

Une Investigation Expérimentale de l'Approche Possibiliste du Raisonnement Non Monotone

An experimental analysis of possibilistic non monotonic reasoning

Salem Benferhat¹

Jean François Bonnefon²

Rui Da Silva Neves²

¹CRIL-CNRS, Université d'Artois
Faculté Jean Perrin
Rue Jean Souvraz, SP 18, 62300 Lens
benferhat@cril.univ-artois.fr

²D.S.V.P., Université Toulouse Le Mirail
Maison de la Recherche
5, Allées A. Machado, 31058 Toulouse Cedex 9, France
{bonnefon, neves}@univ-tlse2.fr

Résumé :

Cet article introduit en premier lieu la représentation possibiliste des règles par défaut, et passe en revue trois relations de conséquence non-monotones dans le cadre de cette approche : l'inférence fondée sur le principe du minimum de spécificité (MSP), la fermeture lexicographique (LC), et les fonctions de croyance infinitésimales (LCD). Ces trois relations ont certaines propriétés en commun mais diffèrent quant à d'autres, ce qui autorise (dans la seconde partie de l'article) un test expérimental de leur plausibilité psychologique. Le raisonnement des participants manifeste des propriétés similaires à celles de LC, mais non à celles de MSP et de LCD.

Mots clefs :

Raisonnement non monotone, théorie des possibilités, étude expérimentale.

Abstract

The paper first provides a survey of the possibilistic handling of default rules, and presents three different non-monotonic consequence relations: minimum specificity inference (MSP), lexicographical closure (LC), and epsilon-belief functions (LCD). These three consequence relations share some properties but differ on others, which allows (in the second part of the paper) for an experimental test of their psychological plausibility. Actual human reasoning is shown to manifest properties which are similar to those of LC but not to those of MSP or LCD.

Keywords :

Nonmonotonic reasoning, possibility theory, experimental studies.

1. Introduction

Durant les dix dernières années, de nombreux modèles pour le maniement des règles par défaut ont été suggérés dans la littérature IA, et diverses relations de

conséquence non-monotones ont été considérées. Bien que ces relations aient en général en commun un noyau de propriétés de base (par exemple, la réflexivité), elles diffèrent souvent pour ce qui est de certaines propriétés plus subtiles (par exemple, la sensibilité à la redondance). Pourtant, il ne semble pas possible, d'un point de vue formel, de décider lesquelles de ces propriétés sont les plus désirables – et, par conséquent, quelle relation de conséquence est la plus satisfaisante. La solution retenue dans cet article a été de procéder à une investigation expérimentale, c'est-à-dire de vérifier lesquelles de ces propriétés sont manifestes dans le raisonnement humain non-monotone.

Nous commençons par énumérer les diverses propriétés qui sont habituellement considérées souhaitables pour une relation de conséquence non-monotone. Puis, après avoir introduit la logique possibiliste et le maniement possibiliste des règles par défaut, nous passons brièvement en revue trois relations de conséquence non-monotone, l'inférence fondée sur le principe du minimum de spécificité (MSP), la fermeture lexicographique (LC), et les fonctions de croyance infinitésimales (LCD). Bien que ces trois relations partagent certaines propriétés, elles diffèrent sur un certain nombre de points. Ainsi, (a) seule LC est sensible à la redondance, (b) l'héritage des propriétés n'est bloqué que par MSP, et (c) contrairement à LC et MSP, LCD ne satisfait pas la Monotonie Rationnelle. Nous rapportons ensuite une expérience mettant en lumière lesquelles de ces propriétés sont effectivement manifestes dans le raisonnement des participants. Les résultats de cette expérience nous permettent d'avancer que LC est plus proche du raisonnement humain par défaut que ne le sont MSP et LCD.

2. Propriétés Désirables pour le Raisonnement non monotone

Nous souhaitons donner une représentation de règles par défaut (aussi appelées assertions conditionnelles, règles avec exceptions, ou simplement défauts) de la forme

“normalement, si α est le cas, alors β est aussi le cas”, où α et β sont des formules d’un langage propositionnel fini L . Une *interprétation* de L est l’assignation d’une valeur de vérité parmi $\{V, F\}$ à toutes les formules de L , suivant les règles classiques du calcul propositionnel ; nous notons Ω l’ensemble de ces interprétations. Une interprétation ω satisfait une formule α (noté $\omega \models \alpha$) si et seulement si α est vraie dans ω .

Nous notons $\alpha \rightarrow \beta$ une règle par défaut “normalement, si α est le cas, alors β est aussi le cas”. Remarquons que “ $\rightarrow$ ” est une implication non-classique, qui ne doit pas être confondue avec l’implication matérielle. Une *base de défauts* est un ensemble $\Delta = \{\alpha_i \rightarrow \beta_i, i = 1, \dots, n\}$ de défauts.

Nous utilisons les bases de défauts afin de représenter les connaissances générales sur ce qui est normalement le cas. Nous souhaitons définir une relation de conséquence $\sim$ entre les formules de L , qui indiquera les conséquences que nous pouvons « raisonnablement » tirer de chaque fait, compte tenu des connaissances générales représentées dans Δ . Les propriétés souhaitables pour $\sim$ ont fait l’objet d’un large débat, ainsi dans Gabbay (1985), Kraus, Lehmann, et Magidor (1990), Lehmann et Magidor (1992), ou Gärdenfors et Makinson (1994).

