.

End of proof.

$\square$

Theorem 1. *For any consistent statement set E that is obtained by the set R of inference rules from a consistent inheritance net Γ (with priority ordering as specified in the Ordering lemma) there is a corresponding consistent path set Φ such that:*

(*i*) *for any $\pi \in \Phi$, $\langle \Gamma, \Phi \rangle \mid\!\sim \pi$, and*

(*ii*) *for any isa-statement, $isa(x, z, s) \in E$ iff for some path $\sigma = (x, \sigma', z, s)$, $\sigma \in \Phi$.*

Proof. The proof is by induction on the construction of E. Our order of priority on the choice of rule application is as follows (the reason we need it will become clear as the proof proceeds):

> (1) The first rule; (2) the fourth rule; (3) the second rule; (4) the third rule.

Base case. All formulae of E_0 are of the $isax$-form. Now, since there are no formulae of the isa-form in E_0, we can let $\emptyset$ be the set corresponding to E_0.

Inductive step. We suppose that the claim holds for all statement sets that have been constructed in $n-1$ steps and show that it hold for the n-th step as well. There are four inference rules in our system and thus four cases to consider.

(1) E_n is acquired by an application of the first rule. Then $isax(x, z, s) \in E_{n-1}$ and $E_n = E_{n-1} \cup \{isa(x, z, s), interm(x, x, y, s)\}$. Clearly, we have $(x, z, s) \in \Gamma$. Now, by the inductive hypothesis, we know that there is a path set Φ' such that $\langle \Gamma, \Phi' \rangle \mid\!\sim \pi$ for any $\pi \in \Phi'$ and that it

contains all and only the paths corresponding to the *isa*-formulae of E_{n-1}. Since (x, z, s) is a link of Γ, we can be sure that $\langle\Gamma, \Phi'\rangle \mid\!\sim (x, z, s)$. We set $\Phi' \cup \{(x, z, s)\}$ to be the path set corresponding to E_n.

(2) E_n is acquired by an application of the second inference rule. Since E_n does not contain any additional *isa*-formulae but those already present in E_{n-1}, in this case the claim holds by the inductive hypothesis alone.

(3) E_n is a result of applying the third inference rule. Here $E_n = E_{n-1} \cup \{isa(x, z, s), \neg precl(x, y, z, s), interm(x, y, z, s)\}$. We can be sure that $isa(x, y, +)$ and $isax(y, z, s) \in E_{n-1}$ and also that $precl(x, y, z, s)$ and $isa(x, z, -s) \notin E_{n-1}$ (negative conditions of applying the 3rd rule). By the inductive hypothesis we know that there is a path set Φ' such that $\langle\Gamma, \Phi'\rangle \mid\!\sim \pi$ for every $\pi \in \Phi'$ and that it contains all and only the paths corresponding to the *isa*-formulae of E_{n-1}. Since $isa(x, y, +) \in E_{n-1}$, there must be a path of the form $(x, \sigma, y, +) \in \Phi'$. Since $isax(y, z, s) \in E_{n-1}$, Γ must contain a link of the form (y, z, s). This is enough to conclude that the path (x, σ, y, z, s) is constructible in the context $\langle\Gamma, \Phi'\rangle$. Since $isa(x, z, -s) \notin E_{n-1}$, we can be sure that (x, σ, y, z, s) is not conflicted in $\langle\Gamma, \Phi'\rangle$. Now, it remains to show that it is not precluded in this context as well. Since $precl(x, y, z, s) \notin E_{n-1}$, we can be sure that for no v we have both $interm(x, v, y, +)$ and $isax(v, z, -s) \in E_{n-1}$ (recall the constraints that we have put on the order of rule application). Our next step is to claim the following:

Claim 2. *There is no node v such that we would have both (i) v is an intermediary to the path $(x, \sigma, y, +)$ in the path set Φ' and (ii) Γ contains a direct link of the form (v, z, s).*

(Proof of claim) Suppose to the contrary, i.e., that there is such a node v'. (ii) tells us that $isax(v', z, s) \in E_{n-1}$. (i) tells us that either v' is a node of $(x, \sigma, y, +)$ itself or there is a path of the form $(x, \rho, v', \rho', w, y, +) \in \Phi'$. If the former is the case, $interm(x, v', y, +)$ must have been added at some point in the process of constructing E_{n-1}, i.e., either simultaneously or after the addition of $isa(x, y, +)$. Otherwise, E_n could not have been acquired by the application of the third rule (recall our priority ordering on rule instantiation). If the latter is the case, at some point in the construction of our statement set the third rule must have been applied on $isa(x, w, +)$ and $isax(w, y, +)$. Now, given that all the other rules must have been applied before the acquisition of E_n (again, constraints on constructions), $interm(x, v', y, +)$ must already be in E_{n-1}. In either case we get a contradiction.

