More Notes on 'A Clash of Intuitions[9]

R. Al-Asady A. Narayanan
ras@uk.ac.exeter.dcs ajit@uk.ac.exeter.dcs
Computer Science Department
Exeter University
Exeter EX4 4PT
U.K.

Abstract

A path-based inheritance logic is proposed for handling nonmonotonic inheritance structures. The formalism is called *exceptional inheritance logic* (EIL). Exceptional information which is nevertheless inheritable is encapsulated in a class, called an exceptional class. EIL also takes into account a difference between acquired and inheritable properties. The paper describes ways in which exceptional classes are identified and subsequently located in inheritance structures. The on-path versus off-path debate is reassessed using EIL.

1 Introduction

A major problem for network representations of defeasible and default knowledge (called 'nonmonotonic inheritance structures') concerns the use of 'redundant' links. Whilst the use of shortest-path algorithms for deriving network conclusions returns the (intuitively) correct result in many cases, in many other cases it does not [Touretzky, 1984]. One obvious answer is to assert 'redundant' links explicitly into the network representation so that shortest-path algorithms continue to work correctly. Such links represent binary assertions which explicitly relate by one link two nodes which previously were related by two or more intervening links. But the introduction of such 'redundant links' can make previously non-problematic network structures problematic: unambiguous structures can become ambiguous, and there may be side-effects in other parts of the network not involved with the redundant link.

Touretzky [1986] describes a formal mechanism called *on-path preemption* which comes into play when a redundant link causes such problems. The general idea is that where there is a redundant link, another path can override paths containing the redundant link (as will be described later). The effect is to implement formally the basic requirement that, despite the presence of redundant links, more general information should be overridden by *more specific,* i.e. that information from more specific superclasses should take precedence over information from more general superclasses in cases of conflict. However, Sandewall [1986] notes that Touretzky's definitions do not take into account cases where the redundant link is itself interrupted by another node (i.e. where a redundant link contains two or more sublinks). In such cases (exemplified by the so-called *Clyde—African.elephant* structure) counterintuitive results are sometimes returned. Sandewall proposes *off-path preemption*, where the idea that more *explicit* information should win in cases of conflict. Touretzky, Horty and Thomason [1987] present a structure, called the 'George-marine' example, in which off-path preemption also returns counterintuitive results. Unfortunately, in the 'George-marine' example chosen by Touretzky *et al.* and which will be examined later, the intuitive interpretation is only obtained by allowing more general information to override more specific information - - a point often overlooked. This raises the question of how general the on-path preemption principle (that more specific information overrides more general) really is. Whilst Touretzky *et al.* argue that on-path preemption is more intuitive than off-path preemption, some (e.g. [Etherington, 1987]) remain neutral, whilst others (e.g. [Neufeld, 1991]) prefer the more permissive off-path preemption approach.

2 Exceptional inheritance reasoning

This paper proposes a way of representing and handling nonmonotonic inheritance structures which supports the principles of both on-path and off-path preemption (that more general information is overridden by more specific, and explicit information overrides derived). The formalism is called *exceptional inheritance logic* (EIL) [Al-Asady and Narayanan, 1992]. Exceptional inheritance reasoning depends on distinguishing between (a) *typical* and *exceptional* classes, and (b) *inherited* and *acquired* properties.

2.1 Typical and exceptional classes

A typical class is one which passes information to its subclasses or instances by default, as is the case with standard class specialization. If the inheritance structure consists solely of typical classes linked by *isa* links, there is little problem in identifying the typical properties to be inherited by instances. But a class may also be typical by virtue of providing exceptions to more general typical classes. Such a class is an *exceptional class.*

In the network depicted in Figure 1, *Clyde,* whilst *elephant* (through two *isa* (⟶) links), is *Royal.elephant*

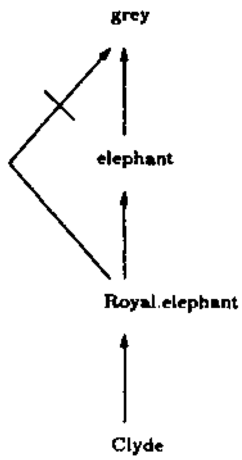

Figure 1: The Clyde—RoyalElephant network

and so (typically) not grey. In this example elephant is a typical class in that elephants typically are grey. Royal.elephant, though, contains an exception to the typical class to which it belongs (elephant). The E1L interpretation here is that Royal.elephant inherits the property of typical non-greyness from an exceptional class (Ψ) for which non-greyness is typical. Information which is exceptional to the typical class and which belongs to the exceptional class is inherited as typical by subclasses and instances of the exceptional class, unless overridden.

