

## LOGIQUES NON MONOTONES

Gérard CHAZAL

### Introduction

#### Définition de la monotonie

Dans le cadre de la logique classique, le calcul des prédicats du premier ordre par exemple, si l'on se donne un ensemble  $D$  de formules cohérentes entre elles, on peut, en appliquant les règles d'inférence habituelles, obtenir un nouvel ensemble de formules que l'on notera  $\text{Th}(\Delta)$  avec  $\Delta \subseteq \text{Th}(\Delta)$ . Si l'on ajoute une formule à  $\Delta$  cohérente avec l'ensemble des formules de  $\Delta$ , de manière à obtenir un ensemble  $\Delta'$  toujours cohérent,  $\text{Th}(\Delta')$  comportera au moins autant de formules que  $\text{Th}(\Delta)$ .

Cette propriété de la logique classique est appelée la monotonie. On pourrait la formuler plus simplement de la manière suivante: *si l'on ajoute à un ensemble de prémisses une nouvelle prémisses de telle sorte que la cohérence soit maintenue, le nombre de conclusions ne peut que croître.*

On peut donner une définition plus formelle de cette propriété:

Si  $A_1, A_2, \dots, A_n \vdash B$  alors  $A_1, A_2, \dots, A_n, A_{n+1} \vdash B$

#### Problèmes conduisant à adopter des logiques non monotones

##### *1 - raisonner en absence d'information complète*

Cette propriété de la logique classique qui, dans le cadre de la déduction, la rend très robuste, présente cependant quelques inconvénients lorsqu'il s'agit de modéliser certaines formes de raisonnement, soit que l'on veuille rendre compte de comporte-

ments humains, soit que l'on souhaite les reproduire avec une machine.

Prenons un exemple simple. Pour venir de Dijon à Neuchâtel, deux routes s'offraient à moi et évidemment je souhaitais prendre l'itinéraire le plus rapide. L'un de ces itinéraires emprunte des routes nationales et départementales, traverse beaucoup de villages, l'autre se fait par autoroute. Les connaissances que j'ai des routes, des autoroutes, des limitations de vitesse en vigueur, etc. me permettent de conclure que l'itinéraire par l'autoroute est plus rapide. Cependant, si j'apprends avant de partir qu'il y a d'importants travaux sur l'autoroute ou un grave accident provoquant d'énormes bouchons, je conclurai que l'itinéraire par les routes nationales et départementales est plus rapide. On peut considérer le problème de deux manières. Premièrement je peux, avant de partir me renseigner sur l'état des routes et autoroutes que je suis susceptible d'emprunter. Dans le cas où j'apprendrais l'existence de travaux ou d'un accident, je serai obligé de retirer ma première conclusion «le trajet autoroutier est le plus rapide» pour la remplacer par la conclusion «le trajet par la route est plus rapide». Nous sommes dans un cas typique de non-monotonie puisque l'ajout d'une prémisse (concernant l'état des routes) m'oblige à retirer une conclusion. Deuxièmement, je peux ne pas avoir les moyens de connaître l'état des routes. Dans ce cas j'en resterai à ma première conclusion et, bien que devant raisonner avec un état de connaissance incomplet, je partirai, quitte à réviser ma décision si cet état de connaissance est modifié. Le raisonnement non monotone va intervenir chaque fois que l'on devra raisonner ou faire «raisonner» une machine avec une base de connaissances incomplète.

De telles situations sont nombreuses dans la vie. Nous tenons fréquemment des raisonnements et aboutissons à des conclusions qui nous servent pour agir alors que nos connaissances sont insuffisantes pour obtenir ces conclusions de manière ferme et définitive avec les règles d'inférence habituelles de la logique classique. On rencontre aussi ce problème dans l'élaboration de la succession des tâches dans un processus industriel, ce qu'on appelle parfois les gammes de montage. La gamme la plus logi-

quement déduite peut être remise en cause par de nouvelles informations (état des stocks, événements sociaux, rupture chez un fournisseur, etc.).

