A Coherence-Based Approach to Default Reasoning

Salem BENFERHAT - Laurent GARCIA
IRIT, CNRS-Université Paul Sabatier, 118, route de Narbonne
31062 TOULOUSE Cedex - France
email: {benferhat, garcia}@irit.fr

Abstract. In the last 15 years, several default reasoning systems have been proposed to deal with rules having exceptions. Each of these systems has been shown to be either cautious (where some intuitive conclusions do not follow from the default base), or adventurous (some debatable conclusions are inferred). However, the cautiousness and the adventurous aspect of these systems are often due to the incomplete way of describing our knowledge, and that plausible conclusions depend on the meaning (semantics) assigned to propositional symbols. This paper mainly contains two parts. The first part discusses, with simple default bases (where the used symbols have no a priori meaning), which assumptions are assumed when a given conclusion is considered as intuitive. The second part investigates a local approach to deal with default rules of the form "generally, if α then β" having possibly some exceptions. The idea is that when a conflict appears (due to observing exceptional situations), we first localize the sets of pieces of information which are responsible for conflicts. Next, using a new definition of specificity, we attach priorities to default rules inside each conflict. Lastly, three proposals are made to solve conflicts and restore the consistency of the knowledge base. A comparative study with some existing systems is given.

1. Introduction

One of the most important problems encountered in knowledge based systems is the handling of exceptions in generic knowledge. A rule having exceptions (also called default rule or conditional assertion) is a piece of information of the form "generally, if α is believed then β is also believed", where α and β are assumed here to be propositional logical formulas. In the presence of a fact, one may jump to conclusions which are just plausible and can be revised in the light of a new fact.

Several proposals have been done to deal with default information. Gabbay (1985), Lehmann and his colleagues (Kraus et al., 1990; Lehmann and Magidor, 1992) and Gärdenfors and Makinson (1994) have tried to provide basic set of properties (or postulates) that a default reasoning system should satisfy. However, inference machinery emerged from these proposals are either very cautious (like System P of Kraus et al. (1990)), or adventurous (like the rational closure (Lehmann and Magidor, 1992)). This means that these postulates are not enough to define a default reasoning system which correctly provides intuitive conclusions. In our opinion, what is missing is to characterize which assumptions are assumed when a given postulate is applied or when a given conclusion is derived. For instance, even if it is agreed that the contraposition rule (i.e., "normally, from α infer β" implies that "normally, from $\neg\beta$ infer $\neg\alpha$") is not a desired postulate, we believe that there are particular cases where this rule can be applied. The first part of this paper (Section 2) tries to explain, with sample default bases, what "intuitive" conclusion means, and what are the implicit assumptions done to derive plausible conclusions.

The second part of the paper (Section 3) investigates a local and coherent-based approach to deal with default information. This approach is based on restoring the consistency of the default base due to observing exceptional situations. Contrary to some existing systems, like System Z of Pearl (1990), we neither proceed to a global handling of conflicts, nor to a global ranking of the defaults rules. Independent conflicts (sets of pieces of information which are responsible for the inconsistency of the database) are handled separately, and hence inferring unwanted conclusions is avoided. Moreover, priorities, based on a new definition of specificity, are attached to default rules inside each conflict. We propose three ways to solve conflicts. One of these ways (the interesting one) is compared with some existing default reasoning systems.

2. What "Intuitive" Conclusions Mean?

The cautiousness and the adventurous aspect of default reasoning systems are often due to the incomplete way of describing the knowledge, and that plausible conclusions depend on the meaning assigned to propositional symbols. Two different meaning assignments to propositional symbols in the same set of rules can lead to different sets of intuitive conclusions. Indeed, when we give a precise meaning to the symbols in a default base, we are leaning to infer conclusions which are based on general information about the real world that are not explicitly mentioned in the default base. We will illustrate this situation in the example of Figure 1. Clearly, one of the important problems in the common sense reasoning remains how to correctly write our knowledge.

In this section, we try to explain with simple default bases (where the used symbols have no a priori meaning) which assumptions are assumed when a given conclusion is considered as intuitive. This section is not intended to give a detailed discussion of the meaning of default rules and ways to deal with them but just to point out, with some examples, what implicit assumptions should be made to do inferences in the presence of incomplete information.

In the following, $\mathcal{L}$ denotes a finite propositional language. Formulas of $\mathcal{L}$ are denoted by Greek letters α, β, δ,... $\top$ represents tautology(ies), and $\perp$ denotes any inconsistent formula. We write a default rule "generally, if α then β" as $\alpha \rightarrow \beta$. The implication "$\rightarrow$" is a *non-classical* arrow, and it should not be confused with material implication ($\neg\alpha \vee \beta$). Default rules considered here are not general default rules in the sense of Reiter (1980), but more a sort of normal default rules or conditional assertions in the sense of Kraus & al. (1990). A *default base* is a set $\Delta = \{\alpha_i \rightarrow \beta_i, i=1,...,n\}$ of default rules.

Before starting the analyzing of some default bases, we consider the following principle as fundamental:

Auto-deductivity principle: For any default $\alpha \rightarrow \beta$, if we observe α and only α then β must be considered as intuitive.

Example 1: Let us consider a simple default base containing only one default rule $\Delta = \{x \rightarrow y\}$. Given an observed fact $\neg y$, we are interested to know if x (resp. $\neg x$) follows or not.