En particulier, Kraus, Lehmann, et Magidor (1990) ont proposé un ensemble de postulats (les postulats KLM) communément considéré comme le noyau minimal de tout système non-monotone « raisonnable », et ont défini un système non-monotone appelé **Système P** (pour « préférentiel ») fondé sur les six postulats suivants :

- (a) *Réflexivité* : $\alpha \sim \alpha$;
- (b) *Equivalence Logique à Gauche* (LLE) :
de $\alpha \models \alpha'$ et $\alpha \sim \beta$ déduire $\alpha' \sim \beta$;
- (c) *Affaiblissement à Droite* (RW) :
de $\beta \models \beta'$ et $\alpha \sim \beta$ déduire $\alpha \sim \beta'$;
- (d) *Disjonction à Gauche* (OR) :
de $\alpha \sim \gamma$ et $\beta \sim \gamma$ déduire $\alpha \vee \beta \sim \gamma$;
- (e) *Monotonie Prudente* (CM) :
de $\alpha \sim \beta$ et $\alpha \sim \gamma$ déduire $\alpha \wedge \beta \sim \gamma$;
- (f) *Coupure* (CUT) : de $\alpha \wedge \beta \sim \gamma$ et $\alpha \sim \beta$ déduire $\alpha \sim \gamma$.

À partir de ces règles, une relation de conséquence $\sim_P$ peut être définie pour toute base Δ par: $\phi \sim_P \psi$ ssi $\phi \sim \psi$ peut être dérivée de Δ par les règles du **Système P**. La fermeture préférentielle de Δ utilisant le **Système P** sera notée Δ^P .

Une conséquence remarquable du **Système P** est la propriété de *Conjonction à Droite* (AND) :

de $\alpha \sim \beta$ et $\alpha \sim \gamma$ déduire $\alpha \sim \beta \wedge \gamma$.

Une autre règle ayant rencontré un large consensus (mais pas une approbation unanime), sans être une conséquence du **Système P**, est la *Monotonie Rationnelle* (RM) :

de $\alpha \sim \delta$ et $\text{non}(\alpha \sim \neg\beta)$ déduire $\alpha \wedge \beta \sim \delta$.

Cette règle a été proposée par Lehmann et Magidor (1992) afin de minimiser l’information perdue en ajoutant un élément β à un élément pré-existant α lorsque β est cohérent avec α . Bien que nul argument définitif ne soit venu démontrer la nécessité de cette règle, elle est habituellement considérée désirable, et est validée par de nombreuses extensions actuelles du **Système P**.

De plus, il est communément admis qu’une relation de conséquence non-monotone devrait satisfaire les cinq propriétés suivantes, telles que résumées par Benferhat, Saffiotti, et Smets (2000):

(a) *Spécificité* : les règles par défaut portant sur une sous classe doivent prendre le pas sur les règles par défaut portant sur la super classe correspondante;

(b) *Non-pertinence* : si δ est une conséquence plausible de α , et si β n’est pas pertinent vis-à-vis de α et δ , alors δ devrait également être une conséquence plausible de $\alpha \wedge \beta$;

(c) *Héritage des propriétés* : une sous classe exceptionnelle quant à une certaine propriété doit tout de même hériter des autres propriétés de sa super classe, tant qu’aucune contradiction ne s’ensuit ;

(d) *Préservation de l’Ambiguïté* : dans une situation où nous disposons d’un argument en faveur d’une proposition et d’un argument indépendant en faveur de sa négation, nous ne devrions rien conclure à propos de cette proposition ; enfin (e) *Indépendance vis-à-vis de la Syntaxe* : Les conséquences dérivées d’une base de connaissances ne devraient pas être fonction de la forme syntaxique employée pour représenter ces connaissances – en particulier, elle ne devraient pas être sensible à la duplication des règles dans la base de connaissances : la non-satisfaction de ce critère est désignée comme le *problème de la redondance*.

3. Théorie des possibilités et règles avec exceptions

La théorie des possibilités (voir Dubois, Lang, et Prade, 1994, pour plus de détails) est basée sur la notion de distribution de possibilité π qui est une fonction de l’ensemble des interprétations Ω vers l’intervalle $[0,1]$. La distribution de possibilité fournit un pré-ordre total entre les interprétations où les éléments les plus plausibles ont le degré maximal qui est 1. $\pi(\omega)$ représente le degré de compatibilité de l’interprétation ω avec les croyances disponibles si on représente des connaissances incertaines, ou un degré de satisfaction si on représente des préférences. Par convention, $\pi(\omega) > 0$ signifie que ω est quelque peu possible, alors que $\pi(\omega) = 0$ signifie que ω est complètement impossible. De façon plus générale, $\pi(\omega) > \pi(\omega')$ signifie que ω est plus plausible que ω' .