## 2 - utiliser des règles avec exceptions

Les systèmes experts mis au point par les techniques d'intelligence artificielle utilisent le plus souvent le *modus ponens* avec des règles de la forme *Si A alors B, or A, donc B*. L'élaboration d'un système expert suppose donc que l'on traduise les connaissances de l'expert auquel on veut substituer un ordinateur sous forme de règles conditionnelles. Or, lorsqu'il s'agit de traduire ainsi des connaissances empiriques, on s'aperçoit que beaucoup de règles, comme dans la grammaire française par exemple, comportent des exceptions. Si les exceptions sont peu nombreuses, on peut pallier au problème de manière simple en ajoutant des conditions à la règle. Par exemple la règle *si x est un nom singulier et si x se termine par «ou» alors au pluriel on écrira x+s* deviendra *si x est un nom singulier et si x se termine par «ou» et si( x<> «caillou» et x<>«chou», et....) alors au pluriel on écrira x+s*. Cependant une telle technique devient inutilisable si le nombre d'exceptions est très grand (lourdeur et longueur des règles) ou si les exceptions ne sont pas connues à l'avance. Dans ce cas là, il sera préférable d'utiliser le formalisme d'une logique non monotone.

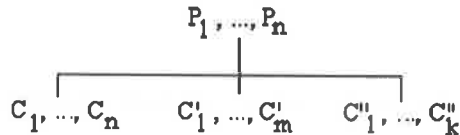
Nous serons donc amenés à utiliser une logique non monotone chaque fois que nous voudrions formaliser un raisonnement conjectural ou devant s'effectuer avec des informations incomplètes, incertaines, évolutives.

## 1. Caractères généraux des logiques non monotones

### La pluriextensionnalité

Si nous considérons un ensemble  $\Delta$  de connaissances sur les caractères des routes et des autoroutes reliant Dijon à Neuchâtel nous pourrions déduire de cet ensemble deux conclusions incompatibles: *prendre les routes nationales et départementales* et

*prendre l'autoroute.* Le choix entre l'une et l'autre se fera en fonction de l'ajout    $\Delta$  d'une information concernant l' tat des routes. D'une mani re g n rale, dans le cadre d'une logique non monotone,   partir d'un ensemble de pr misses donn es, plusieurs ensembles incompatibles de conclusions possibles peuvent  tre obtenues.



Cela entra ne un certain nombre de cons quences:

- Pour une formule  $A$ , la non-monotonie demande la d finition d'une cons quence s mantique permettant d'obtenir des conclusions qui ne sont pas v rifi es pour tous les mod les de  $A$  (en prenant «mod le» au sens habituel de la s mantique).
- Des conclusions pourront  tre consistantes avec les pr misses mais non valides au sens que la logique classique donne   ce terme. D'o  une distinction entre la validit  au sens classique du terme (une formule est valide si elle se d duit des axiomes et des pr misses) et la simple consistance.
- L'ensemble des conclusions inf r es, parmi les ensembles possibles d pendra le plus souvent de l'ordre dans lequel seront effectu es les inf rences. (Cf. L'exemple ci-dessous).
- Enfin, il faudra  videmment  viter d'inf rer conjointement des conclusions non consistantes entre elles.

**Remarque.** Pour  viter l'existence de plusieurs ensembles incompatibles de conclusions possibles, on peut utiliser la monotonie restreinte avec un syst me logique classique. Cette monotonie restreinte correspond   la r gle d'inf rence suivante:

Si  $(A_1, \dots, A_n) \vdash B$  et  
 si  $(A_1, \dots, A_n) \vdash C$ , alors  $(A_1, \dots, A_n, B) \vdash C$

Cependant cela s'avère le plus souvent insuffisant pour formaliser des formes de raisonnement du type de ceux que nous venons de voir.

### **Quelques définitions.**

Un certain nombre de notions de la logique classique devront prendre une définition un peu différente de celle que l'on a coutume d'en donner.

Tout d'abord on définira la notion de point fixe de la manière suivante:

**Un point fixe** est un ensemble stable de formules (de connaissances ou de croyances) pour lequel aucune nouvelle formule ne peut être inférée de manière consistante.

**Un théorème** sera une formule présente dans toutes les extensions pouvant être inférées.

**Une démonstration** d'une formule A visera à établir l'existence d'un point fixe pour des prémisses données, qui contienne A

### **Les règles d'inférence non monotones.**

Une logique non monotone suppose donc l'existence, sous une forme ou sous une autre, de règles d'inférence prenant en compte le caractère partiel des connaissances exprimées par les prémisses ou l'existence d'exceptions à une règle générale. Ce seront des règles révisables. Transposées dans un code informatique, un système expert par exemple, le caractère révisable de ces règles supposent qu'elles puissent être bloquées ou inhibées en cours de déduction (par apparition d'une nouvelle connaissance) ou bien qu'elles comportent des conditions d'application dont la vérification puisse évoluer.