The reasoning behind this interpretation is as follows. In the current literature the tsa links are used to specify superclass relationships as well as infer inheritance relationships in the typical class case. In exceptional cases the not-isa link specifies a non-superclass relationship but nevertheless requires a property to be inherited, namely, the opposite of what would be typically inherited. This is questionable: if isa is used to specify class relationships as well as to infer inheritance relationships, then not-isa should be used for explicitly ruling out certain class relationships (therefore stronger than no links at all) as well as prohibiting inheritance of any sort. However, not-isa is usually used to convey the information that the negation of the property referred to by a class name is to be inherited. The proposed exceptional class (Ψ) is meant to address this point explicitly, by allowing the opposite or negation of what would be typically inherited to be inheritable in its own right. This would leave not-isa to perform the simple but important task of explicitly ruling out certain subclass/superclass relationships, without necessarily being involved in aspects of inheritance.

EIL's approach here is to create an exceptional class which can be inserted into the structure so that the following new path results:

$$Clyde \rightarrow RoyalElephant \rightarrow \Psi \not\rightarrow grey$$

where Ψ is the exceptional class which is just like the elephant class, except that it is a subclass to a class (non-grey things) which stands for the opposite of what elephants typically are a subclass of (grey things). Mechanisms for locating Ψ in inheritance structures will be provided later.

2.2 Acquired and inheritable properties

There is also a distinction between acquired properties, which are properties that an object picks up from the environment, and inheritable properties, which are properties that an object can inherit from its class(es). The underlying rationale is that an inheritable property has the same meaning no matter by what or whom it is inherited, whereas an acquired property, since it is acquired by an object or class from the environment, cannot be passed by inheritance to other objects/classes without assuming that those objects/classes have also acquired this property from the environment. Since this assumption cannot always be guaranteed to hold, acquired properties can only be attached to objects/classes if it is asserted that they have acquired that property from the environment.

3 Conceptual foundations of multiple inheritance with exceptions

The following definitions will make the assumptions and motives of EIL clear. First, multiple inheritance, and inheritable and acquired properties, are defined.

Definition 1 Multiple inheritance is the heredity of one or more properties by an individual or by a class, from two or more classes, either of which can be typical or exceptional.

Definition 2 An inheritable property is a property that may pass via a class to a subclass or individual.

Definition 3 An acquired property is a property derived from experience and training or as a result of interaction with the environment. This kind of property cannot be inherited by subclasses or individuals of a class which has an acquired property.

Next, exceptional class is defined.

Definition 4 An exceptional class (called Ψ) for an individual or class X is a typical superclass (if X is a class) or typical class (if X is an individual) created in order to explain inconsistencies (involving inheritable properties only) arising from multiple inheritance with exceptions. It is similar in all respects except one to another typical superclass or typical class (let us say Y) of X. That is, Ψ carries with it everything that Y does except that, m the case of one property that Y carries, Ψ carries as inheritable the opposite of that property.

The basic rule of inference in EIL is defined as follows.

Definition 5 Given Ψ, Y and X, where X is a subclass of both Ψ and Y, X will inherit from Ψ rather than from Y. Also, subclasses of, and individuals belonging to, X will inherit from Ψ rather than Y.

That is, Ψ has more explicit information than Y. The exceptional information is more explicit because what was previously implicitly inherited by means of a not isa link (e.g. *Royal.elephant* $\not\rightarrow$ *grey*) is now explicitly encapsulated in a class of its own so that the exceptional information becomes typically inheritable in its own right. Also, the exceptional information prevents more general information overriding more specific, as will be shown later.

