SALEM BENFERHAT ET AL.

AN OVERVIEW OF POSSIBILISTIC HANDLING OF DEFAULT REASONING, WITH EXPERIMENTAL STUDIES

ABSTRACT. This paper first provides a brief survey of a possibilistic handling of default rules. A set of default rules of the form, "generally, from α deduce β", is viewed as the family of possibility distributions satisfying constraints expressing that the situation where α and β is true has a greater plausibility than the one where α and $\neg\beta$ is true. When considering only the subset of linear possibility distributions, the well-known System **P** of postulates proposed by Kraus, Lehmann and Magidor, has been obtained. We also present two rational extensions: one based on the minimum specificity principle and the other is based on the lexicographic ordering. The second part of the paper presents an empirical study of three desirable properties for a consequence relation that capture default reasoning: Rationality, Property Inheritance and Ambiguity Preservation. An experiment is conducted to investigate 13 patterns of inference for the test of these properties. Our experimental apparatus confirms previous results on the relevance of System **P**, and enforces the psychological relevance of the studied properties.

1. INTRODUCTION

One important feature of human reasoning is the ability to draw conclusions which are only plausible, i.e., conclusions being derived from incomplete information and susceptible of revision in the light of upcoming information. As goes the canonical penguin example, Tweety is a bird, hence Tweety flies – however, we should withdraw this conclusion if we learn that Tweety is a penguin besides being a bird.

This kind of reasoning has been called "default" or "nonmonotonic" reasoning, for instance by Reiter (1980). Since Reiter's paper, and in particular during the last decade, the AI literature has featured many proposals regarding default reasoning, and it has been discussed at length which could be the desirable properties for a nonmonotonic consequence relation. Yet, some came to wonder whether there could be some way to assess the naturalness of those properties, besides their being intuitively reasonable (Pelletier and Elio 1997). One possibility is to look at actual human reasoning, that is, to experimentally investigate the issue of whether human reasoners indeed abide to the properties AI researchers judge natural.

Synthese **00:** 1–18, 2003.

In the next section, we introduce a commonly agreed-on set of desirable properties for default reasoning. Then we proceed to the two objectives of this paper: (a) We introduce possibilistic logic and the possibilistic handling of default rules, in order to show how a base of default rules Δ will induce a family of possibility distributions satisfying constraints associated with each rule in Δ – we then offer a brief survey of the different non-monotonic consequence relations (together with their properties) that can be defined from this family of possibility distributions; (b) We report an experimental study investigating the psychological plausibility of rationality postulates (Kraus et al. 1990), property inheritance, redundancy, and ambiguity preservation. The last section of the paper is devoted to a concluding discussion.

2. DESIRABLE PROPERTIES FOR DEFAULT REASONING

We are interested in representing default rules (also called conditional assertions, or simply defaults) of the form "normally, if we have α, then β is the case", where α and β are formulas of some underlying finite prepositional language L. An *interpretation* for L is an assignment of a truth value in {T, F} to each formula in L according to the classical rules of propositional calculus; we denote by Ω the set of all such interpretations. We say that an interpretation ω *satisfies* a formula α, and write $\omega \models \alpha$, iff α is true in ω.

We write a default rule "normally, if we have α, then β is the case" as $\alpha \to \beta$. Note that "$\to$" is a non-classical arrow, not to be confused with material implication. A *default base* is a set $\Delta = \{\alpha_i \to \beta_i, i = 1, \ldots, n\}$ of defaults.

We use default bases to represent background knowledge about what is normally the case. We want to define a consequence relation $\mid\!\sim$ between formulas of L, that will tell us which consequences we can "reasonably" draw from each fact, given the background knowledge in $Delta$. Desirable formal properties for $\mid\!\sim$ have been largely discussed, for instance by Gabbay (1985), Kraus et al. (1990), Lehmann and Magidor (1992), and Gärdenfors and Makinson (1994).

In particular, Kraus et al. (1990) proposed a set of postulates (known as the KLM postulates) which is commonly regarded as the minimal core of any "reasonable" non-monotonic system, and defined a non-monotonic system labelled System P (for "Preferential") based on the following six postulates: (a) *Reflexivity*: $\alpha \mid\!\sim \alpha$; (b) *Left Logical Equivalence* (LLE): from $\alpha \models \alpha'$, $\alpha' \models \alpha$ and $\alpha \mid\!\sim \beta$ deduce $\alpha' \mid\!\sim \beta$; (c) *Right Weakening* (RW): from $\beta \models \beta'$ and $\alpha \mid\!\sim \beta$ deduce $\alpha \mid\!\sim \beta'$; (d) *OR*: from $\alpha \mid\!\sim \gamma$ and $\beta \mid\!\sim \gamma$ deduce

$\alpha \vee \beta \mid\sim \gamma$; (e) *Cautious Monotony* (CM): from $\alpha \mid\sim \beta$ and $\alpha \mid\sim \gamma$ deduce $\alpha \vee \beta \mid\sim \gamma$; (f) *Cut*: from $\alpha \wedge \beta \mid\sim \gamma$ and $\alpha \mid\sim \beta$ deduce $\alpha \mid\sim \gamma$.