Dans l'exemple des trajets routes ou autoroutes, la règle classique serait:

Si x est un trajet autoroutier de A à B alors x est le trajet le plus rapide de A à B.

Avec condition de consistance nous aurions:

Si  $x$  est un trajet autoroutier de  $A$  à  $B$  et s'il est consistant (avec ce qui est déjà connu) que  $x$  soit le trajet le plus rapide de  $A$  à  $B$ , alors  $x$  est le trajet le plus rapide de  $A$  à  $B$ .

Avec condition de non-inférabilité d'une assertion:

Si  $x$  est un trajet autoroutier de  $A$  à  $B$  et si l'on ne peut pas inférer qu'il existe un trajet plus rapide que  $x$  de  $A$  à  $B$ , alors  $x$  est le trajet le plus rapide de  $A$  à  $B$ .

## 2. Quelques exemples de logiques non monotones

De manière à rendre plus clair ces quelques notions, nous allons présenter rapidement deux exemples de logiques non monotones. Nous verrons après quelles conséquences épistémologiques nous pouvons en tirer.

### La logique avec défauts de Reiter.

#### *Définitions.*

Reiter (1980, 1987a et b) propose d'ajouter à la logique classique des règles d'inférence d'un type particulier appelées *défauts*. Un défaut se présente ainsi:

$$\frac{\alpha : M\beta}{\gamma}$$

Si  $\alpha$  est connu (est une prémisses ou a été déduit des prémisses) et si  $\beta$  est consistant avec tout ce qui est connu, alors on peut inférer  $\gamma$ .

On peut remplacer connu par cru suivant que l'on se place dans une logique de la connaissance ou dans une logique de la croyance.

$\alpha$ ,  $\beta$ ,  $\gamma$  sont des énoncés prédicatifs implicitement quantifiés universellement.

On aura donc des règles de défaut (ou des défauts) de la forme

$$d: \frac{\alpha(X): M\beta_1(X_1), \dots, M\beta_n(X_n)}{\gamma(Y)}$$

$\alpha(X)$ ,  $\beta_i(X_i)$ ,  $\gamma(Y)$  sont des énoncés bien formés du langage formel utilisé,  $X$ ,  $X_i$  et  $Y$  sont des ensembles de variables pris dans l'ensemble des variables du langage.

$\alpha(X)$  est le prérequis du défaut  $d$ .

$\beta_i(X_i)$  est la justification du défaut.

$\gamma(Y)$  est le conséquent du défaut.

$:M$  appartient au métalangage.

*Exemple.*

Voici un exemple de défauts emprunté à Reiter:

*Une personne habite généralement avec son conjoint*

$$d1 \frac{\text{Conjo int}(x, y) \wedge \text{Habite}(y, z) : M \text{Habite}(x, z)}{\text{Habite}(x, z)}$$

Si  $x$  est le conjoint de  $y$  et si  $y$  habite avec  $z$  et s'il est consistant avec ce qui est déjà connu de dire que  $x$  habite avec  $z$  alors  $x$  habite avec  $z$

*Une personne habite généralement dans la ville de son employeur*

$$d2 \frac{\text{Employeur}(x, y) \wedge \text{Ville}(y, z) : M \text{Habite}(x, z)}{\text{Habite}(x, z)}$$

Si  $x$  a pour employeur  $y$  et si la ville de  $y$  est  $z$  et s'il est consistant avec ce qui est déjà connu de dire que  $x$  habite  $z$  alors  $x$  habite  $z$

Si l'époux de Joséphine habite Neuchâtel et si son employeur est installé à Paris,

(1)  $\text{Habite}(\text{Joséphine}, \text{Neuchâtel})$  et

(2)  $\text{Habite}(\text{Joséphine}, \text{Paris})$

sont des formules appartenant à deux extensions incompatibles de la théorie possédant les défauts  $d_1$  et  $d_2$ . Si l'on a, par ailleurs, les prémisses suivantes:

Conjoint(Joséphine,Pierre),  
Habite(Pierre,Neuchâtel),  
Employeur(Joséphine,IBM) et  
Ville(IBM,Paris)