This claim allows us to concluded that $\pi' = (x, \sigma, y, z, s)$ is not i-precluded in the context $\langle\Gamma, \Phi'\rangle$. By the Preclusion lemma, we can be sure that it is not off-path precluded as well. Now, since (x, σ, y, z, s) is (a) constructible, (b) not conflicted, and (c) not preempted in $\langle \Gamma, \Phi' \rangle$, we can be sure that $\langle \Gamma, \Phi' \rangle \mid\!\sim \pi'$. So we set $\Phi' \cup \{\pi'\}$ to be the set corresponding to E_n.

It remains to check whether we still have $\langle \Gamma, \Phi \rangle \mid\!\sim \pi$ for all $\pi \in \Phi$. So let us suppose to the contrary, i.e., that there is some path $\rho = (x_1, \rho', x_{n-1}, x_n, s) \in \Phi$ such that ρ is not inheritable in $\langle \Gamma, \Phi \rangle$. Notice that $\rho \neq \pi'$ and $\langle \Gamma, \Phi' \rangle \mid\!\sim \rho$. The latter means that ρ was constructible, not conflicted, and not precluded in $\langle \Gamma, \Phi' \rangle$. Clearly, it must still be constructible and not conflicted in $\langle \Gamma, \Phi \rangle$. Hence, it can only be precluded. By the Preclusion lemma, we can be sure that there must be some node v' such that v' is an intermediary to the path $(x_1, \rho', x_{n-1}, +)$ in the path set Φ and $(v', x_n, -s) \in \Gamma$. Since $(x_1, \rho', x_{n-1}, +) \in \Phi'$, we have $isa(x_1, x_{n-1}, +) \in E_{n-1}$. Likewise, E_{n-1} must contain $interm(x_1, v', x_{n-1}, +)$, $isax(x_{n-1}, x_n, s)$, and $isax(v', x_n, -s)$ (again, recall the order of rule instantiation). Since E_n was acquired by the third rule, we can be sure that $precl(x, x_{n-1}, x_n, s) \in E_{n-1}$ (the order). Given that $\rho \in \Phi'$, E_{n-1} must contain the corresponding statement $isa(x_1, x_n, s)$ (inductive hypothesis). Note that ρ is a compound path. Hence, given the way we map statement sets into path sets, we can be sure that $isa(x_1, x_n, s)$ has been included in E_{n-1} by an application of rule (3). Hence, we have $\neg precl(x_1, x_{n-1}, x_n, s) \in E_{n-1}$ as well. This, however, implies that E_{n-1} is inconsistent, which contradicts our assumption. Therefore, we can be sure that, for all $\pi \in \Phi$, it is the case that $\langle \Gamma, \Phi \rangle \mid\!\sim \pi$.

(4) Same as (2).

This concludes the proof.

□

Theorem 2. *For any expansion Φ of a consistent network Γ there exists a corresponding extension E such that, for any path $\pi = (x, \pi', z, s)$, $\pi \in \Phi$ iff $isa(x, z, s) \in E$.*

Proof. We need to define the notion of a *path length* for this proof. Let the length of a path π, $len(\pi)$, be the number of links it consists of. Thus, for $\pi' = (a, B, C, D, +)$ $len(\pi') = 3$. Notice that the number of nodes of any given path π is $len(\pi) + 1$. Now we turn to the proof.