From these rules, a consequence relation $\mid\sim_{\mathbf{P}}$ can be defined for any given Δ by: $\phi \mid\sim_{\mathbf{P}} \psi$ iff $\phi \mid\sim \psi$ can be derived from Δ using the rules of System **P**. Preferential closure of Δ using System **P** will be denoted by Δ^P.

A remarkable consequence of System **P** is *AND*: from $\alpha \mid\sim \beta$ and $\alpha \mid\sim \gamma$ deduce $\alpha \mid\sim \beta \wedge \gamma$. Another rule which has found wide (yet not unanimous) consensus, while not being a consequence of System **P**, is *Rational Monotony* (RM): from $\alpha \mid\sim \delta$ and $\alpha \not\mid\sim \neg\beta$ deduce $\alpha \wedge \beta \mid\sim \delta$. This rule has been proposed by Lehmann and Magidor (1992) in order to minimise the amount of information lost when we add a new consistent piece of information γ to a pre-existing α. Although no conclusive reason has been given for the necessity of this rule, it is usually regarded as desirable, and is validated by many current extensions of System **P**.

Moreover, it is commonly agreed on that a nonmonotonic consequence relation should satisfy the five following properties, as summarised in Benferhat et al. (2000): (a) *Specificity*: Results obtained from more specific classes should override results obtained from more generic classes; (b) *Irrelevance*: If δ is a plausible consequence of α, and if β is unrelated, “irrelevant” to α or δ, then δ should also be a plausible consequence of $\alpha \wedge \beta$; (c) *Property inheritance*: A subclass that is exceptional with respect to some property should still inherit other properties from super-classes, unless some contradiction obtains; (d) *Ambiguity preservation*: In a situation where we have one argument in favour of a proposition and one independent argument in favour of its negation, we should not conclude anything about this proposition; and (e) *Syntax independence*: The consequences of a knowledge base should not depend on the syntactical form used to represent the available knowledge – in particular, they should not be sensitive to duplications of rules in the knowledge base: Failure to do this is referred to as the *redundancy* problem.

3. POSSIBILITY THEORY AND DEFAULT RULES

Possibility theory (see Dubois et al. 1994, for more details) is based on the notion of a *possibility distribution* π which is a mapping from the set Ω to the interval [0, 1] and thus provides a complete ordering of interpretations where the most plausible ones get the highest value 1. $\pi(\omega) > 0$ means that ω is only some what plausible, while $\pi(\omega) = 0$ means that ω is impossible. π restricts the set of interpretations according to the available knowledge about the normal course of things. $\pi(\omega) > \pi(\omega')$ means that ω is more plausible than ω'. Two set-functions are associated with π:

- the possibility degree $prod(\phi) = \sup\{\pi(\omega) \mid \omega \models \phi\}$ which evaluates to what extent ϕ is consistent with the available knowledge expressed by π. $prod$ satisfies the characteristic property:

$$\forall\phi, \forall\psi, \prod(\phi \vee \psi) = \max(\prod(\phi), \prod(\psi));$$

- the dual necessity (or certainty) degree $N(\phi) = 1 - \prod(\neg\phi)$ which evaluates to what extent ϕ is entailed by the available knowledge. We have: $\forall\phi, \forall\psi, N(\phi \wedge \psi) = \min(N(\phi), N(\psi))$.

3.1. *Comparative Possibility Distributions*

The unit interval [0, 1] can be understood as a mere ordinal scale, which means that possibility theory is a qualitative theory of uncertainty. Therefore, to each possibility distribution π, we can associate its comparative counterpart, denoted by $>_\pi$, defined by $\omega >_\pi \omega'$ if and only if $\pi(\omega) > \pi'(\omega')$. $>_\pi$ is called a *comparative possibility distribution*, which can also be viewed as a well-ordered partition $(E_1, \ldots, E_n)$ of the set of classical interpretations Ω. E_1 gathers the worlds which are the most plausible ones and E_n gathers the worlds which are the least plausible ones.

We denote by $[\phi]_\psi$ the set of π-preferred models of the formula ϕ, where ω is a π-preferred model of a ϕ iff: (i) $\omega \models \phi$ and (ii) $\exists\omega', \omega' \models \phi$ and $\omega' >_\pi \omega$.

Given $\geq_\pi$, we define $\phi \geq_\Pi \psi$ (resp. $\phi \geq_\Pi \psi$) iff there exists $\omega \in [\phi]_\pi$ such that for each $\omega' \in [\psi]_\pi$, we have $\omega \geq_\pi \omega'$ (resp. $\omega >_\pi \omega'$).

A formula ψ is a possibilistic consequence of a consistent formula ϕ w.r.t. the comparative possibility distribution $\geq_\pi$, denoted by $\phi \models_\pi \psi$, iff each π-preferred model of ϕ satisfies ψ, i.e.,

$$\phi \models_\pi \psi \quad \text{iff} \forall\omega \in [\phi]_\pi, \omega \models \psi \quad \text{iff} \phi \wedge \psi >_\Pi \phi \wedge \neg\psi.$$

3.2. *Default Rules as a Family of Comparative Possibility Distributions*

In Benferhat et al. (1997), it has been proposed to model default rules of the form "normally if α then β" by "$\alpha \wedge \beta$ is more possible than $\alpha \wedge \neg\beta$". This minimal requirement, called the *auto-deduction principle*, is very natural since it guarantees that each rule $\alpha \to \beta$ in the default base is preserved, namely if α (and only α) is observed then β should follow.