Si l'on applique d'abord le défaut  $d_1$ , on déduira l'énoncé (1) et le défaut  $d_2$  sera bloqué. Si l'on applique d'abord le défaut  $d_2$ , on déduira l'énoncé (2) et le défaut  $d_1$  sera bloqué. On a donc deux extensions:

Conjoint(Joséphine,Pierre), Habite(Pierre,Neuchâtel), Employeur(Joséphine,IBM)  
Ville(IBM,Paris) et Habite(Joséphine,Neuchâtel)

Conjoint(Joséphine,Pierre), Habite(Pierre,Neuchâtel), Employeur(Joséphine,IBM)  
Ville(IBM,Paris) et Habite(Joséphine,Paris)

### *Théories normales.*

On peut n'utiliser que des défauts dont le conséquent est formé du même énoncé que la justification

$$d: \frac{\alpha(X): M\beta(X_n)}{\beta(X_n)}$$

Dans ce cas on parlera de théories normales. Reiter a montré que de telles théories possédaient les propriétés suivantes:

- elles admettent toujours au moins une extension;
- elles sont semi-monotones, c'est-à-dire que si on ajoute à une théorie normale un défaut, en conservant la normalité, alors la nouvelle théorie ainsi obtenue, toujours normale, admet une extension qui inclut une extension de la théorie de départ.

### *Quelques problèmes*

On se heurte cependant à des problèmes.

1) Soit  $\Delta$  une théorie normale et  $f$  une formule du langage formel utilisé; comment savoir si  $\Delta$  possède une extension  $E$  telle que  $f \in E$ ? Autrement dit, existe-t-il une théorie de la preuve pour les théories normales?

Reiter propose une solution. Soit  $\Delta = (D, F)$ , une théorie normale, avec  $D$  ensemble des défauts et  $F$  ensemble des formules de la théorie.

Une séquence finie  $D_0, \dots, D_k$  où chaque  $D_i$  est un sous-ensemble fini de  $D$ , est une preuve de  $f$  si et seulement si:

1.  $F \cup \{CSQ(D_0)\} \vdash f$
2.  $F \cup \{CSQ(D_i)\} \vdash PRQ(D_{i-1})$  pour  $i$  de 1 à  $k$
3.  $D_k = \emptyset$
4.  $F \cup \{CSQ(D_i)\} / 0 \leq i \leq k$  est consistant.

$CSQ(D_i)$ : conjonction des conséquents des  $D_i$ .

$PRQ(D_i)$ : conjonction des prérequis des  $D_i$ .

Or pour la construction d'une preuve de  $f$ , on partira de  $D_k = \emptyset$ , puis on construira successivement  $D_{k-1}, \dots, D_0$ , mais il n'existe pas d'algorithme permettant de passer de  $D_i$  à  $D_{i-1}$ . La théorie de la preuve dans le cadre de cette logique demeure donc un peu faible.

2) Par ailleurs, on sera parfois amené à utiliser des défauts dont le conséquent n'est qu'une partie de la justification. On aura alors affaire à des théories semi-normales. Cependant les théories semi-normales perdent les propriétés des théories normales (existence d'au moins une extension et semi-monotonie) ce qui rend leur usage plus délicat.

### *Conclusion*

La logique de Reiter fournit donc un formalisme assez proche de ce qui se passe lors d'un raisonnement humain en l'absence d'une connaissance exhaustive ou avec des règles comportant des exceptions. Cependant ce formalisme présente des faiblesses particulièrement s'il doit être utilisé pour des représentations en machine. La sécurité exige des théories normales et celles-ci sont souvent insuffisantes par rapport aux besoins des logiciels envisagés.

## Les logiques modales de la connaissance et de la croyance

### *Généralités*

Un autre moyen de formaliser le raisonnement non monotone est d'utiliser le formalisme fourni par les logiques modales en réinterprétant les opérateurs modaux L et M.

LA qui signifiait «il est nécessaire que A» signifiera «A est connu (ou cru)».

MA qui signifiait «il est possible que A» signifiera «non-A n'est pas connu (ou cru)» (donc A est possible et M conserve, d'une certaine manière sa signification initiale).

Les axiomes de la logique modale trouvent aussi une nouvelle interprétation:

Le schéma d'axiome de distribution (axiome K) reste valide:

$$L(A \supset B) \supset (LA \supset LB)$$

Si  $A \supset B$  est connu alors la connaissance de A implique la connaissance de B.