• If we accept *x* then this means that the observation of $\neg y$ is due to the existence of an exceptional x. For us this conclusion is non-intuitive and disagrees with the idea of default reasoning where it is preferred to believe in normal situations rather than in the exceptional ones. Moreover, we can check that any possibility distribution where $N(y|x)>0$ and $N(x|\neg y)>0$ implies $\Pi(y)>\Pi(\neg y)$ which means that we have a prior default information "generally, y is accepted" (see (Dubois & al., 1994) for a complete overview of possibility theory). This is of course non-intuitive. As far as we know, there is no existing system which infers x.

• Now if we assume that $\neg x$ is plausible then this means that the contraposition rule (from $\alpha \to \beta$ deduce $\neg\beta \to \neg\alpha$) is accepted. However this conclusion is intuitive, since $\neg y$ does not contradict our knowledge base. Moreover accepting $\neg x$ means that the rule $x \to y$ has a very small number of exceptions. Here, we do not have any prior default information that is implied from all possibility distributions where $N(y|x)>0$ and $N(\neg x|\neg y)>0$. Some systems like System Z (1990), Geffner's conditional entailment (1992) infer such a conclusion.

• If *neither x nor* $\neg x$ is inferred, then this means that we do not decide between the two previous cases. This is a very cautious attitude because we take into consideration the two possibilities which make sense since the two following default bases $\Delta_1=\{x \to y, \neg y \to x\}$ and $\Delta_2=\{x \to y, \neg y \to \neg x\}$ are consistent. In a probability theory, it is both possible to construct a probability distribution where $P(y|x)>0.5$ and $P(x|\neg y)>0.5$ and another where $P(y|x)>0.5$ and $P(\neg x|\neg y)>0.5$. Some systems, like Reiter's default logic (1980) or System P of Kraus et al. (1990), prefer to choose this attitude.

In the approach that we will develop in Section 3, we consider, from the rule $x \to y$ and a fact $\neg y$, that the conclusion $\neg x$ is plausible, and more generally we will assume the following consistency principle:

Consistency principle: When an observed fact does not contradict the default base, then plausible conclusions are exactly the same as the ones of classical logic; hence, all the properties of the classical logic are considered valid.

The following examples consider the case when adding a fact contradicts the default base.

Example 1 (continued): We add a second default rule $(y \to z)$ to the base of Example 1 and obtain $\Delta = \{x \to y, y \to z\}$.

$$x \longrightarrow y \longrightarrow z$$

Figure 1

Given a fact $x \wedge \neg z$, we are interested to know if y can be inferred or not? Clearly, adding this fact leads to an inconsistent knowledge base which means that at least one of these two rules should not be applied. Let us consider two cases:

• From $x \wedge \neg z$, we prefer inferring $\neg y$. This intuitively means that the rule $x \to y$ meets more exceptions than the rule $y \to z$ (namely, we prefer to give priority to $y \to z$ rather than to $x \to y$). Indeed, if $y \to z$ accepts a very small number of exceptions then we are leaning to apply the "contraposition rule". Intuitive situation of this case can be obtained by interpreting *x* by "students", *y* by "young" and *z* by "have a special rate for travelling". The fact "Tom is a student and has no special rates for travelling" leads intuitively to prefer concluding "Tom is not young". The reason is that the rule "young people have special rates for travelling" accepts a very small number of exceptions, while it is not very surprising to find some students which are not young. Note that using the possibilistic logic machinery, we can check that from $(\neg x \vee y,\ a)$,

$(\neg y \vee z,\ b), (x \wedge \neg z,\ 1)$ with $0<a<b\leq 1$ the conclusion $\neg y$ follows. As far as we know, there is no system which infers $\neg y$ from $x \wedge \neg z$.

• If from $x \wedge \neg z$, we prefer to infer y then applying the contraposition rule to $y \rightarrow z$ in this case can be debatable. In Figure 1, let us interpret x by "Danish people", y by "tall people" and z by "play basket-ball". In the presence of Bjarne Riis (who is Danish and does not play basket-ball), one would like to infer that he is a tall person. The reason is that it is less surprizing to find a tall person which does not play basket-ball rather than to find a small Danish person. Similarly to the precedent case, in possibilistic logic, from $(\neg x \vee y,\ a), (\neg y \vee z,\ b), (x \wedge \neg z,\ 1)$ with $0<b<a\leq 1$ we can check that the conclusion y follows. A probabilistic interpretation where y is inferred from $x \wedge \neg z$ can be the following. Let denote by $|\alpha|$ the number of complete situations (e.g. classical interpretations) where α holds. Let us assume that the two default rules are such that $|x \wedge y|/|x|=|y \wedge z|/|y|=a$ with $a>0.5$. When there is no information about the proportion of α being β, we assume that $|\alpha\beta|=|\alpha\neg\beta|$. Then, using these *strict* assumptions, we conclude that $|xy\neg z|=a.(1-a)>|x\neg y\neg z|=(1-a)/2$, since $a>0.5$, and hence, the conclusion y is preferred. An example of default reasoning system where y is inferred from $x \wedge \neg z$ is Reiter's default logic.

Of course there is also a cautious attitude where neither y nor $\neg y$ is inferred. Most of the existing systems (e.g. System P (Kraus & al., 1990) , System Z (Pearl, 1990), Boutilier's approach (1992), ε-belief functions (Benferhat et al., 1995)) adopt this attitude.

The previous example shows that the meaning assignment to propositional symbols can alter our decisions. It also illustrates the need of providing explicit information to get a desired conclusion. This can be done by adding some ordering between default rules. In (Benferhat & al., 1994) a possibility theory-based encoding of pieces of information of the form "in the context α, δ has no influence on β" has been suggested as a way to express explicit information. In the previous example, adding the explicit independence information "in the context x, z (resp. $\neg z$) has no influence on y" leads to infer y in the presence of $x \wedge \neg z$ in the possibility theory framework, while without this independence information neither y nor $\neg y$ is inferred in this framework.