A set of default rules $\Delta = \{\alpha_i \to \beta_i, i = 1, n\}$, can thus be viewed as a family of constraints restricting a family $\Pi(\Delta)$ of comparative possibility distributions. Elements of $\Pi(\Delta)$ are said to be *compatible* with

Δ. Namely, $\geq_\pi$ is said to be compatible with Δ iff it satisfies the auto-deduction principle, namely for each default rule $\alpha_i \to \beta_i$, of Δ, we have $\alpha_i \wedge \beta_I >_{Pi} \alpha_i \wedge \neg\beta_I$.

A conditional assertion $\alpha \to \beta$ is a *universal possibilistic consequence* of Δ, denoted by $\Delta \models_{\forall\Pi} \alpha \to \beta$, if and only if, β is a possibilistic consequence of α in each $\geq_\pi$ of $\Pi(\Delta)$.

Let $\Delta^\pi = \{\alpha \to \beta\colon \alpha \models_\pi \beta$ and $\geq_\pi$ is compatible with $\Delta\}$be the set of conditional assertions which are inferred from a comparative possibility distribution $\geq_\pi$ of $\Pi(\Delta)$. It turns out that the universal possibilistic consequence relation leads to exactly the same conclusions given by System **P**, namely:

$$\Delta^P = \bigcup_{\geq\pi\in\Pi(\Delta)} \Delta^\pi.$$

3.3. *System* **P** *and Linear Comparative Possibility Distributions*

A comparative possibility distribution $\geq_\pi = (E_1, \ldots, E_n)$ of $\Pi(\Delta)$ is said to be *linear* iff each E_i is a singleton (i.e., contains exactly one interpretation). In Benferhat et al. (1999) it has been shown that each $\geq_\pi$ of $\Pi(\Delta)$ can be represented by a subset $\mathbb{A}$ of $\Pi_L(\Delta)$ such that:

$$\Delta^\pi = \bigcup_{\geq\pi'\in\mathbb{A}(\Delta)} \Delta^\pi.$$

As a corollary of (1) and (2), the entailment only based on all the linear comparative possibility distributions (namely the entailment from α to β w.r.t. Δ defined by $\forall\geq_\pi \in \Pi_L(\Delta), \alpha \models_\pi \beta$) yields the preferential closure Δ^P exactly, namely:

$$\Delta^P = \bigcup_{\geq\pi\in\Pi_L(\Delta)} \Delta^\pi.$$

This means that a subset of $\Pi(\Delta)$ (here $\Pi_L(\Delta)$) is enough to recover Δ^P.

3.4. *MSP Closure Inference*

The possibilistic universal consequence is cautious since there generally exist several comparative possibility distributions compatible with a given default base. A more adventurous entailment consists in selecting one comparative possibility distribution. Selecting one particular rational extension means to accept “rational monotony” as a natural property for default reasoning.

The problem is then to find the "best" possibility distribution compatible with Δ that defines a rational extension of Δ. One possible way is to use the minimum specificity principle (MSP),[1] since it considers each interpretation to be as normal as possible, namely it assigns to each world ω the highest possibility level without violating the constraints. It can be checked that there exists exactly one comparative possibility distribution in $\Pi(\Delta)$ which is the least specific one, denoted by $\geq_{\pi\text{spe}}$.

Moreover, this way of selecting a single possibility distribution is equivalent to the *the rational closure* of Lehmann and Magidor (1992), and to Pearl's System Z (1990). The inference relation based on $\geq_{\pi\text{spe}}$ is called *MSP closure Inference*.

We denote by Δ^{spe} the rational closure of Δ obtained using the least specific possibility distribution $\geq_{\pi\text{spe}}$. A syntactic algorithm which checks if $\phi \rightarrow \psi$ belongs to Δ^{spe} has been provided (Benferhat et al. 1997). First, a set of default rules Δ is transformed into a stratified knowledge base $\sum = \Delta_1 \cup \cdots \cup \Delta_m$ such that Δ_1 contains the more general rules in Δ while Δ_m contains the most specific ones.

Once the stratification $\sum = \Delta_1 \cup \cdots \cup \Delta_1 \cup \cdots \cup \Delta_m$ is produced, the possibilistic logic machinery is applied. Namely $\phi \rightarrow \psi \in \Delta^{\text{spe}}$ iff there exists $m \leq i \leq 0$ such that: $\{\phi\} \cup \Delta_i \cdots \cup \Delta_m$ classically entails ψ and that $\{\phi\} \cup \Delta_i \cdots \cup \Delta_m$ is consistent.