L'axiome T devient le schéma d'axiome de la connaissance:

$$LA \supset A$$

Si A est connue alors A est vraie, c'est-à-dire une connaissance est une information vraie.

Le schéma d'axiome 4 devient celui de l'introspection positive:

$$LA \supset LLA$$

Si A est connu alors on sait que A est connu, autrement dit, on sait ce qu'on sait.

Le schéma d'axiome 5 devient celui de l'introspection négative

$$MA \supset LMA$$

Si je ne connais pas non-A alors je sais que je ne connais pas non-A, autrement dit je sais ce que j'ignore.

Ce dernier axiome suppose, si on utilise un système qui le retient, une parfaite connaissance des limites de la connaissance par un agent parfaitement rationnel.

On gardera la règle d'inférence modale de nécessité  $A \vdash \text{LA}$  (Si A est valide alors je sais que A).

La sémantique des mondes possibles utilisée avec les logiques modales reste valable dans le cadre de cette interprétation.

### *Logiques de McDermott et autoépistémique*

Des exemples d'utilisation d'une telle interprétation de la logique modale ont été donnés, par exemple, par McDermott et par Moore, par ce dernier sous le nom de logique autoépistémique.

McDermott recherche un système axiomatique universel indépendant des domaines d'application possibles. Le but n'est pas de déterminer telle ou telle extension d'une théorie, mais plutôt de s'intéresser à l'ensemble des formules communes à toutes les extensions possibles.

Comme dans les systèmes modaux habituels on peut bâtir des axiomatiques retenant tout ou partie seulement des axiomes modaux (K, T, 4, 5).

McDermott ajoute une règle d'inférence non monotone sous la forme:

«On ne peut inférer non-A»  $\vdash \text{MA}$ .

Cette règle permet:

- D'inférer qu'une assertion est possible, c'est-à-dire consistante d'un point de vue logique avec ce qui est déjà connu.
- D'adopter comme vrai un énoncé simplement consistant avec ce qui est connu, en ajoutant l'axiome  $\text{MA} \supset A$  et en appliquant le *modus ponens*.
- De conférer au système axiomatique un caractère non monotone puisque, si on a inféré MA, puis B à partir de MA, et si on ajoute une prémisses permettant de déduire non-A, alors on devra rétracter B.

La logique autoépistémique de Moore utilise aussi la logique modale en interprétant LA par «A est cru» et MA par «non-A n'est pas cru».

*Comparaison entre la logique de McDermott et la logique auto-épistémique de Moore.*

McDermott définit la notion de point fixe de la manière suivante. Soient un système axiomatique modale utilisant tout ou partie des axiomes modaux ( $S_{TK}$  ou  $S_{TK4}$  ou  $S_{TK5}$ ) et un ensemble de prémisses  $\Delta$ .

$\Gamma$  est un point fixe pour  $\Delta$  si et seulement si  $\Gamma$  est l'ensemble des conséquences modales par  $S_{TK}$  ou  $S_{TK4}$  ou  $S_{TK5}$  de

$$\Delta \cup \{\neg Lf \mid f \notin \Gamma\}.$$

La logique autoépistémique définit la notion d'expansion stable (comme équivalent de celle de point fixe) de la manière suivante:

$\Gamma$  est une expansion stable pour  $\Delta$  si et seulement si  $\Gamma$  est l'ensemble des conséquences modales par  $S_{TK}$  ou  $S_{TK4}$  ou  $S_{TK5}$  de  $\Delta \cup \{\neg Lf \mid f \notin \Gamma \cup Lf \mid f \in \Gamma\}$ .

*Quelques problèmes*

Premièrement les difficultés que la logique modale connaît lorsqu'il s'agit de passer du calcul des propositions à celui des prédicats se retrouvent nécessairement dans le cas d'une utilisation de ces logiques pour des logiques non monotones.

Deuxièmement, on s'aperçoit que si l'on utilise un système axiomatique possédant l'axiome 5  $MA \supset LMA$  et si  $A \vdash f$  selon la logique de McDermott alors  $A \vdash f$  selon la simple logique modale c'est-à-dire sans la règle d'inférence non monotone. Autrement dit il n'y a pas de théorème non monotone si l'on adopte l'axiome 5. McDermott choisit donc un système axiomatique ne comportant que les axiomes T, K et 4.