À partir d’une distribution de possibilité π , on définit deux évaluations pour une formule : le degré de cohérence (ou de possibilité), et le degré de certitude (ou

de nécessité) :

- le degré de possibilité d'une formule ϕ ,

$$\Pi(\phi) = \max\{\pi(\omega) \mid \omega \models \phi\},$$

évalue dans quelle mesure ϕ est cohérente avec les croyances disponibles exprimées par π . La mesure Π satisfait :

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

- le degré de nécessité d'une formule ϕ ,

$$N(\phi) = 1 - \Pi(\neg \phi),$$

évalue dans quelle mesure ϕ peut être déduite à partir des croyances disponibles. Nous avons:

$$\forall \phi, \forall \psi, N(\phi \wedge \psi) = \min(N(\phi), N(\psi)).$$

La dualité $N(\phi) = 1 - \Pi(\neg \phi)$ étend celle de la logique classique où une formule ϕ est inférée de la base si et seulement si son opposée $\neg \phi$ est incohérente avec la base.

3.1. Distribution de possibilités comparatives

L'intervalle numérique $[0,1]$ peut être vu comme une simple échelle ordonnée. A chaque distribution de possibilité π , on peut lui associer sa contrepartie comparative, notée $>_{\pi}$, et définie par $\omega >_{\pi} \omega'$ si et seulement si $\pi(\omega) > \pi(\omega')$. $>_{\pi}$ est appelé une *distribution de possibilité comparative*, qui peut également être vue comme une partition bien ordonnée $(E_1, \dots, E_n)$ de l'ensemble des interprétations Ω . E_1 contient les interprétations les plus plausibles alors que E_n contient les interprétations les moins plausibles.

Nous notons par $[\phi]_{\pi}$ l'ensemble des modèles π -préférés de la formule ϕ , où ω est un modèle π -préféré de ϕ ssi:

(i) $\omega \models \phi$ et ii) il n'existe pas ω' , tels que $\omega' \models \phi$ et $\omega' >_{\pi} \omega$. Etant donnée $\geq_{\pi}$, nous définissons $\phi \geq_{\Pi} \psi$ (resp. $\phi >_{\Pi} \psi$) si et seulement s'il existe une interprétation $\omega \in [\phi]_{\pi}$ tel que pour toute interprétation $\omega' \in [\psi]_{\pi}$, nous avons $\omega \geq_{\pi} \omega'$ (resp. $\omega >_{\pi} \omega'$).

Finalement, une formule ψ est une conséquence possibiliste d'une observation cohérente ϕ , par rapport à la distribution de possibilité comparative $\geq_{\pi}$, noté $\phi \models_{\pi} \psi$, si et seulement si chaque modèle π -préféré de ϕ satisfait ψ , i.e.,

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

3.2. Système P représenté par une famille de distributions de possibilités comparatives

Dans Benferhat, Dubois, Prade (1997), une règle avec exceptions de la forme "normalement si α alors β " est interprétée par une contrainte qui impose que dans le

contexte α , β est préférée à $\neg \beta$, ou encore la conjonction " $\alpha \wedge \beta$ " est plus plausible que " $\alpha \wedge \neg \beta$ ". Cette contrainte minimale est très naturelle puisqu'elle garantit que chaque règle $\alpha \rightarrow \beta$ de la base est préservée, c'est-à-dire que si α (et rien que α) est observée alors β doit être plausiblement inférée.

Un ensemble de règles avec exceptions $\Delta = \{\alpha_i \rightarrow \beta_i, i=1, n\}$, avec des antécédents α_i cohérents, peut être vu comme un ensemble de contraintes restreignant une famille $\Pi(\Delta)$ de distributions de possibilités comparatives. Les éléments de $\Pi(\Delta)$ sont appelés les distributions de possibilités comparatives compatibles avec Δ . Une distribution de possibilité $\geq_{\pi}$ est dite compatible avec Δ si et seulement si pour chaque règle avec exceptions $\alpha_i \rightarrow \beta_i$ de Δ , nous avons $\alpha_i \wedge \beta_i >_{\Pi} \alpha_i \wedge \neg \beta_i$.

Une assertion conditionnelle $\alpha \rightarrow \beta$ est dite une conséquence possibiliste universelle de Δ , notée par $\Delta \models_{\forall \Pi} \alpha \rightarrow \beta$, si et seulement si, β est une conséquence possibiliste de α dans chacune des distributions de possibilités comparative $\geq_{\pi}$ de $\Pi(\Delta)$.

La relation d'inférence possibiliste universelle donne exactement les mêmes résultats que le Système P (Benferhat et al., 1997, Dubois et Prade, 1995). C'est-à-dire, notons

$$\Delta \pi = \{\alpha \rightarrow \beta : \alpha \models_{\pi} \beta \text{ et } \geq_{\pi} \text{ compatible } \Delta\}$$

l'ensemble des assertions conditionnelles qui peuvent être déduites à partir d'une distribution de possibilités comparative $\geq_{\pi}$ de $\Pi(\Delta)$. Alors :

$$\Delta^P = \bigcap_{\geq_{\pi} \in \Pi(\Delta)} \Delta \pi.$$

3.3. Inférence basée sur le principe du minimum de spécificité (MSP)

L'inférence possibiliste universelle est très prudente puisqu'il existe généralement plusieurs distributions de possibilités compatibles avec une base de règles avec exceptions. Une relation de conséquence plus aventureuse serait de choisir une seule distribution de possibilité compatible. Sélectionner une seule distribution de possibilité compatible revient à accepter la propriété de la "monotonie rationnelle" comme une propriété naturelle pour le raisonnement avec des règles ayant des exceptions. Dans la suite nous appellerons $\Delta \pi$ une extension rationnelle de Δ .