Let Φ be an expansion of some inheritance network Γ. Let n be the length of the longest path in Φ. Now for each i, $1 \leq i \leq n$, we take the corresponding subset Φ_i of Φ:

$$\Phi_i = \{\pi \in \Phi : len(\pi) = i\}$$

These subsets form a sequence $\Phi_1, \Phi_2, \ldots, \Phi_n$ such that $\bigcup_{i=1}^{n} \Phi_i = \Phi$. Notice that for any path π of length i such that $1 < i \leq n$ we have the following: π is of the form (x, π', y, z, s), $(x, \pi', y, +) \in \Phi_{i-1}$, and $(y, z, s) \in \Phi_1$. Othrewise, π would not be constructible and, thus, Φ could not have been an expansion. More generally, for any i, $1 < i \leq n$, and any $\pi \in \Phi_i$, we have $\langle \Gamma, \Phi_{i-1} \rangle \mid\!\sim \pi$.

Now we will construct a sequence of increasing sets of statements E_0, E_1, E_2, ..., making use of our sequence $\Phi_1, \Phi_2, \ldots$ as follows:

- Set $E_0 = \{isax(x, z, s) : (x, z, s) \in \Gamma\}$.

- Let $k = |\Phi_1|$. E_1, ..., E_k are constructed from Φ_1. Let each E_j, where $1 \leq j \leq k$, be constructed as follows: for each $\pi = (x, z, s) \in \Phi_1$, apply the 1st inference rule on the corresponding $isax(x, z, s) \in E_{j-1}$. Notice that $\bigcup_{j=1}^{k} E_j = \{isa(x, z, s) : (x, z, s) \in \Gamma\} \cup \{interm(x, x, z, s) : (x, z, s) \in \Gamma\}$.

- Now we specify what to do with each Φ_i such that $1 < i \leq n$. Let E_l be the last statement set obtained from Φ_{i-1} and $m = |\Phi_i|$. We show how to construct E_{l+1}, ..., E_{l+m}. We construct each E_j, where $l < j \leq m$ as follows: for each $\pi = (x, \pi', y, z, s) \in \Phi_i$, apply the 3rd inference rule on the corresponding $isa(x, y, +)$ and $isax(y, z, s) \in E_{j-1}$. Then proceed to Φ_{i+1}, until Φ_n is reached.

It is easy to verify that at each step j the third rule can be applied. Thus, in case $isa(x, y, +)$ or $isax(y, z, s)$ where not in E_{j-1}, either $(x, \pi, y, +)$ would not be in Φ_{i-1} or (y, z, s) would not be in $\Phi_1 = \Gamma$. Either way we would not have $\langle \Gamma, \Phi_{i-1} \rangle \mid\!\sim (x, \pi, y, z, s)$, and, *per consequens*, (x, π, y, z, s) could not have been in Φ_i. Similarly, we can check that at each step the negative condition of the rule is also satisfied. For suppose it was not. Recall that it specified that neither $precl(x, y, z, s)$ nor $isa(x, z, -s)$ may be present in E_{j-1}. Since by construction we can be sure that E_{j-1} does not contain any formulae of the *precl*-form, it must be $isa(x, z, -s)$. Now, again by construction, we can be sure that this $isa(x, z, -s)$ has been acquired from some path $\rho = (x, \rho', z, -s) \in \Phi$ such that $len(\rho) \leq len(\pi)$. This, however, means that π is not inheritable in $\langle \Gamma, \Phi \rangle$ and Φ is no expansion.

- Keep instantiating the fourth rule until it is not possible to apply it anymore.

- Keep instantiating the second rule until no further applications are possible.

It is obvious that each E_i, $1 < i$ of the sequence E_0, E_1, E_2, ... is acquired by an application of one of the inference rules from R. It remains to show that the limit of this sequence E is an extension of Γ. To this end we prove the two following claims.

Claim 3. *E is consistent.*

(Proof of claim) Suppose to the contrary. Then there are some two formulae p and $\neg p \in E$. Given our set of rules, p has to be of the form $precl(x, y, z, s)$.