The inference based on the least specific possibility distribution suffers from the "blocking of property inheritance" problem. It corresponds to the case when a class is exceptional for a superclass with respect to some property, then the least specific possibility distribution does not allow to conclude anything about whether this class is normal with respects to other properties. Let us consider the following example, where we assume that scientists have discovered some new life forms, called **Glacyceas**, in the arctic ocean, that can develop in extreme cold. Assume that we have the following default base $\{\text{T} \rightarrow \text{LC}, \text{Cr} \rightarrow \text{T}, \text{Cr} \rightarrow \text{LC}, \text{T} \rightarrow \text{M1}\}$ where the rules respectively mean: "Translucent glacyceas generally live in large colonies", "Crusts glacyceas are generally translucent", "Crusts glacyceas generally live in small colonies", "Translucent glacyceas are generally shorter than 1 mm". From this example, we cannot deduce the expected result "Crusts are generally shorter than 1 mm".

3.5. *Lexicographical Closure*

Another way to select one particular rational closure of Δ takes its inspiration from the approaches based on a lexicographic order. The main idea is again to start from a stratification $\Delta = \Delta_1 \cup \cdots \cup \Delta_m$ of Δ using the algorithm proposed in (Benferhat et al. 1997), and regard each formula in the Δ_i layer as being equally important, and more important than any set of formulas in subsequent layers. An interpretation ω is said to be *lex-prefered* than ω' if and only if there exists $1 \leq i \leq m$ such that: $\forall m \geq j > i$, $|[\omega]_j| = |[\omega']_j|$, and $|[\omega]_i| = |[\omega']_i|$, where $|[\omega]i|$ is the number of rules in Δ_i satisfied by ω.

This approach favours interpretations satisfying a maximal number of rules in Δ. The lexicographical preference induces a total pre-order on interpretations (hence a comparative possibility distribution denoted by $\geq_{\text{lex}}$) which refines the one obtained with the minimum specificity principle, which means that $\Delta^{\text{lex}} \subseteq \Delta^{\text{spe}}$, where Δ^{lex} is obtained using the comparative possibility distribution $\geq_{\text{lex}}$.

We now present an algorithm to syntactically check whether $\phi \rightarrow \psi$ belongs to Δ^{lex} or not (see Benferhat et al. 2001, for more details). This algorithm avoids the computation of conflicts which is known to be a hard problem. The idea is to construct a classical knowledge base KB which is initialized with the formula ϕ. Then at each step i (from m to 1), we check if KB classically entails ψ. If it is the case, then $\phi \rightarrow \psi$ belongs to Δ^{lex}. Otherwise, we expand KB by adding Δ_i if KB $\cup \Delta_i$ is consistent. If KB $\cup \Delta_i$ is inconsistent, we consider a weaker form of Δ_i by adding to KB all possible pairwise disjunctions of Δ_i. If the result is still inconsistent, then we replace the formulas in Δ_i by all possible disjunctions involving 3 formulas of Δ_i and again if the result is inconsistent we consider disjunctions of size 4, 5, etc. We denote by $d_k(\Delta_i)$ the set of all possible disjunctions of size k of Δ_i.

Input: (i) a stratified knowledge base $K = \Delta 1, \ldots, \Delta m$. (ii) An observation ϕ Result: check if $\phi \rightarrow \psi$is in Δ^{lex} or not.

a. KB $= \phi$
b. For $i = m$ to 1 repeat b1–b2
  - b.1. If KB $\models \psi$ then return $\phi \rightarrow \psi$ is in Δ^{lex}
  - b.2. If (KB $\cup \Delta_i$ is consistent) then KB $=$ KB $\cup \Delta_i$
    - b.2.1. else
      - b.2.1.1. $k = 2$
      - b.2.1.2. While (KB $\cup d_k(\Delta_i)$ is inconsistent and $k \Leftarrow |\Delta_i|$) do $k = k + 1$
      - b.2.1.3. If $k \Leftarrow |\Delta_i|$ then KB $=$ KB $\cup d_k(\Delta_i)$

c. Return $\phi \rightarrow \psi$ is not in Δ^{lex}.

The main limitation of the lexicographic inference is that it does not satisfy the requirement of syntax-independence. The repetition of the same default in Δ, or the presence of different arguments in favour of a conclusion, may change the result. Indeed, consider the following example $\Delta = \{H \rightarrow So, N \rightarrow \neg So, Bu \rightarrow So\}$, where the rules respectively stand for "Hermaphrodite Glacyceas generally live in a solid environment", "Necrophagous Glacyceas do not generally live in a solid environment", and "Bulging Glacyceas generally live in a solid environment". There is only one strate in Δ which contains all rules of Δ. Let us apply the above algorithm, to check if "So" is a lexicographical consequence of "H, N, Bu". We start with KB = {H, N, Bu}.

The 3 rules cannot be added to KB since it leads to an inconsistency. $D_2(\Delta) = \{\neg H \vee So \vee \neg N \vee \neg So, \neg H \vee \neg Bu \vee So, \neg N \vee \neg So \vee neg Bu \vee So\}$ which is consistent with KB. Therefore, $KB = KB \cup d_2(\Delta)$ from which "So", which is supported by two arguments, is inferred.

4. EXPERIMENTAL STUDY

Our main objective is to evaluate the psychological plausibility of lexicographic closure (LC) and MSP closure inference, by means of determining the set of properties satisfied by human inference in default reasoning, among Rationality (System **P** + Rational Monotony + And), Property Inheritance, Ambiguity Preservation and Redundancy.