Le problème est alors de trouver la "meilleure" distribution de possibilité compatible avec Δ . Une voie possible est d'utiliser le Principe du minimum de

spécificité (MSP)¹. Le principe du minimum de spécificité minimise le nombre de classes en plaçant les interprétations dans les premières classes possibles sans violer les contraintes induites par la base, ce qui revient à essayer d'affecter à chaque interprétation le plus grand degré possible. On peut montrer qu'il existe exactement une seule distribution de possibilité, notée $\geq_{\pi_{spe}}$, dans $\Pi(\Delta)$ qui est la moins spécifique.

La relation d'inférence possibiliste basée sur $\geq_{\pi_{spe}}$ est appelée la *MSP-Inférence* ou tout simplement MSP. Cette relation d'inférence est équivalente à la fermeture rationnelle de Lehmann et Magidor (1992), et au Système Z de Pearl (1990).

Nous noterons Δ_{spe} la fermeture rationnelle de Δ obtenue en utilisant la distribution de possibilité $\geq_{\pi_{spe}}$. Un algorithme syntaxique qui vérifie si une assertion conditionnelle $\phi \rightarrow \psi$ appartient Δ_{spe} a été proposé (Benferhat et al., 1997). D'abord la base de règles avec exceptions Δ est transformée en bases stratifiée $\Sigma = \Delta_1 \cup \dots \cup \Delta_m$ tel que Δ_1 contient les règles les plus générales ou normales de Δ et que Δ_m contient les règles les plus spécifiques. Une fois la stratification $\Sigma = \Delta_1 \cup \dots \cup \Delta_m$ est produite, la machinerie inférentielle possibiliste est appliquée. C'est-à-dire $\phi \rightarrow \psi \in \Delta_{spe}$ si et seulement si il existe $m \geq i \geq 0$ tels que: $\{\phi\} \cup \Delta_1 \dots \cup \Delta_m$ infère classiquement ψ et que $\{\phi\} \cup \Delta_1 \dots \cup \Delta_m$ est cohérent.

La relation MSP-inférence n'est pas entièrement satisfaisante et souffre du problème du « blocage d'héritage ». Ce problème correspond à la situation où si une classe est exceptionnelle par rapport à une propriété alors elle ne pourra plus hériter les autres propriétés de la classe mère.

Considérons l'exemple suivant, où nous supposons que des scientifiques ont découvert une nouvelle forme de vie, appelée **Glacyceas**, dans l'océan arctique, qui peut se développer dans des conditions de froid extrêmes, et dont les Crusts sont une sous espèce. Supposons que nous avons la base suivante : $\{T \rightarrow LC, Cr \rightarrow T, Cr \rightarrow \neg LC, T \rightarrow M1\}$ où les règles signifient que:

"généralement, les glacyceas translucides (T) vivent dans de larges colonies (LC)", "généralement les Crusts (Cr) sont translucides", "généralement, les Crusts vivent en petites colonies", "les glacyceas translucides mesurent généralement moins d'un mm (M1)". A partir de cet exemple, on ne peut pas inférer le résultat attendu que "généralement, les Crusts mesurent moins d'un mm".

3.4. La fermeture lexicographique (LC)

Une autre façon de sélectionner une fermeture rationnelle particulière de Δ est d'utiliser l'approche dite lexicographique. L'idée principale est de nouveau de partir de la stratification $\Sigma = \Delta_1 \cup \dots \cup \Delta_m$ de Δ donnée par l'algorithme proposé dans (Benferhat et al. 1997). Les formules de la strate Δ_i ont le même niveau d'importance, et sont strictement plus importantes que n'importe quelle formule des strates inférieures. Une interprétation ω est dite *lex-préférée* à ω' si et seulement si il existe $1 \leq i \leq m$ tels que: $\forall m \geq j > i, |\omega|_j| = |\omega'|_j|$, et $|\omega|_i| > |\omega'|_i|$, où $|\omega|_i|$ est le nombre de formules de Δ_i satisfaites par ω .

Cette approche favorise les interprétations qui satisfont le maximum de règles prioritaires (plus spécifiques) de Δ . La relation de préférence *lex-préférée* induit un pré-ordre entre les interprétations, c'est-à-dire peut être vue comme une distribution de possibilité comparative, notée $\geq_{lex}$.

Cette distribution de possibilité $\geq_{lex}$ raffine celle obtenue à partir du principe du minimum de spécificité. C'est-à-dire que $\Delta_{lex} \supseteq \Delta_{spe}$, où Δ_{lex} est l'ensemble des assertions conditionnelles, appelé fermeture lexicographique de Δ , qui se déduisent de la distribution de possibilité comparative $\geq_{lex}$. La relation d'inférence non monotone basée sur la distribution de possibilité comparative $\geq_{lex}$ est appelée inférence lexicographique, ou tout simplement LC-inférence.

Un algorithme qui syntaxiquement vérifie si une assertion conditionnelle $\phi \rightarrow \psi$ appartient ou non à Δ_{lex} est proposé dans (Benferhat et al., 2001).