Thus, we have $precl(x, y, z, s)$ and $\neg precl(x, y,\ z, s) \in E$. The latter must has been added after an instantiation of the third rule together with $isa(x, z, s)$. Now, this $isa(x, z, s)$ has a corresponding path $(x, \sigma, y, z, s) \in \Phi$. $precl(x, y, z, s)$, on the other hand, must have been added by the second rule, which tells us that E also contains formulae $isa(x, y, +)$, $interm(x, v, y, +)$, $isax(y, z, s)$, and $isax(v, z, -s)$. Yet again, $isa(x, y, +)$ has a corresponding path $(x, \rho, y, +) \in \Phi$, while $isax(y, z, s)$ and $isax(v, z, -s)$ tell us of the corresponding links (y, z, s) and $(v, z, -s) \in \Gamma$. It is easy to see that $interm(x, v, y, +)$ implies existence of a node v such that v is an intermediary to the path $(x, \sigma, y, +)$ in Φ. Now, given this and $(v, z, -s)$, we can conclude that the path (x, σ, y, z, s) is precluded in Φ, and, therefore, Φ cannot be an expansion. Thus, we have a contradiction.

Claim 4. *E is a fixed point.*

(Proof of claim) Suppose to the contrary. Then it must be possible to apply one of the rules from R and add some new statement to E. It is obvious that this cannot be rule (1), (2), or (4). Since it has to be (3), E must contain two statements of the form $isa(x, y, +)$ and $isax(y, z, s)$. Likewise, it must be the case that no statements of the form $isa(x, z, -s)$ and $precl(x, y, z, -s)$ are in E. Now, $isa(x, y, +)$ and $isax(y, z, s)$ tell us that Φ contains some path of the form $(x, \sigma', y, +)$ and that Γ contains the link (y, z, s). This means that the path $\sigma = (x, \sigma', y, z, s)$ is constructible in $\langle \Gamma, \Phi \rangle$. Further, the fact that there are no statements of the form $isa(x, z, -s)$ is enough to conclude that there are no path of the form $(x, \rho, z, -s)$ in Φ. This shows that σ is also not conflicted in $\langle \Gamma, \Phi \rangle$. Finally, it is not difficult to show that the absence of $precl(x, y, z, -s)$ from E implies that σ is not precluded in $\langle \Gamma, \Phi, \rangle$. Given all this, we have $\langle \Gamma, \Phi \rangle \mid\!\sim \sigma$. Recall that Φ is a fixed point. Hence, we can be sure that $\sigma \in \Phi$. However, given that σ is a part of Φ, at some point in the process of constructing our sequence E_0, E_1, ..., we must have already applied rule (3) on $isa(x, y, +)$ and $isax(y, z, s)$, and expanded it accordingly. Since we have arrived at a contradiction here, we can be sure that E is a fixed point.

This concluded the proof of the theorem. □

Appendix B: Simonet's Inference Rules

In [15] Geneviève Simonet presents three sets of non-monotonic inference rules. They are devised with a view of transferring three different path-based theories of inheritance to Erik Sandewall's *isa*-framework. The sole thing that makes these theories different is the notion of preclusion they rely on. Consequently, the three sets of inference rules also differ only by the rule that is responsible for preclusion. The first set, call it S, is supposed to implement a path-based theory that is very much alike the one we have developed in the first section. Here is how S looks:

1. If $isax(x, y, s) \in E$,
   then add $isa(x, y, s)$ to E.
2. Inference of $precl(x, y, z, s)$:
   1. If $isa(x, v, +)$ $isax(v, y, +)$, $isa(x, y, +)$ and $isax(v, z, -s) \in E$,
      and $precl(x, v, y, +)$ and $isax(x, y, -) \notin E$,
      then add $precl(x, y, z, s)$ and $\neg precl(x, v, y, +)$ to E.
   2. If $precl(x, w, z, s)$, $isax(w, y, +)$ and $isa(x, y, +)$ are in E,
      and $precl(x, w, y, +)$ and $isax(x, y, -)$ are not in E,
      then add $precl(x, y, z, s)$ and $\neg precl(x, w, y, +)$ to E.
3. If $isa(x, y, +)$ and $isax(y, z, s) \in E$,
   and $precl(x, y, z, s)$ and $isa(x, z, -s) \notin E$,
   then add $isa(x, z, s)$ $\neg precl(x, y, z, s)$ and $\neg isa(x, z, -s)$ to E.

The second rule is the one that implements preclusion; off-path preclusion in this case. According to Simonet, $1 - 3$ are a generalisation of the rules initially put forward by Sandewall in [13] together with her revision proposal. Be it as it may, there are at least four problems with S. I now consider them one by one.