The other reason why y should be preferred to $\neg y$ in Figure 1 is to use the specificity principle. Indeed the rule $x \rightarrow y$ is preferred to the rule $y \rightarrow z$. Specificity (Touretzky, 1984 ; Simari and Loui, 1992) is at the core of defeasible reasoning, and it is modelled by some preference relation, generally extracted from the syntax of the knowledge base, which guarantees that results issued from sub-classes override those obtained from super-classes. The following well-known triangle example illustrates the specificity principle:

Example 1 (continued): Now, let us expand the example of Figure 1 by adding the rule $x \rightarrow \neg z$. So, we have $\Delta = \{x \rightarrow y, y \rightarrow z, x \rightarrow \neg z\}$.

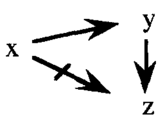

Figure 2

This is the famous penguin example when x is interpreted by “penguin”, y by “birds” and z by “fly”. Given an observed fact x then it is well agreed that y and ¬z are plausible conclusions of Δ. The main argument used to justify these conclusions is that the class x is more specific than the class y. Figure 2 in some way is similar to Figure 1 when we have a fact $x \wedge \neg z$. The difference is that in Figure 1, ¬z is observed, while ¬z is deduced in Figure 2. However we have no problems to conclude y in Figure 2 while this conclusion in general does not follow in Figure 1! Let's notice that Reiter's default logic doesn't allow to infer neither z nor ¬z but the principle of specificity is inherent to most of the inference relations devoted to default reasoning.

Example 1 (continued): Irrelevance

Now let us again consider the example of Figure 2. We assume that the language contains four propositional symbols x, y, z and v (where v does not appear in the default base). We are interested to know if from a fact *x and v* we get ¬z or not? Some systems, like System P of Kraus et al. (1990), do not allow to infer neither z nor ¬z. This cautious attitude is justified by the fact that we have no information to decide if "x and v" is an exceptional x with respect to the property "¬z" or not. This problem is known as the irrelevance problem. In the following, we consider the next principle as valid:

Irrelevance Principle: Let δ be a propositional formula composed of propositional symbols which do not appear in the default base; if some conclusion ψ is a plausible consequence of a given fact ϕ w.r.t. Δ, then ψ is also a plausible consequence of a given fact $\phi \wedge \delta$ w.r.t. Δ.

Reiter's default logic, System Z and its extensions have no problem to satisfy the irrelevance problem.

Let us study now the importance of specificity beside the length of the path of deduction. For this, let us consider the example given in the following Figure 3: $\Delta=\{a1 \rightarrow a2, a2 \rightarrow a3, ..., an\text{-}1 \rightarrow an, a1 \rightarrow \neg an\}$.

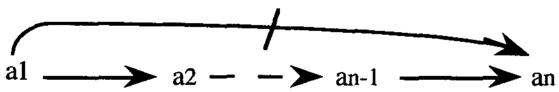

Figure 3

We want to know what must be deduced from a1. We have yet discussed, in Figure 2, the case where n=3. Let now consider the case where n=4 and let us give two interpretations to propositional symbols.

A meaning that can be applied to this case is again the penguin example where a1 is interpreted by "penguin", a2 by "bird", a3 by "has feathers" and a4 by "fly". Here, we want to deduce that a penguin has feathers. This solution follows the idea of Touretzky (1984) or Moinard (1987), where they justify this conclusion by applying the specificity principle. Another meaning is to consider that a3 is interpreted by "moving in the sky" (the other variables have the same meaning). Here, we want to deduce that a penguin does not move in the sky (since, intuitively, it is very surprising to find a non-flying object moving in the sky). Of course, there is a cautious alternative where neither an-1 nor ¬an-1 is inferred. That's the choice of System Z or Geffner's conditional entailment.

In both situations, we are using the knowledge of the real world to decide what conclusions we want to infer (here, we implicitly use the properties of penguins). This example can be seen as an extension of the example of Figure 1 but here the facts a2 and a4 are not given but deduced. So, it should be natural to keep on deducing the same results.
We must notice that some systems like Pearl's System Z (1990) or Geffner's system (1992) do not deduce an-1 (nor a3, ..., an-2) from this base. The reason is that their definition of specificity is based on the notion of contradiction between rules i.e. a rule is specific if, when it is verified, there is an inconsistency due to the presence of other rules. But these definitions do not take into account that specificity can also exist between classes when no exception explicitly holds. If we take the base $\Delta=\{A\rightarrow B, B\rightarrow C\}$, the two rules are not conflicting and are in the same level even if $A\rightarrow B$ is more specific than $B\rightarrow C$. Here, the notion we advocate is rather to consider that a class is more specific when it concerns a subset of properties w.r.t. a certain set with more properties (the rules are more precise if they are applied to the subset). This idea does not take into account the notion of exceptionality and, in fact, refines it since, when a subclass is exceptional, it concerns a subset of another set (which is the general class).