4 A semi-formal introduction to exceptional inheritance

Let

- x represent an object,
- α represent x's typical class or superclass which has an *inheritable* property,
- β be x's class or superclass which has an *acquired* property (β is called 'acquired class' in the definitions that follow),
- δ be the inheritable property of α,
- ζ be an exceptional class which is just like α except for the inheritable property with which it conflicts,
- σ is the property in respect of which α and ζ differ (i.e. the exceptional property),
- γ stand for the conjunction of $\neg\sigma(x)$ and $\neg\delta(x)$ (neither exceptional nor typical inheritable),
- $\odot$ stand for 'provided that',
- $\otimes$ signify an inconsistency or conflict between two classes with respect to the inheritable property,
- $\triangleright$ stand for a *class similarity* relationship between two classes of an object, *except for* one property (σ) on which they differ, and
- $\bowtie$ stand for a *bidirectional* relationship between Ψ and α, both of which are immediate (super)classes of x.

The general form of an inheritance rule in EIL is as follows:

$$\Gamma \odot \Delta \rightarrow \Theta$$

which stands for: 'If Γ can be derived and provided that the exceptions Δ have been tested for, then derive Θ.' There are also rules of the form

$$\odot\Delta \rightarrow \Theta$$

where inheritance conclusions require only the exceptions to be tested for. The exception Δ consists of a σ (an exceptional property) or a γ (an acquired property), or both. If an element of Δ is positive, the interpretation is: 'provided that the exception holds', and if an element of Δ is negative: 'provided that the exception doesn't hold'.

Three primitive relations (where $\phi(x)$ has the interpretation of class or object x belonging to ϕ where ϕ stands for a class) are now described. First,

$$PR1 \qquad \zeta(x) \triangleright \alpha(x)$$

i.e. x's exceptional class (if there is one) is exactly like a typical superclass α of x except for an exceptional property σ in respect of which α and ζ differ (Definition 4). Then,

$$PR2 \qquad \alpha(x) \odot \neg\zeta(x) \rightarrow \delta(x)$$

and

$$PR3 \qquad \odot\,\zeta(x) \not\rightarrow \delta(x)$$

i.e. if x belongs to a typical class and provided that x does not belong to an exceptional class, then x inherits the inheritable property δ of typical class α ($PR2$); and provided that x belongs to an exceptional class then x does not have the typical inheritable property δ of α ($PR3$).

An *exception indicator rule* is represented as follows (bearing in mind that ζ stands for a class which contains the negation of some inheritable property δ):

$$EIL1a \qquad \alpha(x) \odot \zeta(x) \rightarrow \otimes\delta(x)$$

The following definition describes this rule:

Definition 6 *If x is an α and provided that it belongs to an exceptional class ζ (where $\zeta(x) \triangleright \alpha(x)$ by $PR1$ above), then there is a conflict between these classes regarding the typical inheritable property δ.*

Similarly, an *acquired property* conflict is defined as follows:

$$EIL1b \qquad \alpha(x) \odot \beta(x) \rightarrow \otimes\delta(x)$$

where the conflict concerns whether a typical inheritable property or an acquired property should be inherited. To draw the right inheritance conclusion:

Definition 7 *If x has typical superclass α, and provided that x is neither an instance of an exceptional class ζ nor an instance of an acquired class β, then x inherits property δ.*

The above definition is formalized by:

$$EIL2 \qquad \alpha(x) \odot [\neg\zeta(x) \wedge \neg\beta(x)] \rightarrow \delta(x).$$

This leads to the following exception rules:

Definition 8 *If x has typical superclass α, and provided that x is an instance of an exceptional class ζ but not an instance of the acquired class β, then x inherits the exceptional property σ.*

The rule can be represented as follows:

$$EIL3 \qquad \alpha(x) \odot [\zeta(x) \wedge \neg\beta(x)] \rightarrow \sigma(x),$$

Definition 9 *If x has typical superclass α, and provided that x is an instance of the acquired class β but not an instance of the exceptional class ζ, then x has the acquired property by assertion.*

The above definition can be represented as follows:

$$EIL4 \qquad \alpha(x) \odot [\beta(x) \wedge \neg\zeta(x)] \rightarrow \gamma(x).$$

Then,

Definition 10 *If x has typical superclass α, and provided that x is an instance of an exceptional class ζ and also an instance of the acquired class β, then x has the acquired property by assertion.*

The above definition can be represented in a rule as follows:

$$EIL5 \qquad \alpha(x) \odot [\beta(x) \wedge \zeta(x)] \rightarrow \gamma(x)$$

In other words, the acquired property overrides an inheritable property. This is because the acquired property is an explicit property (declared and asserted by the user), whereas the inheritable property is a typical (implicit) one.