Troisièmement la notion de point fixe chez McDermott conduit à la remarque suivante de Moore: *Un agent de la logique non monotone de McDermott est omniscient à propos de tout ce qu'il ne croit pas, mais peut tout ignorer de ce qu'il croit.* En fait il manque à la logique de McDermott le  $\{Lf \mid f \in \Gamma\}$  qu'ajoute la notion d'expansion stable de Moore.

Quatrièmement, l'axiome T ( $LA \supset A$ ) ne pose pas de problème si  $LA$  signifie  $A$  est connu. En revanche, si  $LA$  signifie  $A$  est cru, cet axiome revient à dire que toute proposition crue par un agent est valide donc que tout ce qui est cru est vrai. Or la

non-monotonie exclut la notion de validité ou de vérité au sens de la logique classique puisqu'elle suppose la révisabilité des croyances. Or avec cet axiome rien chez McDermott ne permet de différencier le valide du simplement consistant avec des prémisses. C'est pourquoi Moore propose de rejeter l'axiome T et non l'axiome 5.

### 3. Enjeux épistémologiques

Nous venons d'avoir un aperçu rapide des formalismes utilisables pour une logique non monotone. J'ai parfaitement conscience du caractère très incomplet de cet aperçu, d'une part quant à l'exposition des logiques de Reiter, de McDermott ou de Moore et, d'autre part du fait que nous avons ignoré beaucoup d'autres travaux. Il aurait probablement fallu parler des techniques dites «de minimisation» telles que l'hypothèse du monde clos, la complétion de prédicats ou la circonscription qui, chacune à leur manière, introduisent la non-monotonie. Cependant, j'espère en avoir suffisamment dit pour pouvoir clairement aborder quelques enjeux épistémologiques de la non-monotonie.

#### Le vrai et le consistant

Il est tout d'abord intéressant de comparer la conception de la vérité qui ressort de l'usage de logiques non monotones à celle qui est habituellement associée à la sémantique de la logique classique. Nous pouvons partir de la définition que donne Tarski [TAR 44]

*X est vrai si et seulement si p*

où X est le nom de la proposition «p» dont l'exemple canonique est:

*La proposition «la neige est blanche» est vrai si et seulement si la neige est blanche.*

Ainsi, selon Tarski, reprenant explicitement une conception d'Aristote, la vérité d'une proposition consiste en son accord (ou sa correspondance) avec la réalité.

Or, toute la logique depuis Aristote, consiste à établir les règles qui permettent de faire passer le caractère vrai d'un énoncé à un autre. Les notions de déduction, d'inférence trouvent implicitement leur définition dans cette conception. En effet dire qu'un énoncé B se déduit d'un énoncé A, c'est dire que A étant vrai, B le sera aussi s'il existe une règle d'inférence reconnue permettant de passer de A à B. Dans le cadre de la logique classique, des énoncés pris comme prémisses étant vrais, soit parce qu'ils consistent en énoncés observationnels intersubjectivement reconnus, soit parce qu'ils ont été déduits selon des règles d'inférence valides d'énoncés observationnels, soit enfin parce qu'ils ont un caractère tautologique, les conclusions qui s'ensuivent selon encore des règles d'inférence valides, sont aussi vraies.

Dans la conception de Tarski, la vérité coule donc de deux sources: l'observation de ce qui est (les énoncés observationnels) et la tautologie, et son cours parcourt le discours correctement dans la mesure où il est guidé par des règles d'inférence admises et éprouvées. Nous sommes donc placés dans une perspective empiriste. Il suffirait d'ajouter que seuls les énoncés ainsi caractérisés sont scientifiques et pourvus de sens pour retrouver la position de Carnap (tout du moins dans ses premiers écrits).

Or il en va assez différemment dans le cadre des logiques non monotones. En effet, un ensemble de prémisses peut-être composé d'énoncés vrais au sens où nous venons de le définir mais le fait qu'il puisse s'en déduire plusieurs extensions mutuellement incompatibles, nous oblige à dissocier la notion de vérité de la notion de consistance ou de non contradiction. Ces deux notions vont de pair dans le cadre de la logique classique puisque les règles de déduction assurent qu'il n'y aura pas de contradiction entre les conclusions et les prémisses d'une part, et entre les conclusions entre elles d'autre part. Tout le problème de la consistance réside donc dans les prémisses ou dans le système d'axiomes adopté dans le cas d'une axiomatisation. Dans un système non monotone la consistance doit demeurer entre les prémisses et les conclusions mais non entre les conclusions.