La principale limitation de l'inférence lexicographique est qu'elle est dépendante de la syntaxe. En particulier, elle est sensible à la répétition d'une même règle de Δ , ou encore la présence dans la base de plusieurs arguments en faveur d'une conclusion peut modifier l'ensemble des conséquences plausibles. En effet, considérons l'exemple suivant $\Delta = \{H \rightarrow So, N \rightarrow \neg So, B \rightarrow So\}$, où les règles signifient respectivement que "généralement les glacyceas hermaphrodites (H) vivent dans des environnements solides (So)", "généralement, les glacyceas nécrophages (N) ne vivent pas dans des environnements solides", et que "généralement, les glacyceas bombés (B) vivent dans des environnements solides".

On peut vérifier que si on utilise la relation d'inférence lexicographique alors "les glacyceas qui sont à la fois hermaphrodites, nécrophages et Bombés vivent dans des environnements solides" puisque la conclusion « vivre dans des environnements solides » est soutenue par deux arguments, alors que sa conclusion contraire n'est soutenue que par un seul argument.

¹ $\geq_{\pi} = \{E_1, \dots, E_n\}$ est dite moins spécifique que $\geq_{\pi'} = \{E'_1, \dots, E'_m\}$ ssi: $\forall \omega$, si $\omega \in E'_i$ alors $\omega \in E_j$ avec $j \leq i$.

3.5. Le système LCD dans le cadre de la théorie des possibilités

Il existe un autre système qui va au delà du Système **P**, et qui traite le problème du blocage d'héritage. Il s'agit du système LCD (Least Committed + Dempster), développé dans (Benferhat et al., 2000) dans le cadre des fonctions de croyances. L'idée est de considérer des fonctions de croyances particulières ou les degrés de croyances sont soit proches de 0 soit proches de 1, comme dans les probabilités infinitésimales de Adams (1975).

Cette sous-section montre comment le système LCD peut être codé dans le cadre de la théorie des possibilités.

La première étape du codage consiste à voir les règles avec exceptions de la base Δ comme des informations élémentaires issues des sources *distinctes*. A chaque règle avec exceptions $\alpha \rightarrow \beta$, on lui associe une distribution de possibilité π_d de la forme:

$$\pi_d(\omega) = 1 \text{ si } \omega \models \neg\alpha \vee \beta \\ = \varepsilon_d \text{ sinon,}$$

où ε_d est un nombre infinitésimal associé à la règle avec exceptions d .

La deuxième étape du codage est de combiner ces distributions de possibilités avec l'opérateur produit, pour obtenir la distribution de possibilité associée à la base Δ . L'opérateur produit peut-être vue comme la contrepartie de la règle de combinaison de Dempster dans le cadre de la théorie des possibilités.

Etant donnée une base de règles avec exceptions $\Delta = \{d_1, \dots, d_n\}$ et un ensemble $E = \{\varepsilon_1, \dots, \varepsilon_n\}$ d'infinitésimaux ε_i associés aux règles d_1 à ε_i , nous construisons la distribution de possibilité combinée comme suit:

$$\pi_* = * \{\pi_d \mid d \in \Delta\}. \quad (1)$$

C'est-à-dire pour chaque ω dans Ω , nous avons

$$\pi_*(\omega) = * \{\varepsilon_d \mid d \in \Delta \text{ et } \omega \text{ falsifie } d\},$$

et $\pi_*(\omega) = 1$ si ω satisfait toutes les règles de Δ . Et ω falsifie $d = \alpha \rightarrow \beta$ si ω satisfait α et falsifie β .

Notons $\Pi_*(\Delta)$ la famille de toutes les distributions de possibilités construites à partir de Δ en utilisant la règle (1). Les éléments de $\Pi_*(\Delta)$ diffèrent dans le choix de l'ensemble E des infinitésimaux associés aux règles avec exceptions de Δ .

La dernière étape est de restreindre la famille $\Pi_*(\Delta)$ aux distributions de possibilités qui sont compatibles et qui sont les moins spécifiques. Notons $\Pi_{*MSP}(\Delta)$ le sous-ensemble de $\Pi_*(\Delta)$ qui contient les distributions de possibilités compatibles et qui sont les moins spécifiques.

La proposition suivante donne l'équivalence entre le système LCD et son codage en théorie des possibilités :

Proposition: Une assertion conditionnelle $\alpha \rightarrow \beta$ est vraie dans le système LCD si et seulement si $\forall \pi \in \Pi_{*MSP}(\Delta)$, $\alpha \models_{\pi} \beta$.

Contrairement à la relation MSP-inférence, le système LCD ne souffre pas du problème du blocage d'héritage. Et contrairement à la relation d'inférence lexicographique (LC) le système LCD n'est pas sensible à la répétition des mêmes règles avec exceptions dans la base.

Concernant les propriétés de rationalité, le système LCD satisfait toutes les règles du système **P**, mais viole la propriété de monotonie rationnelle.

La section suivante fournit une étude expérimentale de ces trois systèmes : La MSP-inférence, l'inférence lexicographique LC et le System LCD.