Problem 1: effect of preclusion without preclusion

Let us consider the inheritance network Γ_{12} that is an even simpler version of the Tweety Triangle discussed above. It consists of the links $(tweety, Bird, +)$, $(Bird, Fly, +)$, and $(tweety, Fly, -)$. Clearly, the corresponding E_0 must look as follows:

$$\{isax(tweety, Bird, +),\\ isax(Bird, Fly, +),\\ isax(tweety, Fly, -)\}.$$

One of the ways to acquire the unique extension E of Γ_{12} by means of S is this:

E: Apply (1) on $isax(tweety, Fly, -)$ and add $isa(tweety, Fly, -)$ to E_0. Apply (1) two more times, adding $isa(tweety, Bird, +)$ and $isa(Bird, Fly, +)$ to the statement set. That is it.

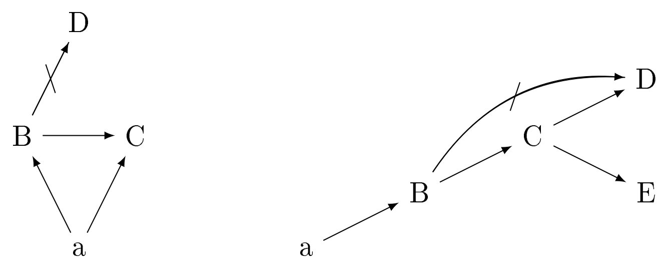

Figure 13: Γ_{13} and Γ_{14}

At this point already no non-redundant instantiations of Simonet's inference rules are possible. In particular, (3) cannot be applied on $isa(tweety, Bird, +)$ and $isax(Bird, Fly, +)$ due to the presence of $isa(tweety, Fly, -)$ (see the second negative condition of the rule). Likewise, (2) cannot be applied for the lack of positive conditions. Given the way the inference rules are defined, E would have to contain $isa(tweety, tweety, +)$ to enable inference of $precl(tweety, Bird, Fly, +)$. Now, the conclusion set of E — $\{\varphi \in E : \varphi = isa(x, z, s)\}$ — is identical to the conclusion set of the corresponding expansion, but we can hardly be entirely content with how the inference rules fare in this situation. First, the extension does not contain any statement saying that the inference of $isa(tweety, Fly, -)$ is precluded. This is weird, given that (*i*) in any path-based approach that relies on off-path preclusion $(tweety, Bird, Fly, +)$ comes out precluded, and that (*ii*) in the *isa*-framework preclusion works by means of *precl*-statements. Second, the would-be extension in which $isa(tweety, Fly, +)$ is derived in two steps is made inconsistent not by (2), as one would expect, but rather by an application of (1). Again, since preclusion is at issue, we would (arguably) want the corresponding rule to be responsible for making the would-be extension inconsistent.

Problem 2: preclusion of phantom defaults

Consider the net Γ_{13} that is depicted in Figure 13, on the left. The problem here is that the unique extension of Γ_{13} contains the statement $precl(a, C, D, +)$. This means that Simonet's rules permit precluding non-existent defaults; $(C, D, +)$ in this particular case. The fault is with (2.1) which allows inferring $precl(a, C, D, +)$ on the basis of $isa(a, B, +)$, $isax(B, C, +)$, $isa(a, C, +)$, and $isax(B, D, -)$. The network Γ_{14} that is situated on the right from Γ_{13} presents a very similar problem: its extension contains $precl(a, E, D, +)$, but in this case the fault is with (2.2).

Problem 3: superfluous conditions

Let us focus on the negative conditions for instantiating the subrules of (2). If the first one is to be instantiated, $precl(x, v, y, +)$ and $isax(x, y, -)$ must not be in E; similarly, if the second one is to be applied, $precl(x, w, y, +)$ and $isax(x, y, -)$ must not be present in E. Now, it turns out that the $isax(x, y, -)$ part is superfluous. In general, negative conditions are present in the rules in order to block their instantiation in certain cases. However, all the cases in which $isax(x, y, -)$ can act as a block are not worth saving, as the following line of reasoning shows.