Lexicographic inference satisfies all these properties (except for one form of ambiguity preservation) whereas MSP closure inference does not satisfy Property Inheritance and is not sensitive to Redundancy. As such, finding out that human inference satisfies all these properties would argue for the psychological plausibility of LC, and against the plausibility of MSP. Conversely, finding out that Property Inheritance and Sensitivity to Redundancy are the only properties which are not satisfied by human inference would argue for the reverse conclusions. Any other pattern of results shall give plausible directions for further formal works. In the following, we first summarise previous, fragmentary experimental results related to these properties, then we present a new experimental study devoted to the test of all the properties altogether.

4.1. *Previous Results*

Rationality: In a previous experiment, making use of a possibilistic semantics for plausible rules, Da Silva Neves et al. (2002) tested whether participants' judgements of possibility on a set of concrete rules were consistent with LLE, RW, CUT, CM, OR, AND, and RM. The main results were that (a) human inference was found consistent with all patterns except

LLE (which was not tested due to some problematical material), (b) about half participants made judgements consistent with all the studied patterns, and (c) 85% of participants made judgements that did not violate more that one pattern.

Property inheritance: Elio and Pelletier (1993) and Hewson and Vogel (1994) showed that whatever the abstract or concrete nature of the material, participants generally allowed a default property to be inherited.

Redundancy: Hewson and Vogel (1994) showed that participants' responses were not affected by a redundant link.

Finally, we did not found any psychological experiment relevant to Ambiguity Preservation.

On the whole, human inference appears to be consistent with the rationality postulates we consider, not to be affected by redundant links, and generally to allow default properties to be inherited. No results were available about ambiguity preservation. Thus, previous results are not conclusive in regard to the psychological plausibility of either LC or MSP inferences, and no study has explored the psychological plausibility of all considered properties altogether.

4.2. *Experiment*

4.2.1. *Participants*

Fifty seven First-year Psychology students at the University of Toulouse-Le Mirail, all native French speakers, contributed to this study. None of them had previously received any formal logical training.

4.2.2. *Material*

Twenty concrete but non-familiar default rules and 18 questions were involved in the experimental test of 13 arguments (see Tables 1 and 2), built to study Rationality, Property Inheritance, Ambiguity preservation and Redundancy. Questions were introduced by the following scenario (appearing on a computer screen).

> Consider the following facts. Scientists have discovered some new life forms in the arctic ocean (north pole), some life forms that can develop in extreme cold. As a generic term, they are called "Glacyceas". Glacyceas come into two main varieties, "Crusts" and "Worms". While scientific knowledge about Glacyceas is still scarce and not totally reliable, scientists reckon that the following is true.

A first set of premises was then displayed. Next, participants were presented with a first set of questions, introduced one after another (see section 4.2.3.).

TABLE I

Rules, questions and expected responses for the test of the Rationality properties, plus the monotony property (MN), AND and RM. For the test of RM, we only consider participants who answered “Yes” to the first question, and did not answer “yes” for the second question

MN	Worms are generally non-translucent.	
	Do you expect this Worm, with more than one year of life expectancy, to be non-translucent?	
LLE	All the Crusts and only Crusts move fast	
	Crusts are not hermaphrodite, generally.	
	Do you expect this fast moving Glacycea to be a hermaphrodite?	No
RW	All Worms have a high pressure tolerance.	
	Translucent Glacyceas are generally Worms.	
	Do you expect this translucent Glacycea to have a high pressure tolerance?	Yes
OR	Hermaphrodite Glacyceas generally live in small colonies.	
	Crusts generally live in small colonies.	
	Do you expect this Glacycea (which is either a Crust or a hermaphrodite) to live in a large colony?	No
CM	Hermaphrodite Glacyceas generally live in a solid environment.	
	Hermaphrodite Glacyceas do not, generally, have mandibles.	
	Do you expect this hermaphrodite Glacycea, living in a solid environment, to have mandibles?	No
CUT	Crusts generally live in small colonies.	
	Varities of Crusts living in small colonies have generally been in existence for some 5 millions years.	
	Do you expect this variety of Crust to have been in existence for 5 millions	Yes
years?		
AND	Crusts are not hermaphrodite, generally.	
	Crusts are generally translucent.	
	Do you expect this Crust to be translucent and not to be a hermaphrodite?	Yes
RM	Hermaphrodite Glacyceas are not generally translucent.	
	Do you expect this hermaphrodite Glacycea with mandibles to be translucent?	Yes
	Do you expect this hermaphrodite Glacycea to have mandibles?	Not Yes

TABLE II

Rules and questions for the test of Ambiguity Preservation (AMBd and AMBi), Property Inheritance (INH and INHg) and Redundancy (RED). "dnc" stands for "does not conclude". "diff" stands for "predicts a difference in Yes responses"