4. Etude expérimentale

Des études expérimentales antérieures ont suggéré que l'inférence humaine non monotone (i) était compatible avec le système **P** plus la monotonie rationnelle (Da Silva Neves, Bonnefon, & Raufaste, 2002 ; Benferhat, Bonnefon, and Da Silva Neves, 2003), (ii) tolérait l'héritage de propriétés (Elio & Pelletier, 1993; Hewson & Vogel, 1994 ; Benferhat et al., 2003), et (iii) était sensible à la redondance (Benferhat et al., 2003, mais voir Hewson & Vogel, 1994, pour un résultat contraire). Les résultats obtenus précédemment sont apparus plus compatibles avec LC qu'avec MSP, mais sans caractère décisif. L'expérience présentée ci après vise d'une part à raffiner ces résultats en vue d'obtenir un panorama plus clair des propriétés de l'inférence non monotone humaine, et d'autre part à étudier la compatibilité du raisonnement humain non monotone avec le système LCD. Elle se concentre tout particulièrement sur les propriétés suivantes : la monotonie (MN), la monotonie prudente (MP), la monotonie rationnelle (MR), l'ambiguïté (AMB), et la sensibilité à la redondance (RED).

4.1. Expérience

4.1.1. Sujets

Quarante étudiants inscrits en première année de psychologie à l'Université de Toulouse-II dont la langue maternelle est le Français ont participé à cette expérience. Aucun d'entre eux n'avait reçu préalablement d'enseignement en logique.

4.1.2. Matériel

Un ensemble de 7 règles (concrètes et non familières) avec des exceptions a été utilisé pour le test expérimental

de 6 arguments (cf. Tableau 1) construits pour l'étude de MN, MP, MR, AMB et RED.

4.1.3. Procédure

L'expérience était présentée sous la forme de deux livrets. Le premier présente les objectifs de l'expérience et certaines caractéristiques des règles avec exceptions. Ensuite, le scénario suivant était introduit :

Considérez les faits suivants. Des scientifiques ont découvert de nouvelles formes de vie dans l'océan arctique (le pôle nord) capables de se développer dans des froids extrêmes et regroupées sous le terme générique de Glacyceas. Les glacyceas se divisent en deux principales variétés, les « Crusts » et les « Vers », et vivent dans un environnement solide ou liquide. Par ailleurs, ils sont hermaphrodites ou sexuellement différenciés, sont plats ou bombés, sont nécrophages ou non, sont opaques ou translucides, ont des mandibules ou pas, ont une

espérance de vie supérieure à un an ou pas, et se déplacent doucement ou rapidement.

Dans la troisième page, on présentait aux sujets l'ensemble suivant de 7 règles avec exceptions (la moitié des sujets recevait ces règles avec les prédicats négatifs, e.g. translucide au lieu de opaque):

Les Vers sont généralement opaques.

Les glacyceas hermaphrodites vivent généralement dans un milieu liquide.

Les glacyceas hermaphrodites ont généralement des mandibules.

Les glacyceas nécrophages vivent généralement dans un milieu solide.

Les Crusts ont généralement une espérance de vie supérieure à un an.

Les glacyceas bombés vivent généralement dans un milieu liquide.

Les glacyceas rapides vivent généralement dans un milieu liquide.

MN	Les Vers sont généralement translucides.			
	Pensez-vous que ce Ver, avec une espérance de vie supérieure à un an, est translucide ?	oui		
MP	Les glacyceas hermaphrodites vivent généralement en milieu solide. Les glacyceas hermaphrodites n'ont pas, généralement, de mandibules			
	Pensez-vous que ce glacycea hermaphrodite vivant dans un milieu solide a des mandibules ?	non		
MR	Les Vers sont généralement translucides	LC	MSP	LCD
	Pensez-vous que ce Ver est nécrophage ? (RMc)	ncp	ncp	ncp
	Pensez-vous que ce Ver nécrophage est translucide ? (RM)	oui	oui	ncp
AMB	Les glacyceas nécrophages vivent généralement dans un milieu liquide. Les glacyceas hermaphrodites vivent généralement dans un milieu solide.			
	Pensez-vous que ce glacycea hermaphrodite et nécrophage vit dans un milieu solide ?	ncp	ncp	ncp
RED1	Les glacyceas nécrophages vivent généralement dans un milieu liquide. Les glacyceas hermaphrodites vivent généralement dans un milieu solide. Les glacyceas bombés vivent généralement dans un milieu solide.			
	Pensez-vous que ce glacycea hermaphrodite, nécrophage et bombé vit dans un milieu solide ?	oui	ncp	ncp
RED2	Les glacyceas nécrophages vivent généralement dans un environnement liquide. Les glacyceas hermaphrodites vivent généralement dans un milieu solide. Les glacyceas bombés vivent généralement dans un milieu solide. Les Glacyceas rapides vivent généralement dans un milieu solide.			
	Pensez-vous que ce glacycea hermaphrodite, nécrophage, bombé et rapide vit dans un milieu solide ?	oui	ncp	ncp

Tableau 1: Règles, questions et réponses attendues pour le test de la monotonie (MN), la monotonie prudente (MP), la monotonie rationnelle (MR), l'ambiguïté (AMB), et la redondance (RED1 et RED2). "NCP" signifie "ne conclut pas". Pour le test de RM, la réponse prédite à la seconde question (RM) est "oui" pour les seuls sujets qui ont répondu ncp à la première question (RMc).

Un deuxième livret de 7 pages avec une question par page était remis aux sujets (cf. tableau 1). Les questions étaient introduites l'une après l'autre dans le même ordre pour tous les sujets, et étaient formulées sur le modèle suivant:

Supposez que vous ayez à examiner [e.g. un glacycea hermaphrodite].