3. A Local Approach to Deal with Default Information

3.1. Why a Local Approach?

We have seen in the previous section that the inference of debatable conclusions can be due to some reasonable assumptions done by default reasoning systems in the absence of complete information. However some systems, like System Z of Pearl(1990) (or equivalently the rational closure of Lehmann and Magidor (1992)), infers debatable conclusions that are due to the global handling of inconsistency appearing when we learn some new fact. To illustrate this situation, let us consider the following example:
Example 2 (ambiguity): Let $\Delta = \{q\rightarrow p, r\rightarrow \neg p\}$. This is the famous Nixon diamond example: "Republicans normally are not pacifists" and "Quakers normally are pacifists". Clearly, from q we prefer to infer p, and from r we prefer to infer $\neg$p and this is justified by the auto-deductivity principle. But however, if we are interested to know if Nixon, who is a Republican and a Quaker, is a pacifist or not, then we prefer to say nothing. This is intuitively satisfying. Now let us add to this example three further rules (not related to pacifism), which give more information about Quakers: "Quakers are Americans", "Americans like base-ball" and "Quakers do not like base-ball". So, let $\Delta = \{q\rightarrow p, r\rightarrow \neg p\} \cup \{q\rightarrow a, q\rightarrow \neg b, a\rightarrow b\}$. This is illustrated by the following Figure 4:

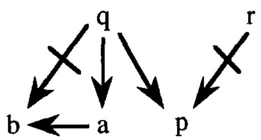

Figure 4

Given a fact $q \wedge r$, we have two conflicts, one $A = \{r \wedge q, q \to p, r \to \neg p\}$ is related to the pacifism and the other $B=\{q, q \to a, q \to \neg b, a \to b\}$ is related to playing base-ball. These two conflicts are independent and, hence, from $q \wedge r$ we prefer to infer $a, \neg b$ but neither p nor $\neg p$ is intuitively considered as a plausible conclusion. Some systems like System Z, prefer to infer p rather than $\neg p$. The reason is that the two conflicts are not handled independently but globally, and hence, in System Z, for the conflict B the class q is a specific class and is given a higher priority over all the general classes especially over the class r, therefore the debatable conclusion p is inferred!

There is another reason where handling default information locally is recommended. It concerns the case of default bases containing cycles. Let's take the example of the base $\Delta=\{A \to B, B \to C, C \to A\}$. It is illustrated by the following Figure 5.

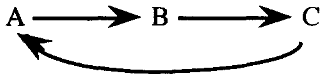

Figure 5

A global treatment of rules does not allow to make a difference between the rules and put the three rules in the same level. This is intuitively satisfying. Considering the fact $A \wedge \neg C$, a global treatment does not allow to infer B. With a local treatment, from the fact $A \wedge \neg C$, only the two rules $A \to B$ and $B \to C$ are involved in the inconsistency and we do not take into account the rule $C \to A$ (namely, $C \to A$ is not activated). Priority is given to the rule $A \to B$ upon $B \to C$ using the specificity principle. Then, B can be inferred.

3.2. Handling Stratified Conflicts

In the example of Figure 4, we have seen that, when it is possible, conflicts must be handled independently and that defaults must be ranked locally inside each conflict. A same default can be involved in different conflicts. This section gives a formal definition of conflicts and ways to solve them.

We represent a fact by a set of its prime implicates $\mathcal{F}=\{p_1, p_2, ..., p_n\}$. We denote by $\sum_\Delta = \{\neg\alpha_i \vee \beta_i \ / \ \alpha_i \to \beta_i \in \Delta\}$ the material counterpart of Δ obtained by replacing each default $\alpha_i \to \beta_i$ in Δ by its corresponding material implication $\neg\alpha_i \vee \beta_i$. $\mathcal{F}$ and $\sum_\Delta$ are assumed to be consistent. Only, $\mathcal{F} \cup \sum_\Delta$ can be inconsistent.

Definition 1: A subbase C of $\sum_\Delta \cup \mathcal{F}$ is called a conflict if it satisfies the two following properties: ① $C \vdash \bot$ (C is classically inconsistent) and ② $\forall \phi \in C$, $C - \{\phi\} \nvdash \bot$ (C is minimal w.r.t. the set inclusion relation).

For each conflict C, we assume the existence of a priority relation $<_C$ between rules inside this conflict. Formulas representing facts in C (i.e., $\mathcal{F} \cap C$) get always the highest rank. Intuitively, this ordering reflects the specificity relation and we will see in Section 3.3. one way to compute this priority relation. In fact this ordering is only used to determine which defaults contain a general antecedent class, in other way which defaults are susceptible to be given up. Hence a conflict C can be simply viewed as a couple ($\underline{C}$, $\bar{C}$) where:

$\underline{C}$ = {φ | φ∈C and ∄ ψ∈C, φ $<_C$ ψ}, and

$\bar{C}$ = C - $\underline{C}$.

Intuitively, $\underline{C}$ contains defaults which can be non-pertinent for the fact and $\bar{C}$ contains defaults which are surely pertinent. Of course $\underline{C}$ must be strictly included in $\sum_\Delta$. We will denote by ℂ={C_i = ($\underline{C}_i$, $\bar{C}_i$) | C_i is a conflict in Δ and $\mathcal{F}$} the set of all the conflicts of $\sum_\Delta \cup \mathcal{F}$.

When $\sum_\Delta \cup \mathcal{F}$ is consistent, we simply apply classical logic to deduce plausible conclusions from $\sum_\Delta \cup \mathcal{F}$. However, an interesting case is when ℂ is not empty. Hence we have to solve all the conflicts inside ℂ. Solving a conflict C in ℂ means removing at least one formula in C, and more precisely in $\underline{C}$. Of course, we can have several possibilities, and the following subsections present three of them.