Finally,

Definition 11 *If x has immediate (super)classes Ψ and α, then $\Psi \bowtie \alpha$, i.e. Ψ and α are bidirectionally related in that they conflict over δ. The bidirectional link cannot be used for drawing inheritance conclusions.*

The above definition can be represented in a rule as follows:

$$EIL6 \qquad \alpha(x) \odot \Psi(x) \rightarrow \alpha \bowtie \Psi$$

Bidirectionality (signified by $\leftrightarrow$ in the networks below) is meant to express the possibility of mutual specialization between a typical class and its related exceptional class. That is, just as an exceptional class is similar to a typical class except for an exceptional property, a typical class can be regarded as similar to an exceptional class except for a typical property.

5 The Clyde—Royal Elephant problem revisited

Consider the following primitive paths:

$$Clyde \rightarrow elephant,$$
$$elephant \rightarrow grey.$$

These paths are depicted in Figure 2. *Clyde* is neither a member of an exceptional class nor a member of a class which has an acquired property, using $EIL2$:

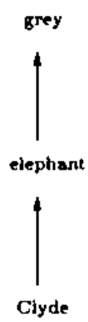

Figure 2: The *Clyde—grey* structure

$$elephant(Clyde) \odot [\neg\zeta(Clyde) \wedge \neg\beta(Clyde)] \rightarrow grey(Clyde).$$

If the following primitive paths are included (giving us the structure depicted in Figure 1):

$$Clyde \rightarrow Royal.elephant,$$
$$Royal.elephant \not\rightarrow grey,$$
$$Royal.elephant \rightarrow elephant,$$

first a check is made as to whether there is any conflict with the previous assertions, using $PR2$ and $PR3$:

$$Elephant(Clyde) \odot \neg Royal.elephant(Clyde) \rightarrow grey(Clyde)$$
$$\odot Royal.elephant(Clyde) \not\rightarrow grey(Clyde)$$

These primitive relations identify the conflict:

$$elephant(Clyde) \odot Royal.elephant(Clyde) \rightarrow \otimes grey(Clyde)$$

The condition, exception and conclusion in the conflict rule are used to locate Ψ, namely, an exceptional, immediate superclass of *Royal.elephant* (the exception in the rule). The following two primitive paths are added:

$$Royal.elephant \rightarrow \Psi,$$
$$\Psi \not\rightarrow grey.$$

Ψ is the exceptional class which has the missing, complementary aspect regarding the typical superclass property, and its negative link with that property, namely, $\Psi \not\rightarrow grey$, signifies this. It follows by $EIL6$:

$$\Psi \bowtie elephant.$$

Figure 3 shows the EIL representation of the structure. There is a new path:

$$Clyde \rightarrow Royal.elephant \rightarrow \Psi \not\rightarrow grey$$

Using $EIL3$:

$$elephant(Clyde) \odot [Royal.elephant(Clyde) \wedge \neg\beta(Clyde)] \not\rightarrow grey(Clyde)$$

Since there are no acquired properties (γ) to consider, the inheritance conclusion is $\neg grey(Clyde)$.