Pensez-vous que [e.g. ce glacycea]:

- ☐ A. [e.g. vit dans un milieu solide]?
- ☐ B. [e.g. vit dans un milieu liquide]?
Ou bien pensez-vous que:
- ☐ C. on ne peut pas conclure?

Il était par ailleurs demandé aux sujets le degré auquel ils étaient sûr de la réponse donnée sur une échelle ordinale en 6 points ("Tout à fait certain, presque certain, plutôt certain, peu certain, très peu certain, pas certain du tout"). Les participants avaient accès aux informations relatives aux glacyceas et aux crusts à tout moment et pouvaient modifier leurs réponses s'ils le désiraient.

4.1.4. Analyses

Etant donnés les pourcentages de réponses "oui", "non" et "on ne peut pas conclure", on calcule un χ^2 dans le but de tester les différences entre ces pourcentages. On conclut que l'inférence humaine est cohérente avec une propriété ou pattern d'inférence étudiés si la réponse modale est la réponse prédite (par le système considéré) et si l'hypothèse nulle H_0 : « Il n'y a pas de différence entre les pourcentages de réponses » est rejetée, i.e. la probabilité d'obtenir une valeur plus grande que la valeur du χ^2 observé n'est pas supérieure à .05, comme il est d'usage en psychologie expérimentale.

4.2. Résultats expérimentaux

On n'observe pas d'effet du caractère positif versus négatif de la formulation des prédicats.

Monotonie:

Le tableau 2 montre que 33 sujets sur 40 raisonnent de manière monotone en ce qui concerne l'argument étudié (MN dans tableau 1). Ce nombre est de 36 sur 40 en ce qui concerne MP. Une comparaison entre les degrés de certitude moyens pour MN et MP montre que les sujets sont significativement moins confiants ($t(38) = -3.55$; $p < .001$) en ce qui concerne leur conclusion à MN (moyenne (m) = 3.23, écart type (σ) = 1.11, $n = 40$) par rapport à leur conclusion pour MP ($m = 3.92$, $\sigma = 1.03$, $n = 40$). Par conséquent, la prémisse additionnelle impliquée dans MP conduit à conclure monotoniement avec une confiance plus grande.

Les résultats antérieurs relatifs à la monotonie rationnelle (MR) n'ont pas permis d'établir clairement sa plausibilité psychologique pas plus que sa non plausibilité. Dans cette expérience, dans le but de tester MR, la règle "les Vers sont généralement translucides" était présentée comme plausible, et les sujets devaient répondre aux deux questions suivantes:

1. "Pensez-vous que ce Ver nécrophage est translucide?"
(test de $\alpha \wedge \beta \sim \gamma$)
2. "Pensez-vous que ce Ver est nécrophage ?"
(test de $\alpha \sim \neg \beta$)

Sans la deuxième question, la première permet seulement de tester la monotonie. Vingt huit sujets sur 40 ont produit une réponse cohérente avec la conséquence de forme " $\alpha \wedge \beta \sim \gamma$ ", et 26 sujets sur ces 28 ont conclu "on ne peut pas conclure" à la question de forme " $\alpha \sim \neg \beta$ ". Ces résultats fournissent un support important à l'hypothèse que la monotonie rationnelle a une bonne plausibilité psychologique.

Il convient de noter qu'une large majorité de sujets (36 sur 40), quelle que soit leur réponse à la question de forme " $\alpha \wedge \beta \sim \gamma$ ", produisent la réponse C à la question " $\alpha \sim \neg \beta$ ". La comparaison entre les jugements de certitude moyens fournis pour MN et MR montre une différence non significative ($t(38) = -1.1$; ns) entre MN ($m = 3.20$, $\sigma = 1.04$, $n = 26$) et MR ($m = 3.44$, $\sigma = 1.16$, $n = 26$).

Finalement, une comparaison entre les jugements de certitude moyens pour MP et MR montre une différence significative ($t(39) = 3.11$; $p = .003$) et une plus grande confiance dans la conclusion à MP par rapport à MR. Globalement, ces résultats sont cohérents avec les inférences LC et MSP. En revanche LCD, ne satisfaisant pas MR, reçoit moins de support de ces résultats.

Redondance:

Le tableau 2 montre que la réponse modale à la question relative à l'ambiguïté est "on ne peut pas conclure". Ce résultat est cohérent avec LC, MSP et LCD. En accord avec l'hypothèse que les raisonnements des sujets sont sensibles à la redondance, on devrait observer une baisse dans la proportion de réponses "on ne peut pas conclure" quand le nombre de prémisses redondantes augmente. Le tableau 2 montre que quel que soit le nombre de prémisses additionnelles, la réponse modale est la réponse A. Ainsi, une seule prémisse redondante est suffisante pour lever l'ambiguïté chez une proportion non négligeable de sujets. L'augmentation du nombre de prémisses redondantes ne produit quant à elle aucun effet supplémentaire.