Suppose that all the positive conditions for instantiating (2.1) (or (2.2)) are in E, but its instantiation still cannot take place due to the presence of $isax(x, y, -)$. Clearly, E must contain $isa(x, y, +)$. As any other statement of the *isa*-form it must have been derived either by the first, or by the third inference rule. Suppose it was derived by rule (1). Then we have $isax(x, y, +) \in E$. Since we have both $isax(x, y, +)$ and $isax(x, y, -) \in E$, the inheritance network at issue must contain $(x, y, +)$ and $(x, y, -)$, and contradictory nets of such kind are certainly of not much interest. Suppose now that $isa(x, y, +)$ was derived by the third rule. Given the way it is defined, E would also have to contain $\neg isa(x, y, -)$. Now, having $isax(x, y, -) \in E$ means that at some point we have to apply the first rule on it, infer $isa(x, y, -)$, and, thus, render the statement set at issue inconsistent.

To summarise, if instantiation of (2.1) or (2.2) is blocked by its second negative conditions, we are dealing with either a contradictory network, or with a contradictory would-be extension. In either case, there is no reason for blocking the rule, and, therefore, having $isax(x, y, -)$ as its part amounts to nothing.

Problem 4: doubts about correspondence

The three problems considered thus far are hardly grave and could be fixed by some relatively minor adjustments to the rules. The next one, however, is of a different sort: it undermines the most important result of (Simonet, [15]) and there seems to be no obvious way to fix it.

Simonet's main result is a correspondence theorem between three path-based theories and three sets of non-monotonic inference rules. In particular, the set S we have focused on hitherto is said to correspond to a path-based theory which is very much alike the one I presented in the first section. Therefore, the result (restricted to S) can be restated as follows:

> *For any network Γ, the conclusion sets associated with the extensions of Γ obtained by the set of inference rules S are exactly the sets supported by the credulous expansions of Γ.*

It turns out, that there are networks for which this does not hold. Γ_{15} that is depicted in Figure 14 is one such. In any path-based approach that relies on

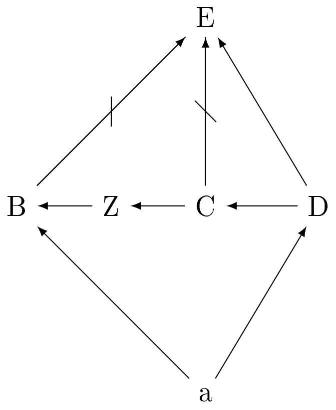

Figure 14: Γ_{15}

off-path preclusion, including ours and the one used by Simonet, $(a, D, E, +)$ has to preclude $(a, B, E, -)$. However, this is just what does not happen for S. There is no way $precl(a, B, E, -)$ could be derived by (2.1). Likewise, there is no way for preclusion to transfer from $(C, E, -)$ to $(B, E, -)$ by means of (2.2).[17] It could happen only in case $\Gamma_1 5$ contained a positive link connecting the nodes C and D (see the positive conditions). However, such a link is simply not there. Still the extension of Γ_{15} that contains $isa(a, E, +)$, call it E, disallows extending it by $isa(a, E, -)$. This happens, though, due to the presence of $isa(a, E, +)$, and not $precl(a, B, E, -)$. Thus, in the case E we have an effect of preclusion without preclusion, just as we did in the Tweety Triangle example before.

However, this is not all there is to say. As a matter of fact, E is not the only extension of Γ_{15}. There is another one — call it E' — that contains the statement $isa(a, E, -)$, among others. It can be acquired if the third rule is applied on $isa(a, B, +)$ and $isax(B, E, -)$ before it is applied on $isa(a, D, +)$ and $isax(D, E, +)$. Then the presence of $isa(a, E, -)$ blocks the inference of $isa(a, E, +)$. It still contains $precl(a, C, E, -)$, but it does not matter. What really matters is this: on the path-based theory Simonet uses, Γ_{15} has one expansion that supports $isa(a, E, +)$, but her rule set S gives two extensions E and E' such that one contains $isa(a, E, +)$ and the other one — $isa(a, E, -)$. This shows that the above result does not hold, implying that something has to be wrong with Simonet's correspondence proof.

[17] Essentially what (2.2) does, as far as I can see, is exactly transferring preclusion from one link to another.