3.2.1. A Cautious Consequence Relation

The first way to do is to remove all the defaults which are in the last layer of each conflict in ℂ. Then we get the following non-monotonic consequence relation:

Definition 2: Let Del = {$\neg\alpha_i \vee \beta_i$ / $\neg\alpha_i \vee \beta_i \in \sum_\Delta$ and ∃C∈ℂ s.t. $\neg\alpha_i \vee \beta_i \in \underline{C}$} be the set of deleted formulas; let E = $\sum_\Delta$- Del be the remaining set. A formula ψ is said to be a *cautious conclusion* of $\sum_\Delta$ and $\mathcal{F}$ if and only if E ∪ $\mathcal{F}$ ⊢ ψ.

The main drawback of this method is that it deletes too many rules. The inference is then too cautious as we can see with the following example:

Example 3: Let Δ={rr→w , rr→ll , rr→¬f , w→f , ll→r , a→¬f∨¬r },

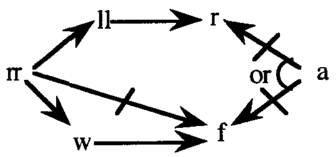

Figure 6

where rr, w, ll, f, a, r respectively mean “roadrunners”, “have wings”, “have long legs”, “fly”, “animals” and “run”.
We are interested to know the properties of Beep-Beep which is a roadrunner animal (a∧rr). Clearly, there is no problem to infer that Beep-Beep has wings (w) and long legs (ll). Concerning the properties of flying and running we have two conflicts stratified in the following way:

A={$\underline{A}$={*w→f*}, $\bar{A}$={rr→w, rr→¬f}} and

B={B̲={w→f, ll→r, a→¬f∨¬r}, B̄ ={rr→w, rr→ll}}.

The stratification is supposed to be given. A and B are the only conflicts in the knowledge base. Intuitively, the conflict A is easily solved since the A̲ contains exactly one formula and hence w→f has to be removed. Hence we prefer inferring that Beep-Beep does not fly. The conflict B concerns the property of running where we have a reason to believe in r composed of {a, rr, rr→ll, ll→r } and a reason to believe in ¬r composed of B={a∧rr, rr→w, w→f, a→¬f∨¬r}}. However, the last reason contains a formula w→f which is defeated, namely has to be removed using the conflict A. Hence the last reason is contestable and therefore we prefer to infer that Beep-Beep runs.
Unfortunately, the cautious inference relation does not allow us to infer that Beep-Beep runs. Indeed, using the previous definition, the sets A̲ and B̲ are deleted from $\sum_\Delta$ then we get the remaining base E = {rr→w, rr→ll, rr→¬f}. Hence, we cannot infer r from E and a∧rr.

3.2.2. Determining the First Solvable Conflicts

In the previous example, we have seen that solving a conflict (here the conflict A) can lead to solve other conflicts (the conflict B). Hence, it is very important to decide which conflicts must be first solved. The aim of this sub-section is to define a ranking between conflicts which indicates their solving influence:

Definition 3: A conflict A has a *positive influence* on a conflict B (or solving A solves B), denoted by A $\leq_I$ B, iff ① A̲ = B̲ or, ② A̲ ∩ B ≠ ∅ but B̲ ⊄ A̲.

Note that, when A̲⊂B̲, removing any rule from A̲ necessarily leads to solve B. However, if A̲ ∩ B ≠ ∅ then removing a rule from A̲ can possibly solve the conflict B (if the removed rule belongs both to A and to B). Note that A̲⊂B̲ implies A̲∩B≠∅. To define the first conflicts to solve, we represent the relation $\leq_I$ by a graph where the nodes are the set of conflicts and an edge is drawn from A to B iff A has a positive influence on B.

Definition 4: A conflict A is said to be in the *set of first solvable conflicts*, denoted by min(ℂ), if and only if there is no conflict B such that ① there is a path from B to A, but ② there is no path from A to B.

The two following sub-sections present two approaches which use min(ℂ) to define less cautious non-monotonic consequence relations.

3.2.3. A Less Cautious Consequence Relation

The idea in this approach is the following: for each conflict C in min(ℂ) we remove from $\sum_\Delta$ all the formulas which are in C̲, and from ℂ all the solved conflicts (conflicts containing at least one formula of C̲). We repeat again the previous step until solving all the conflicts. This approach is described by the following algorithm:

a. Let Del $= \emptyset$, $\mathbb{C}$ be the set of conflicts in $\mathcal{F} \cup \Sigma_\Delta$.

b. Repeat until $\mathbb{C} = \emptyset$

For each C in min($\mathbb{C}$),

b.1. Del $=$ Del $\cup\ \underline{C}$,

b.2. $\mathbb{C} = \mathbb{C} - \{C' / C' \in \mathbb{C}$ and $C' \cap \underline{C} \neq \emptyset\}$

c. Return E $= \Sigma_\Delta$ - Del.

Let E be the sub-set of Σ_Δ obtained in step c. Then a formula ψ is said to be a *RFS-consequence* (RFS for Removing the First Solvable conflicts) of $\mathcal{F} \cup \Sigma_\Delta$ iff ψ is a classical consequence of $\mathcal{F} \cup$E.

Clearly, any cautious conclusion of $\mathcal{F} \cup \Sigma_\Delta$ is also a RFS-consequence of $\mathcal{F} \cup \Sigma_\Delta$. The converse is false. Unfortunately the RFS-consequence can lead to adventurous conclusion as it is illustrated by the following example:

Example 4:

Let $\Delta = \{p \rightarrow b, p \rightarrow \neg f, b \rightarrow f, mw \rightarrow f\}$

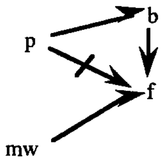

Figure 7

where p, b, f, mw respectively means “penguins”, “birds”, “fly”, “have metal wings”. Given a fact $mw \wedge p$, we are in the situation of ambiguity and we have no reason to believe in f or in $\neg$f. We have two stratified conflicts:

$A = \{\underline{A} = \{p \rightarrow \neg f, mw \rightarrow f\}, \overline{A} = \emptyset\}$ and

$B = \{\underline{B} = \{b \rightarrow f\}, \overline{B} = \{p \rightarrow b, p \rightarrow \neg f\}\}$.