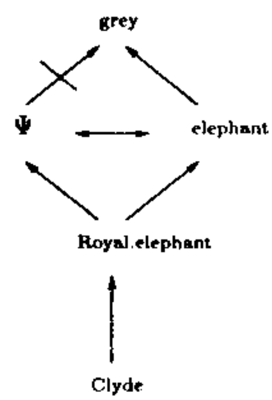

Figure 3: The *Clyde—Royal.elephant* structure represented using EIL

6 On-path preemption revisited

The application of EIL to the on-path versus off-path debate can now be demonstrated. On-path preemption applies if there is redundancy in the inheritance network. Consider Figure 4 which is just like Figure 1 but now contains a redundant link *(Clyde —▶ elephant)* which can be derived from two primitive links already in the structure: *Clyde —▶ Royal .elephant* and *Royal.elephant —▶ elephant.*

There are now three paths starting at *Clyde:*

(a) *Clyde —▶ Royal .elephant* $\not\rightarrow$ *grey*

(b) *Clyde —▶ elephant —▶ grey*

(c) *Clyde —▶ Royal.elephant —▶ elephant —▶* prey

Path (c) is pruned on the basis that it contains a node *(Royal.elephant)* from which there is a link to another

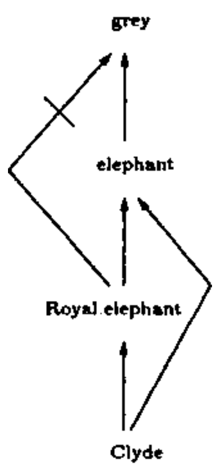

Figure 4: Inheritance network with redundancy

node which occurs later in the same path *{grey)* and is of opposite polarity. *Royal.elephant* and the information it contains is therefore more specific to *Clyde* than the nodes which follow *Royal.elephant* in path (c), i.e. *elephant.* Also, path (b) is pruned because there is another path (path (a)) in which there is a node *(Royal.elephant)* — called an *on-path intermediate* — from which starts a subpath of opposite polarity to that of (b). Since there is a link between the on-path intermediate and a node *(elephant)* in path (b), path (a) contains more specific information despite the redundant link which reduces both paths to equal length. Path (b) is therefore pruned to leave path (a). However, the notion of on-path intermediate crucially depends on the existence of a redundant link which bypasses that on-path intermediate. Sandewall [1986] presents the structure shown in Figure 5 to point out the shortcomings of on-

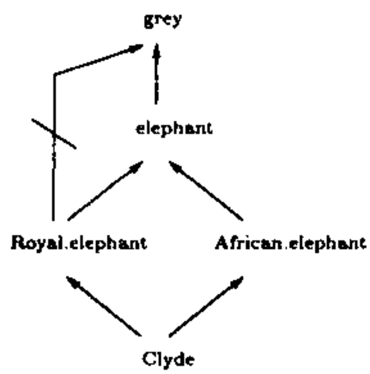

Figure 5: *Clyde* the *Royal, African.elephant*

path preemption. The redundant link in the structure in Figure 4 between *Clyde* and *elephant* is interrupted by the *African.elephant* node, thereby resulting formally in the loss of the redundant link. On-path preemption would then not be applicable, and given a skeptical approach nothing could be inferred about *Clyde's* colour.

Sandewall's point is that the *Clyde* $\not\to$ *grey* conclusion should nevertheless go through, since this supports our intuition that *African.elephant* doesn't affect that part of the network dealing with *Clyde*'s greyness. Since one of *Clyde*'s two immediate classes has some explicit information concerning greyness, this should be preferred to derived information in cases of conflict. Sandewall's proposal of off-path preemption was intended to build this preference for explicit information directly into the inferencing mechanism. However, Touretzky *et al.* [1987] argue that this approach is less intuitive than on-path, citing the structure in Figure 6 as an example of where

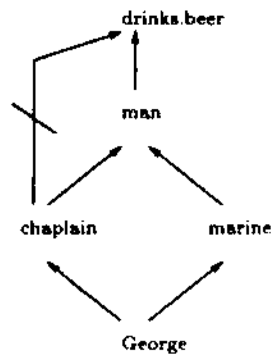

Figure 6: *George* the *chaplain/marine*

the analogous conclusion that *George* does not drink beer is not intuitively supported. The structures depicted in Figures 5 and 6 have the same topology, however. As stated earlier, the fact that the intuitive reading in the *George* structure seems to require more general information to override more specific information (i.e. information attached to *man* (*drinks.beer*) to override information attached to *chaplain* in the case of *George*) questions the claimed usefulness of preemption generally.