	Effectifs des réponses			Chi ²	p
	A	B	C		
MN .001	33	2	5	43.8	<
MP .001	1	36 ^{ELM}	3	57.9	<
MR .001	28 ^{LM}	4	8 ^E	21.8	<
MRc .001	9	9	36	57.8	<
AMB .001	1	6	33 ^{ELM}	44.4	<
RED1 .001	25 ^L	1	14 ^{EM}	21.6	<
RED2 .001	25 ^L	1	14 ^{EM}	21.6	<

Tableau 2: Effectifs des réponses des sujets et valeurs de Chi² par pattern d'inférence. Les lettres en exposant indiquent les réponses prédites: L pour LC; M pour MSP et E pour LCD. N = 40; df = 2.

Des comparaisons entre degrés de certitude moyens montre que les sujets ont significativement plus confiance dans leur conclusion relative au pattern de l'ambiguïté ($m = 3.95$, $\sigma = 1.2$, $n = 40$) qu'à celui de la redondance ($m_{RED1} = 3.08$, $\sigma = 1.21$, $n = 40$); $m_{RED2} = 3.05$, $\sigma = 1.24$, $n = 40$). Ces résultats sont compatibles seulement avec LC.

En conclusion, LC apparaît comme ayant une plausibilité psychologique supérieure à MSP et LCD. Cette conclusion est par ailleurs renforcé par des résultats expérimentaux antérieurs relatifs à l'héritage de propriété (e.g. Elio and Pelletier, 1993; Hewson and Vogel, 1994; Benferhat, Bonnefon and Da Silva Neves, 2003) qui ont montré qu'une classe exceptionnelle sous ordonnée hérite généralement des propriétés de la classe hyper ordonnée en dépit de contradictions observées relatives à une autre propriété.

5. Conclusion

Nous avons commencé cet article en présentant trois relations de conséquence non-monotone, définies à partir de familles de distributions possibilistes: l'inférence fondée sur le principe du minimum de spécificité (MSP), la fermeture lexicographique (LC), et les fonctions de croyance infinitésimales (LCD).

Ces trois relations de conséquences ne présentant pas les mêmes propriétés, nous avons pu conduire un test expérimental de leur plausibilité psychologique: le raisonnement sous-jacent à cette démarche est que la relation de conséquence présentant le plus de similarité (en termes de propriétés) avec le raisonnement humain non-monotone est celle dont la plausibilité psychologique est la plus élevée.

Les résultats de notre investigation expérimentale suggèrent que LC est la plus plausible des trois relations que nous avons considérées. En particulier, les participants ont clairement manifesté une tendance à la monotonie rationnelle et une sensibilité à la redondance. Cela étant, notre expérience ne saurait valider ou confirmer le système LC en tant que tel contre d'autres approches non possibilistes (voir cependant Raufaste, Da Silva Neves et Mariné, 2003, pour une étude de la pertinence psychologique de la logique possibiliste). En effet, dans l'état de nos investigations, tout système qui satisfait à la fois le système **P** et la monotonie rationnelle, qui préserve l'ambiguïté mais est sensible à la redondance, et qui ne bloque pas l'héritage de propriétés est *a priori* psychologiquement plausible.

6. Références

- Adams, E. W. (1975) *The logic of conditionals*. Reidel, Dordrecht.
- Benferhat S., Bonnefon J.F., Da Silva Neves R. (to appear). An overview of possibilistic handling of default reasoning with experimental studies. *Synthese*.
- Benferhat S., Dubois D., Prade H. (1997). Nonmonotonic reasoning, conditional objects and possibility theory. *Artificial Intelligence Journal*, 92, pp. 259-276, 1997.
- Benferhat S., Saffiotti A., Smets P.(2000) Belief functions and default reasoning. *Artificial Intelligence Journal*, vol. 122/1-2, 2000, pp. 1-69.
- Benferhat S., Kaci S., Le Berre D., Williams M. A. (2001b). Weakening conflicting information for iterated revision and knowledge integration. *Proc. of the 17th Intern. Joint Conf. on Artificial Intelligence IJCAI01*, , p. 109-115, 2001.
- Da Silva Neves R.M., Bonnefon J.F. and Raufaste E. (2002). An empirical test of patterns for nonmonotonic inference. *Annals of Mathematics and Artificial Intelligence*, 34 : 107-130.
- Dubois, D. and Prade, H. (1995) Conditional objects, possibility theory and default rules. *In: Conditionals : From Philosophy to Computer Sciences*. Réds: Crocco, G. et Fariñas del Cerro, L. et Herzog, A., Eds: Oxford University Press , 301--336.
- Dubois, D., Lang, J. and Prade, H. (1994) Possibilistic logic. *In: Handbook of Logic in Artificial Intelligence and Logic Programming*, Vol. 3 (D.M. Gabbay, C.J. Hogger, J.A. Robinson, D. Nute, eds.), Oxford University Press, 439-513.
- Elio, R., and Pelletier, F.J. (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, 439-457.
- Gärdenfors, P. and Makinson D. (1994) Nonmonotonic inference based on expectations. *Artificial Intelligence*, 65, 197-245.
- Hewson, C. and Vogel, C. (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., Lehmann, D. and Magidor, M. (1990) Non-monotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44, 167-207.
- Lehmann, D. and Magidor, M. (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*, 121-135.
- Pelletier F.J., and Elio R. (1997). What should reasoning be, by default ? *Computational Intelligence*, Volume 13, Number 2, pp. 165-187.
- Raufaste, E., Da Silva Neves, R., and Mariné, C. (in press). Possibility theory as a normative framework for human judgments of uncertainty. *Artificial Intelligence*.