Clearly, solving A can lead to solve B (if the rule $p \rightarrow \neg f$ is removed from A). Then it is interesting to try to solve A then B. Unfortunately, the RFS-consequence removes all the formulas of $\underline{A} = \{p \rightarrow \neg f, mw \rightarrow f\}$ from Σ_Δ, this leads to infer "f" which is of course non-intuitive.

Note that this approach of solving conflicts is inspired from the ideas of Williams (1996) where her proposal is done in the context of belief revision with a partial epistemic entrenchment.

3.2.4. A Universal Consequence Relation

In the two previous ways of restoring the consistency of $\mathcal{F} \cup \Sigma_\Delta$ only one sub-set, denoted by E, of Σ_Δ is generated. This sub-set in general is not a maximal consistent

sub-set of $\sum_\Delta$ (namely, there exists a formula ϕ in $\sum_\Delta$ and ϕ is not in E while $E \cup \{\phi\} \cup \mathcal{F}$ is classically consistent).
Approaches based on generating one extension are either cautious or adventurous. The cautious entailment and the RFS-entailment described in the previous sub-sections are examples of such approaches. The cautious or the adventurous aspects are due to the fact that when we solve a conflict C, all the formulas of $\underline{C}$ are removed.
In this last approach, we only remove one formula to solve a given conflict. Of course, we consider all the different possibilities and this leads to compute several possible extensions rather than to generate one extension.
The following short algorithm gives one way to construct one extension:

a. Let $Del = \emptyset$, $\mathbb{C}$ be the set of conflicts in $\mathcal{F} \cup \sum_\Delta$.
b. Repeat until $\mathbb{C} = \emptyset$
  b.1. Let C a conflict in $\min(\mathbb{C})$,
  b.2. Let ϕ a formula in $\underline{C}$,
    b.2.1 $\mathbb{C} = \mathbb{C} - \{C' / C' \in \mathbb{C} \text{ and } \phi \in C'\}$
    b.2.2. $Del = Del \cup \{\phi\}$
c. Return $E = \sum_\Delta - Del$.

We denote by $\mathcal{E}(\mathcal{F} \cup \sum_\Delta)$ the set of all extensions E_i obtained using the previous algorithm (using all the possible cases in b.1. and all the possible cases in b.2.).

Definition 5: A formula ψ is said to be a universal consequence of $\mathcal{F} \cup \sum_\Delta$ iff for each extension E_i in $\mathcal{E}(\mathcal{F} \cup \sum_\Delta)$ we have $E_i \cup \mathcal{F} \vdash \psi$.

Clearly, each cautious conclusion of $\mathcal{F} \cup \sum_\Delta$ is also a universal conclusion of $\mathcal{F} \cup \sum_\Delta$ but the converse is false. Indeed, if we consider Example 3, it is enough to notice that r is a universal consequence of $a \wedge rr$ but it is not a cautious conclusion.
Moreover, RFS-consequence relation and the universal consequence relation are incomparable. Indeed, it is enough to consider Example 4 to see that the conclusion f is a RFS-consequence of $mw \wedge p$ while it is not a universal consequence. If we consider the base $\{x \rightarrow y, x \rightarrow z, \neg y \vee \neg z, y \rightarrow t, z \rightarrow t\}$, we can notice that, assuming $\neg b \vee \neg c$ is not in the general layer of a conflict, t is a universal consequence of x but not a RFS-consequence.

3.3. Specificity-Based Default Ranking

In the description of the three different ways of dealing with inconsistency caused by the addition of a fact $\mathcal{F}$ to the default base Δ, we have assumed that each conflict is stratified into two sub-sets: $\underline{C}$ which contains rules with the general antecedent classes, and $\bar{C}$ which contains rules with the specific antecedent classes.
This section gives how to compute $\underline{C}$ for a given conflict C. We will denote by $C_\mathcal{F}$ the factual part of C (namely $C_\mathcal{F} = C \cap \mathcal{F}$), and by C_Σ the knowledge part of $\mathbb{C}$ (namely $C_\Sigma = C \cap \Sigma$). We propose the following definition of specificity:

Definition 6: A formula $\neg\phi\vee\psi$ of C_Σ is said to be in $\underline{C}$ iff there is no rule $\neg\phi'\vee\psi'$ in C_Σ with $\phi\not\equiv\phi'$ such that $\{\phi, \psi\} \cup C_\Sigma \vdash \phi'$.