On peut encore dire les choses d'une autre manière. En logique classique si une déduction est valide, les conclusions pourront être dites aussi valides, et les prémisses étant vraies, la validité des conclusions entraîne nécessairement leur vérité. Il n'en est plus de même en logique non monotone. Accordons encore que les prémisses soient vraies, les conclusions valides ne sont plus nécessairement vraies au sens que Tarski donne au mot.

### **Limites de l'empirisme et logique de la découverte scientifique**

La position empiriste, dans sa forme la plus radicale, celle du premier Carnap par exemple, soulève un certain nombre de problèmes bien connus. Les logiques non monotones, sans apporter une solution définitive et complète peuvent cependant ouvrir quelques voies à la réflexion.

Dans le cadre d'une recherche scientifique, le savant émet des hypothèses, ensemble d'énoncés constituant une théorie au sens logique du terme. Les énoncés de cette théorie ne sont généralement ni des énoncés observationnels, ni des tautologies. Un strict empirisme qualifierait donc de tels énoncés de métaphysiques. Ces hypothèses pourront être confrontées à des procédés expérimentaux donc à des énoncés observationnels, et si ce n'est ces hypothèses elles-mêmes ce seront des énoncés déduits de ces hypothèses selon les règles de la logique. Les énoncés déduits d'hypothèses, comme les hypothèses elles-mêmes, avant toute confrontation à des expériences validantes ou invalidantes (position vérificationniste ou falsificationniste), ne relèvent pas de la catégorie du vrai (seuls en relèvent les énoncés observationnels ou les tautologies), mais d'une simple consistance sur l'ensemble des énoncés considérés (hypothèses, énoncés déduits et éventuellement énoncés constituant le corpus de la science en question). A une hypothèse donnée peut donc correspondre plusieurs extensions au sens des logiques non monotones. Une expérience peut alors permettre d'obtenir un énoncé observationnel qui n'invalidera pas l'hypothèse mais permettra de choisir entre l'une ou l'autre des extensions possibles. Dans le cadre de la logique classique une hypothèse pou-

vait  tre, suite   une exp rience, soit corrobor e, confort e (v rifi e ou non falsifi e) soit rejet e. Dans le cadre des logiques non monotones, l'exp rimentation peut aussi viser   rejeter une hypoth se, mais tout autant   d terminer quelle extension parmi celles qui sont possibles, correspond   la r alit .

On voit donc, que les logiques non monotones peuvent offrir un cadre formel pour une logique de la d couverte scientifique. Bien s r, le formalisme propos  par la logique des d fauts de Reiter ou la logique de McDermott ou la logique auto- pist mique, ne peuvent pr tendre  puiser la question de la logique de la d couverte scientifique. Cependant, la philosophie des sciences qui s'int resse   l' laboration et   la construction des th ories scientifiques en lien avec les proc s exp rimentaux, ne peut pas ignorer de tels formalismes et les services qu'ils peuvent rendre. Nous venons de donner les principes g n raux de tels services, il resterait maintenant   les confronter   la pratique scientifique elle-m me, sur des cas concrets.

### **Le sujet de la connaissance**

Il y a une derni re cons quence des logiques non monotones qu'il faut maintenant aborder: c'est le rapport qu'entretient la connaissance avec le sujet qui la porte. Pour la logique classique, la v rit  ou la fausset  d'un  nonc  ne d pend en aucune fa on du sujet qui l' met. Peu importe qui dit que « $2+2 = 4$ » ou que «la neige est blanche». Ces  nonc s poss dent une v rit  qui peut  tre *a priori*, inn e, le fait d'une r miniscence de type platonicien ou s' tablir sur la base d'une observation ou d'une exp rience, elle ne d pend cependant absolument pas de l'esprit auquel ils sont inn s, ni de celui qui les reconna t, ni m me de l'observateur ou de l'exp rimentateur. On peut appeler cela le caract re objectif du vrai et du faux dans le cadre de la logique classique.