AMBd	Necrophagous Glacyceas do not, generally, live in a solid environment.	LC	MSP
	Hermaphrodite Glacyceas generally live in a solid environment.		
	Do you expect this hermaphrodite, necrophagous Glacycea to live in a solid environment	dnc	dnc
AMBi	Translucent Glacyceas generally live in large colonies.		
	Hermaphrodite Glacyceas generally live in small colonies.		
	Hermaphrodite Glacyceas generally live in a solid environment.		
	Hermaphrodite Glacyceas do not, generally, have mandibles.		
	Glacyceas living in a solid environment generally have mandibles.		
	Do you expect this hermaphrodite, translucent Glacycea to live in a large colony?	No	No
INH	Translucent Glacyceas generally live in large colonies.		
	Translucent Glacyceas are generally shorter than 1 mm.		
	Crusts are generally translucent.		
	Crusts generally live in small colonies.		
	Do you expect this Crust to be longer than 1 mm?	No	dnc
INHg	Translucent Glacyceas generally live in large colonies.		
	Non-translucent Glacyceas are generally shorter than 1 mm.		
	Crusts are generally translucent.		
	Crusts generally live in small colonies.		
	Do you expect this non-translucent Crust to be shorter than 1mm?	Yes	dnc
RED	Hermaphrodite Glacyceas generally live in a solid environment.		
	Necrophagous Glacyceas do not generally live in a solid environment.		
	Bulging Glacyceas generally live in a solid environment.		
	Do you expect this necrophagous, hermaphrodite and Bulging Glacycea to live in a solid environment?	diff.	no diff.

In addition to the questions displayed in Tables 1 and 2, the following four questions were asked to participants:

> Do you expect this Glacycea, living in a large colony, to have mandibles?
> Do you expect this translucent Glacycea to live in a solid environment?
> Do you expect this Crust and this Worm to resist to a temperature below −75 °C?
> Do you expect this Worm to be necrophagous?

This was done in order to introduce questions with no objective response other than: “There is no way to tell”.

4.2.3. *Design and Procedure*
Instructions and questions were displayed on a computer screen. Instructions were broken into several pages (participants were able to go back and forth pages). On the first page, participants were informed about the aim of the experiment and about some characteristics of default rules. Next, a brief scenario (see Section 4.2.2 above) was introduced, followed by a first set of default rules. Questions were introduced one after another, in the same order for all the participants. Questions were framed as follows:

> Suppose you have to examine [e.g., a hermaphrodite Glacycea].
> Would you expect that [e.g., this Glacycea]:
>
> A. [e.g., lives in a solid environment]?
> B. [e.g., does not live in a solid environment]?
>
> Or do you think that:
>
> C. there is no way to tell?

In addition, participants were asked how sure they were that their answer was true (by choosing a single modality on the following ordinal scale: “Quite sure, Almost sure, Rather sure, A bit sure, Weakly sure, Not sure at all”). Participants could access the Glacyceas/Crust information anytime by clicking on a link – when they did so, the information appeared on the right half of the screen, the problem remaining on the left side. Participants were able to modify their answers to one question as long as they had not move to the next question. The second part of the experiment was similar to the first, except that participants had access to some additional information and were asked a different set of questions. Predictions are given in Table 3.

TABLE III

Observed frequencies and Chi^2 values by pattern of inference. The exponent values notify the predicted responses: L for LC; Π for MSP, P for System **P**. N = 49, except for RED and MRc (N = 22); df = 2, except for MN and AND (df = 1)

	Responses frequencies			Chi^2	p
	A	B	C		
MN	23	0	26	0.18	ns
LLE	4	31P	14	22.8	0.000
RW	31^P	2	16	25.8	0.000
OR	1	39^P	9	49.1	0.000
CM	21	21^P	7	8	0.02
Cut	31^P	1	17	27.6	0.000
And	44^P	0	5	31	0000
RM	18 (32)	0 (2)	4 (15)	27.8	0.000
RMc	9	9	4	2.3	ns
AMBd	0 (1)	2 (6)	20 $(15)^{L\Pi}$	29.4 (13.7)	0.000 (0.001)
RED	2 (1)	2 (16)L	18 $(32)^{\Pi}$	23.3	0.000
AMBi	2 (0)	8 $(4)^{L\Pi}$	39 (18)	48.3 (8.9)	0.000 (0.003)
INH	3 (2)	21 $(12)^L$	25 $(8)^{\Pi}$	16.8 (6.9)	0.000 (0.03)
INHg	1 (0)	31 $(18)^L$	17 $(4)^{\Pi}$	27.6 (8.9)	0.000 (0.003)

4.2.4. *Rationale*

Given the percentages of “Yes”, “No” and “there is no way to tell” answers to each question, we computed a Chi-square coefficient in order to test the differences between these percentages. We concluded that human inference was consistent with some property or pattern of inference if both the modal response was the predicted one (depending on the considered system) and the null hypothesis H0 (“there is no significant difference between the percentages of responses”) was rejected, i.e., the probability of obtaining a value as large as the observed Chi-square was not greater than 0.05, as it is usual in experimental psychology.

4.3. *Experimental Results*

Eight out of the 57 participants were excluded from the analyses, for they systematically answered “there is no way to tell” (C response) when at least one premise included the predicate “generally”.

We introduced four questions for which we expected the logical response "there is no way to tell". Such was indeed the modal answer for these four questions.