In the previous definition, it is easy to check that default rules considered in $\underline{C}$ are the most general ones. Indeed a rule is in $\underline{C}$ if activating this rule (by considering true its antecedent) does not activate any other rules. We can check that the stratification of the conflicts given in the Examples 2, 3 and 4 can be obtained using the previous definition.
This definition of specificity extends the one of Touretzky (1984). It is clear that, if we reduce the rules we use to the form of the ones used by Touretzky, the notion of specificity is the same since this notion uses paths of deduction. The difference is that, with our notion, we can deal with formulas. Although it is extended, our notion remains easy to compute.
Our notion of specificity is more refined than the one used by Pearl in his System Z (1990). If we refer to the end of the discussion of Section 2, we must notice that our notion of specificity allows to infer all the results which are not exceptional. Let remind that it corresponds to the base $\Delta=\{A\rightarrow B, B\rightarrow C\}$. Considering $A\wedge\neg C$, our notion allows to consider that $A\rightarrow B$ has priority on $B\rightarrow C$ (then B can be inferred).
The main drawback of our notion is that it can't really deal with rules which form a cycle. But we can notice that it is impossible to have a cycle in a conflict. Suppose it is possible to have a cycle. This means that we can delete at least one rule and obtain the same results. So, when a conflict has a cycle, it is not minimal. This contradicts the definition of a conflict. Let's take the example of Figure 5: $\Delta=\{A\rightarrow B, B\rightarrow C, C\rightarrow A\}$. Considering $A\wedge\neg C$, we do not take into account the rule $C\rightarrow A$ and no cycle is in the conflict. Our definition is then useful in this view.

To sum up, we can say that our notion of specificity is adapted to the treatment of rules having exceptions when dealing with default is made in a local way.

3.4. Comparative Study

This section compares the behaviour of the universal consequence relation (where the stratification of conflicts is based on Definition 6) with some existing default reasoning systems. We will not discuss the behaviour of the cautious entailment or of the RFS-entailment because the first one is too cautious (see Example 3) while the later can lead to non-intuitive and adventurous conclusions (see Example 4).

Let us start this comparative study with System Z. This system proceeds in the following way: first it partitions the set of default rules Δ in $(\Delta_1,\ldots,\Delta_n)$ (this partition is based on a definition of specificity called *tolerance* where Δ_1 contains the most specific rules and Δ_n the most general rules), and next for a given fact ϕ System Z selects one sub-set E of Δ such that: $\{\phi\}\cup\Delta_1\cup\ldots\cup\Delta_i$ is consistent but $\{\phi\}\cup\Delta_1\cup\ldots\cup\Delta_{i+1}$ is inconsistent. An inference in system Z is simply defined by: ψ follows from Δ and a fact ϕ if ψ is classical consequence of $E\cup\{\phi\}$ (for more details see (Pearl, 1990)). Concerning the set of inferred conclusions, the two approaches are incomparable. Indeed, for example, in Figure 4 System Z infers the non-intuitive conclusion p from $(q\wedge r)$ while in our approach neither p nor $\neg p$ is inferred (which is satisfying). Moreover, System Z can be cautious and more precisely suffers from a so-called "blocking property inheritance", illustrated by the following figure:

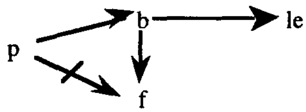

Figure 8

where b→le means that generally birds have legs. System Z does not allow to infer that "penguins have legs" while with our approach we get this intuitive conclusion. Our approach removes only concerned formula and no more.

Several solutions have been proposed to deal with the blocking property inheritance. Boutilier (1992) uses Brewka's preferred subtheories (1989) in System Z to define a new nonmonotonic inference relation. The idea is to start with the default base $(\Delta_1,\ldots,\Delta_n)$ stratified by System Z and construct a preferred subtheory E from Δ by adding to the set of observed facts $\mathcal{F}$ as many formulas of the set Δ_1 as possible (with respect to consistency criterion) then as many as possible formulas of the set Δ_2, and so on. Lastly, Boutilier defines the set of plausible conclusions as a set of assertions which hold in all preferred subtheories of Δ. This approach partially remedies to the blocking inheritance problem, but however it leads to increase the set of plausible conclusions provided by System Z. Hence, using preferred sub-theories do not block the inference of adventurous conclusions produced by System Z. There are also other solutions to the "blocking property inheritance", like the lexicographical approach, based on selecting some preferred sub-theories which contains a highest number of formulas with lower rank (see (Benferhat et al., 1993) and (Lehmann, 1993) for more details)). This approach not only generates undesirable conclusions, like in example of Figure 4, but can be very syntax dependent. For example, duplicating the same formula can change the set of plausible conclusions. For instance, if we take a variant of Nixon diamond, namely $\Delta=\{q\rightarrow p, r\rightarrow\neg p, e\rightarrow p\}$, applying the lexicographical approach to the given fact $e\wedge q\wedge r$ leads to infer p, while in our approach neither p nor $\neg p$ are inferred. Moreover, if we consider the example of Figure 1, then neither Boutilier's approach nor the lexicographical approach allow us to infer the conclusion y from the fact "$x\wedge\neg z$", while in our approach we get it.

Besides, the notion of extensions used in our approach to define the universal consequence relation is not the same as the one used in Reiter's default logic since the direction of the arrow in more restrictive for Reiter. For instance from $\Delta=\{x\rightarrow y\}$ (in Reiter's notation we write x:y/y) and $\mathcal{F}=\{\neg y\}$, using algorithm given in Section 3.2.4., we get one extension $E=\{\neg y, \neg x\vee y\}$ while in Reiter's default logic we obtain $E=\mathcal{F}$. Moreover, Default logic does not use specificity criteria to prefer one extension over the other (see example of Figure 3). Delgrande and Schaub (1994), following ideas of Reiter and Criscuolo (1981), have suggested to transform some normal defaults, in the default base, to semi-normal defaults. This will lead to eliminate unwanted extension. However, their approach suffers from some limits, due to the use of default logic, such as generating a non-maximal extension (e.g., from $\Delta=\{x\rightarrow y\}$ and $\mathcal{F}=\{\neg y\}$ we do not get $\neg x$ even if $\{\neg y, \neg x\vee y\}$ is consistent).