The EIL view is that the two structures above are not of the same topology at a deeper level when analyzed with EIL. Consider the *Royal.elephant* example again (Figure 5). Following the same analysis as was made for the *Clyde—Royal.elephant* example (Figures 1, 3), the resulting structure is depicted in Figure 7. *Royal.elephant* will inherit non-greyness through the exceptional class, just as before, as will *Clyde*. This is because Ψ contains more explicit information, and classes (or individuals of a class) containing exceptional information will be preferred to non-exceptional classes in cases of inheritance conflict. *African.elephant* will inherit greyness through *elephant* $\to$ *grey*. The 'path' *African.elephant* $\to$ *elephant* $\to$ Ψ $\not\to$ *grey* is no inheritance path at all, since the bidirectional relationship between *elephant* and Ψ is not an inheritance relationship (Definition 11).

Now consider again the *George* example (Figure 6). In the *George* structure, both the classes *marine* and *chaplain* acquire their property of *drinks.beer* from the environment. It is the explicitly acquired properties which have the last word in this structure (by EIL5 and EIL6), not the implicit inheritance properties from the

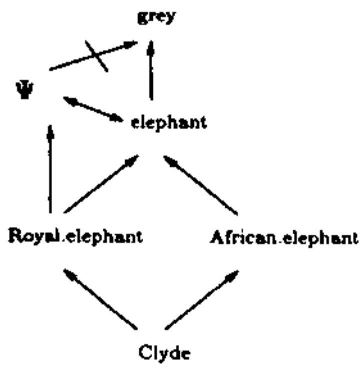

Figure 7: The EIL representation of the Clyde—African.elephant structure

class man. Therefore, since the acquired property of drink8.beer is not associated with either chaplain or marine directly in this structure, nothing can be directly inferred about George. There are then options as to where drinks.beer should be linked in as an acquired property (dashed lines in Figure 8). The property of

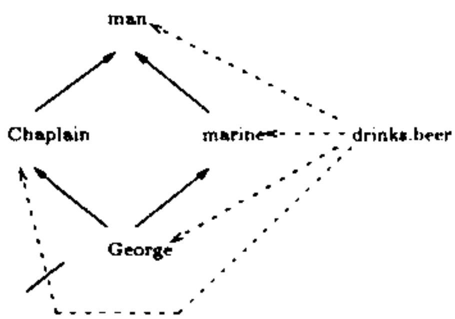

Figure 8: The George example represented in EIL

drinks.beer (or its negation) can be explicitly located with man, chaplain, marine, or George, or with more than one (since acquired properties are not inheritable). The EIL representation of the problem structure is that this is straightforward multiple inheritance, with no inheritable exceptions. If it is claimed that drinks.beer is indeed an inheritable property of man (just as grey is of elephant), this must be argued for. And that argument is a different one from the on-path versus off-path debate. Importantly, the EIL representation of the George structure removes the implication that in certain structures more general information is to be preferred to more specific.

References

[Al-Asady and Narayanan, 1992] R. Al-Asady and A. Narayanan. EIL: An exceptional approach to nonmonotonic reasoning. Technical Report 244, University of Exeter Department of Computer Science, 1992.

[Etherington, 1987] D. Etherington. More on inheritance hierarchies with exceptions: Default theories and inferential distance. In Proceedings of the American Association for Artificial Intelligence, pages 352-357, 1987.

[Neufeld, 1991] E. Neufeld. Notes on 'A clash of intuitions'. Artificial Intelligence, 48(2):225-240, 1991.

[Sandewall, 1986] E. Sandewall. Nonmonotonic inference rules for multiple inheritance with exceptions. Proceedings of the IEEE-86, 74(10):1345~1353, 1986.

[Touretzky et ai, 1987] D. S. Touretzky, J. F. Horty, and R. H. Thomason. A clash of intuitions: The current state of nonmonotonic multiple inheritance systems. In Proceedings of the Tenth International Joint Conference on Artificial Intelligence, pages 476-482, 1987.

[Touretzky, 1984] D. S. Touretzky. Implicit ordering of defaults in inheritance systems. In Proceedings of the American Association for Artificial Intelligence, pages 322-325, 1984.

[Touretzky, 1986] D. S. Touretzky. The Mathematics of Inheritance. Pitman Research Publications, London, UK, 1986.