Rationality: Table 3 shows that 23 participants out of 49 reasoned monotonically about the monotonic argument (MN), and that the remaining participants (26) concluded that "there is no way to tell". There is no significant difference between these proportions. On the contrary, a significant majority of participants selected the predicted response of LLE, RW, OR and CUT patterns. However, a non negligible proportion of participants concluded that "there is no way to tell". This result can be explained by the complexity of the task, which made use of an important quantity of information. Participants also performed well on the AND property. However, a problem occurred with CM, for which the positive and negative answers were selected by the same proportions of participants, while only few participants selected the C response. In addition, only 4 participants gave responses that were perfectly in accordance with LLE, CUT, CM, RW and OR. Nineteen participants gave responses that were consistent with 4 of these 5 patterns, therefore exhibiting some satisfying degree of reationality with respect to System **P**.

Rational Monotony: Previous results neither clearly confirmed nor rejected the psychological validity of RM. In the present study, the rule "Hermaphrodite Glacyceas are not generally translucent" was asserted as plausible, and participants had to answer to the following questions:

1. "Do you expect this hermaphrodite Glacycea with mandibles to be non-translucent?"
2. "Do you expect this hermaphrodite Glacycea to have mandibles?"

Question 1 without question 2 only allows for the test of Monotony. Because of a technical problem, responses to question 2 have been recorded only for 22 participants, while responses to question 1 have been recorded for all the 49 participants (see the frequencies between parentheses in Table 3). Table 3 shows that, on the whole, participants reasoned monotonically (32 participants out of 49 choose the A response). This appears to be independent from the response to question 2. Thus, once again, results do not appear to be conclusive about RM.

Inheritance: Regarding the INH pattern, Table 3 shows that half participants' answers were consistent with LC, that is, the exceptional subclass inherited the property from the super-class despite the contradiction observed with another property (no inheritance blocking). Blocking of inheritance (consistent with MSP) is observed for the remaining participants. The values between parentheses in Table 3 (on the INH line) correspond to those 19 "rational" participants which gave answers that were "consistent" with System P. Although those proportions are slightly different from what they are when one considers all participants, there is no significant proportion reversal here. With respect to the more general case of inheritance (INHg), Table 3 shows a clear dominance of LC over MSP. Indeed, inheritance blocking is almost twice less frequent than property inheritance. With rational participants (in the sense previously defined), inheritance blocking is about four times less frequent than property inheritance.

Ambiguity: Table 3 shows that, whether ambiguity was direct (AMBd) or indirect (AMBi), and whether participants were “rational” or not, modal response was: “There is no way to tell”. This result is consistent with both LC and MSP for AMBd, but it is not for AMBi. However, it must be noticed that the set of premises used for the test of the AMBi pattern is involved in the AMBd one. So, given the fact that the AMBd pattern involves additional information, and given the overall complexity of the task, we cannot exclude that participants neglected this additional information. As such, our results would relate to AMBd only, and would be in accordance with the related prediction.

Redundancy: In order to test redundancy, after the test of direct ambiguity in the first part of the experiment, we introduced an additional premise that enforced one of the two contradictory conclusions (corresponding to the RED line in Table 3). According to the hypothesis that participants' reasoning is affected by redundancy, we should observe a decrease in the proportion of “there is no way to tell” responses, or, at least, a decrease in the mean of certainty judgements.

Results have been computed for a sample of 22 participants only. They show no significant difference between the proportions of responses to the AMBd question and to the RED question. However, a comparison between certainty judgements for the two questions shows that when participants answer “there is no way to tell”, they are significantly less confident with RED question than with AMBd question ($t(21) = 2$; $p = 0.05$), which denotes an impact of redundancy. Thus, our results are consistent with MSP if we consider the proportion of C responses, but consistent with LC if we consider the mean degree of expressed certainty.

In sum, participants' inferences appear to be more consistent with LC than with MSP. In particular, it appears that subclasses which are exceptional with respect to some property inherit other properties from super-classes, despite a contradiction, and that human default reasoning is slightly sensitive to redundancy.

5. CONCLUSION

Starting with the idea that System P and rational monotony are reasonable properties for reasoning with default rules, this paper has proposed different characterisations of possibilistic inference which abide to these properties. The different semantics were given in terms of a family of possibility distributions (i.e., in terms of a complete ranking of the interpretations).

We have presented different forms of nonmonotonic consequence relations, which are defined from families of possibility distributions. Linear possibility distributions in particular are sufficient to recover System P. Moreover, we have presented an algorithm designed to check whether a given conditional assertion belonged to the rational closure associated with some least specific possibility distribution, or with the rational closure based on lexicographic order.

Our experimental study first showed that, on the whole, participants' reasoning was consistent with LLE, RW, OR, AND, and CUT, but not as consistent with CM and RM. Except for CM, those results are in line with previous experimental studies of nonmonotonic reasoning.

Another objective of this experimental study was to investigate which of Minimum Specificity Principle inference or Lexicographic Closure inference was the closest to actual human reasoning, with respect to ambiguity, inheritance of property, and sensitivity to redundancy. Results here lead to a contrasted panorama. Regarding INH, half participants gave answers consistent with the first approach while the other half gave answers consistent with the second approach. However, with respect to the more general case of inheritance INHg, LC inference clearly provided a better fit to participants' answers.