Problems with the two other sets of rules

(Simonet, [15]) contains two other sets of inference rules, and one might think that these do better than S. The first of them — S' — is supposed to implement *off-path preclusion with reinstatement*. In fact, though, it suffers from all of the above problems as much as S does. Besides, S' faces a number of additional difficulties. First, the idea of reinstatement is counter-intuitive. Second (and more importantly), in some examples S' simply does not give us the reinstatement it promises. In particular, the unique extension of Γ_8 discussed above (see Figure 9) contains $precl(a, D, E, +)$, in spite of the fact that $(a, B, C, D, E, +)$ should be reinstated by $(a, B, E, +)$.

Admittedly, the third set of rules S'' fairs somewhat better than the first two, for it faces only two of the problems considered above (namely, the first and the second). However, the notion of preclusion it implements — *preclusion by general subsumption* — has problems of its own. In fact, it is generally agreed that it is too restrictive to be a plausible concept for preclusion, and there are also some well-known cases where it produces counter-intuitive results (see (Horty, [6][pp. 48–51])). Hence, we cannot be content with either S' or S''. No more than we can be with S.

References

[1] Boutilier, C.: Conditional logic of normality: a modal approach. Artificial Intelligence 68, 87–154 (1994)

[2] Boutilier, C.: Conditional logics for default reasoning and belief revision. PhD thesis, Computer Science Department, University of Toronto (1992)

[3] Delgrande, J.: An approach to default reasoning based on a first-order conditional logic: revised report. Artificial Intelligence 33, 63–90 (1988)

[4] Dung, P.M., Son, T.C.: Nonmonotonic inheritance, argumentation and logic programming. Proceedings of the 3rd International Conference on LPNMR pp. 317–329 (1995)

[5] Etherington, D.W., Reiter, R.: On inheritance hierarchies with exceptions. Proceedings of AAAI-89 pp. 104–108 (1983)

[6] Horty, J.F.: Some direct theories of nonmonotonic inheritance. In: Gabbay, D., Hogger, C., Robinson, J. (eds.) Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 3, pp. 111–187. Oxford University Press (1994)

[7] Horty, J.F., Thomason, R.H., Touretzky, D.S.: A skeptical theory of inheritance in nonmonotonic semantic networks. CMU-CS-87-175, Carnegie Mellon University (1987)

[8] Makinson, D., Schlechta, K.: Floating conclusions and zombie paths: two deep difficulties in the "directly skeptical" approach to defeasible inheritance nets. Artificial Intelligence 48, 199–209 (1991)

[9] McCarthy, J.: Circumscription - a form of nonmonotonic reasoning. Artificial Intelligence 13, 27–39 (1980)

[10] McCarthy, J.: Applications of circumscription to formalizing common sense knowledge. Artificial Intelligence 28, 89–116 (1986)

[11] Reiter, R.: A logic for default reasoning. Artificial Intelligence 13, 81–132 (1980)

[12] Sandewall, E.: A functional approach to non-monotonic logic. Proceedings IJCAI-85 pp. 100–106 (1985)

[13] Sandewall, E.: Nonmonotonic inference rules for multiple inheritance with exceptions. IEEE 74, 1345–1353 (1986)

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

[15] Simonet, G.: On Sandewall's paper: Nonmonotonic inference rules for multiple inheritance with exceptions. Artificial Intelligence 86, 359–374 (1996)

[16] Touretzky, D.S., Horty, J.F., Thomason, R.H.: A clash of intuitions: the current state of nonmonotonic multiple inheritance systems. Proceedings of IJCAI-87 pp. 478–483 (1991)

[17] Touretzky, D.S.: The Mathematics of Inheritance Systems. Morgan Kaufmann (1986)

[18] Touretzky, D., Thomason, R.H., Horty, J.F.: A skeptic's menagerie: conflictors, preemptors, reinstaters and zombies in nonmonotonic inheritance. Proceedings IJCAI-91 pp. 478–483 (1991)

[19] Veltman, F.J.M.M.: Defaults in update semantics. Journal of Philosophical Logic 25, 221–261 (1997)

[20] Wang, X., You, J.H., Yuan, L.Y.: A defaults interpretation of defeasible network. IJCAI 1, 156–161 (1997)