Geffner's conditional entailment (1992) and our system are not comparable. If we consider bases according to Figure 3, with Geffner's system it is not possible to deduce a3, ..., an-1 while our system can. On the other hand, if we take the example of the base $\Delta=\{A\wedge Y\rightarrow U, U\rightarrow A, U\rightarrow\neg W, A\rightarrow W\}$, Geffner deduce $\neg W$ from $A\wedge Y$ while it is not possible by our method (the minimality of the conflict leads to not consider the rule $U\rightarrow A$ and, then, to not between $U\rightarrow\neg W$ and $A\rightarrow W$). The wanted result depends on the meaning given to the symbols. Geffner provides an interpretation of propositional symbols where $\neg w$ follows. However, we can also give another meaning where the intuition is the opposite. This problem should be studied more precisely and discussed more deeply since the intuition is not obvious here.

Finally, the universal consequence satisfies the auto-deductivity, the consistency and the irrelevance principles discussed in Section 2. Moreover we get the desired conclusions in examples discussed in Section 2.

4. Conclusion

This paper has tried to point out that intuitive conclusions depend on general information that we have about the real world and which are not explicitly mentioned in the default base. A future work will explore this direction by providing general principles which indicate the implicit assumptions done to considering that some conclusion is plausible or not.

The proposed notion of local and coherence-based approach is appealing for several reasons. First, it extends the classical logic when the observing fact does not contradict the default. Next, it is modular in the sense that the step of computing the specificity ordering of the defaults is independent of the step of solving conflicts. Hence, if one prefers another definition of specificity, then it is not very hard to change our method. Thirdly, our approach to deal with conflicts is local and hence independent conflicts are solved separately. This is not easy to do with approaches based on a global handling of inconsistency, like System Z, where some formulas are removed while they are outside the conflicts. Lastly, the universal consequence relation appears to be particularly attractive. It is not too cautious and avoids to infer unwanted conclusions. A future work will be a deep comparison with other existing systems, especially with the one of Dung's notion of argument (1993).

Acknowledgements

We would like to thank Didier DUBOIS and Henri PRADE for their useful comments.

References

S. Benferhat, C. Cayrol, D. Dubois, J. Lang & H. Prade (1993). Inconsistency management and prioritized syntax-based entailment. Proc. of the 13th Intern. Joint Conf. on A. I. (IJCAI'93). 640-645.

S. Benferhat, D. Dubois & H. Prade (1994). Expressing Independence in a Possibilistic Framework and its Application to Default Reasoning. Proc. of the 11th European Conf. on A. I. (ECAI'94). 150-154.

S. Benferhat, A. Saffiotti & P. Smets (1995) Belief functions and default reasoning. Proc. of the 11th conf. on Uncertainty in Artificial Intelligence (UAI'95). 19-26.

C. Boutilier (1992). What is a Default priority?. Proc. of the 9th Canadian Conf. on Artificial Intelligence (AI'92). 140-147.

G. Brewka (1989). Preferred subtheories: an extended logical framework for default reasoning. Proc. of the 11th Intern. Joint Conf. on A. I. (IJCAI'89). 1043-1048.

J. P. Delgrande & T. H. Schaub (1994). A general approach to specificity in default reasoning. Proc. of the 4th Intern. Conf. on Princ. of Knowledge Representation and Reasoning (KR'94). 146-157.

D. Dubois, J. Lang & H. Prade (1994). Possibilistic logic. In: Handbook of Logic in A. I. and Logic Programming, vol. 3. Oxford University Press. 439-513.

P.M. Dung (1993). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning and logic programming. Proc. of the 13th Intern. Joint Conf. on A. I. (IJCAI'93). 852-857.

D. Gabbay (1985). Theoretical foundations for non-monotonic reasoning in expert systems. In: Logics and models of Concurrent Systems. Springer Verlag. Berlin. 439-457.

P. Gärdenfors & D. Makinson (1994). Nonmonotonic inference based on expectations. In: Artificial Intelligence, 65. 197-245.

H. Geffner (1992). Default reasoning: causal and conditional theories. MIT Press.

S. Kraus, D. Lehmann & M. Magidor (1990). Nonmonotonic reasoning, preferential models and cumulative logics. In: Artificial Intelligence, 44. 167-207.

D. Lehmann (1993). Another perspective on default reasoning. Technical report. Hebrew University, Jerusalem.

D. Lehmann & M. Magidor (1992). What does a conditional knowledge base entail? Artificial Intelligence, 55. 1-60.

Y. Moinard (1987). Donner la préférence au défaut le plus spécifique. Actes du 6ème congrès AFCET-RFIA. 1123-1132.

J. Pearl (1990). System Z: A natural ordering of defaults with tractable applications to default reasoning. Proc. of the 3rd Conf. on Theoretical Aspects of Reasoning about Knowledge (TARK'90). 121-135.

R. Reiter (1980). A logic for default reasoning. In: Artificial Intelligence, 13. 81-132.

R. Reiter & Criscuolo (1981). On interacting defaults. Proc. of the 7th Intern. Joint Conf. on A. I. (IJCAI'81). 270-276.

G. R. Simari & R. P. Loui (1992). A mathematical treatment of defeasible reasoning and its implementation. Artificial Intelligence, 53. 125-157.

D. S. Touretzky (1984). Implicit ordering of defaults in inheritance systems. Proc. of the 1984 National Conf. on A. I. (AAAI'84). 322-325.

M. A. Williams (1996). Towards a Practical Approach to Belief Revision: Reason-Based Change. Proc. of the 5th Intern. Conf. on Princ. of Knowledge Representation and Reasoning (KR'96).