Results related to ambiguity were consistent with both models in the case of direct ambiguity, but with none in the case of indirect ambiguity. It is probable, however, that the complexity of the task led participants to treat indirect ambiguity as direct ambiguity. Finally, results related to redundancy are consistent with MSP inference if one considers participants' choice of answer, but with LC inference if one considers the uncertainty participants expressed about their answer. This result can be taken as an indicator of the very flexible nature of human reasoning, suggesting that human reasoners can follow various models of reasoning depending for example on the pragmatic constraints they have to face.

Our experimental results confirm previous findings on the relevance of System P to human reasoning, and provide new insights on the psychological plausibility of the properties a rational nonmonotonic consequence relation should satisfy. Indeed, our results appear to be in favour of LC inference. However, (in part because of the complexity of the task that was used here) further studies are necessary in order to decide which of Minimum Specificity Principle inference or Lexicographical inference has the highest psychological plausibility.

NOTES

[1] In the qualitative case $>_\pi = \{E_1, \ldots E_n\}$ is said to be less specific than $>_{\pi'} = \{E_1, \ldots, E'_m\}$ iff: $\forall\omega$, if $\omega \in E'_i$ then $\omega \in E_j$ with $j \leq i$.

REFERENCES

Adams, E. W.: 1975, *The Logic of Conditionals*, Reidel, Dordrecht.

Benferhat, S., D. Dubois, and H. Prade: 1997, 'Nonmonotonic Reasoning, Conditional Objects and Possibility Theory', *Artificial Intelligence Journal* **92**, 259–276.

Benferhat, S., A. Saffiotti, and P. Smets: 2000, 'Belief Functions and Default Reasoning', *Artificial Intelligence Journal* **122**(1-2), 1–69.

Benferhat S., S.Kaci, D. Le Berre, and M. A. Williams: 2001b, 'Weakening Conflicting Information for Iterated Revision and Knowledge Integration', *Proc. of the 17th Intern. Joint Conf. on Artificial Intelligence IJCAI01*, pp. 109–115.

Da Silva Neves, R. M., J. F. Bonnefon, and E. Raufaste: 2002, 'An Empirical Test of Patterns for Nonmonotonic Inference', *Annals of Mathematics and Artificial Intelligence* **34**, 107–130.

Dubois, D., J. Lang, and H. Prade: 1994, 'Possibilistic Logic', in D. M. Gabbay, C. J. Hogger, J. A. Robinson, and D. Nute (eds), *Handbook of Logic in Artificial Intelligence and Logic Programming*, Vol. 3, Oxford University Press, pp. 439–513.

Elio, R. and F. J. Pelletier: 1993, 'Human Benchmarks on AI's Benchmark Problems', in *Proc. of the 15th Conference of the Cognitive Science Society*, Boulder, CO, pp. 406–411.

Gabbay, D. M.: 1985, 'Theoretical Foundations for Non-Monotonic Reasoning in Expert Systems', in K. R. Apt (ed.), *Logics and Models of Concurrent Systems*, Springer Verlag, Berlin, pp. 439–457.

Gärdenfors, P. and D. Makinson: 1994, 'Nonmonotonic Inference Based on Expectations', *Artificial Intelligence* **65**, 197–245.

Hewson, C. and C. Vogel: 1994, 'Psychological Evidence for Assumptions of Path-Based Inheritance Reasoning', in *Proc. of the 16th Conference of the Cognitive Science Society*, Atlanta, pp. 403–414.

Kraus, S., D. Lehmann, and M. Magidor: 1990, 'Non-Monotonic Reasoning, Preferential Models and Cumulative Logics', *Artificial Intelligence***44**, 167–207.

Lehmann, D. and M. Magidor 1992, 'What Does a Conditional Knowledge Base Entail? *Artificial Intelligence* **55**, 1–60.

Pearl, J.: 1990, 'System Z: A Natural Ordering of Defaults with Tractable Applications to Default Reasoning', *Procs. of Theoretical Aspects of Reasoning about Knowledge*, pp. 121–135.

Pelletier, F. J. and R. Elio: 1997, 'What Should Reasoning Be, by Default?', *Computational Intelligence* **13**(2), 165–187.
Reiter, R.: 1980, 'A Logic for Default Reasoning', *Artificial Intelligence* **13**, 81–132.

Salem Benferhat
CRIL – CNRS, Université d'Artois
Rue Jean Souvraz
SP 18 62307 Lens Cedex, France
E-mail: benferhat@cril.univ-artois.fr

Jean François Bonnefon
D.S.V.P., Université Toulouse Le Mirail
Maison de la Recherche
5, Allées A. Machado, 31058 Toulouse Cedex 5, France
E-mail: bonnefon@univ-tlse2.fr

Rui Da Silva Neves
UFR de Psychologie
Université Toulouse Le Mirail
5, Allées A. Machado, 31058 Toulouse Cedex 5, France
E-mail: neves@univ-tlse2.fr