En revanche, les logiques non monotones ram nent brutalement le sujet de la connaissance au coeur m me de celle-ci. Dans la mesure o  elles se pr tent   la formalisation d'une connaissance incompl te ou   celle de la simple croyance, elles se r f rent n cessairement   l' tat cognitif d'un sujet   un moment donn , qu'il s'agisse d'un sujet humain ou de ce qu'on

pourrait appeler d'un «sujet artificiel» dans le cas où elles interviennent dans des systèmes informatiques de gestion de bases de connaissances. On peut montrer cet aspect particulier des logiques non monotones aussi bien avec le formalisme des défauts de Reiter qu'avec les logiques modales de McDermott ou auto-épistémiques.

Dans le cas de la logique des défauts de Reiter, le surgissement du sujet est directement lié à la pluri-extensionnalité. Les règles de déduction appliquées à une théorie comportant des défauts conduisent à l'une ou l'autre des extensions possibles en fonction de l'ordre dans lequel sont appliqués les défauts. Nous venons d'en voir un exemple. Or cet ordre n'est absolument pas déterminé par les différents éléments de la théorie (formules ou défauts). Il renvoie à un élément extérieur à la théorie, à celui qui effectue la déduction et donc en fin de compte à un sujet. Évidemment cela ne signifie pas que l'ordre d'utilisation des défauts soit complètement laissé au hasard. Le sujet déduisant peut obéir à des règles. Par exemple, nous pouvons utiliser l'une ou l'autre des règles suivantes: a) *Utiliser les défauts dans l'ordre de leur écriture*; ou b) *Utiliser prioritairement les défauts qui, statistiquement, ont donné le plus de réussite*; ou encore c) *Elaborer un classement des défauts sur une base expérimentale*; etc. Peu importe, dans tous les cas cela revient à introduire des métarègles. Cependant de telles métarègles conduisent, si elles sont correctement formulées et appliquées, à effectuer, pour une théorie donnée, un choix parmi les extensions, qui sera toujours le même, la déduction étant répétée autant de fois que l'on voudra. L'introduction de telles métarègles reviendrait donc à éliminer la non-monotonie, ou plutôt à la déplacer à une autre niveau, celui du choix de la métarègle utilisée. Il ne s'agit pas de s'engager dans une régression à l'infini en créant des métathéories non monotones qui elles-mêmes appelleraient des méta-métarègles et ainsi de suite. Simplement, on est bien obligé de remarquer que, à quelque niveau que l'on se place, il faut faire intervenir à ce niveau-là une procédure de choix (choix d'un ordre d'application des défauts, choix d'une métarègle, ou d'une méta-métarègle...) et que tout choix renvoie inévitablement à celui d'un sujet qui l'effectue.

Dans le cas des logiques non monotones utilisant le formalisme de la logique modale, la référence au sujet se fait de manière peut-être plus formelle, à la fois plus subtilement et plus explicitement. En effet tout se joue autour de trois axiomes que nous avons rapidement abordés mais qui méritent un examen plus attentif.

L'axiome T de la logique modale qui devient l'axiome de la connaissance  $LA \supset A$ , peut se lire «Si A est connu alors A est vrai». Bien sûr l'interprétation classique de l'implication donnent trois possibilités de vérité de cette formule, si l'on en fait un axiome. D'abord LA est vraie (A est connu) et A est vrai. Ensuite LA est faux (A n'est pas connu) mais A est vrai. Enfin LA est faux (A n'est pas connu) et A est faux. Le deuxième cas ménage la possibilité qu'un énoncé soit vrai et pourtant non connu comme tel, et par là il semble séparer la question de la vérité d'un énoncé de celui du sujet qui énonce cette vérité. En effet, comment peut-on dire en même temps qu'un énoncé A est vrai et qu'il n'est pas connu? Cela ne peut raisonnablement se faire que si LA signifie en fait «A est connu comme vrai». Alors effectivement il peut se faire que quelqu'un ne connaisse pas A comme vrai (c'est-à-dire le croit faux ou même l'ignore complètement) et qu'en même temps A soit vrai. Mais cette interprétation oblige à introduire un «quelqu'un», c'est-à-dire un sujet de la connaissance. Le premier cas lie explicitement la vérité d'un énoncé à une connaissance correspondante de sorte qu'une connaissance est un énoncé vrai. Enfin, dans le troisième cas, la fausseté de A est impliquée par le fait que A ne soit pas connu. On a donc trivialement un renvoi à un sujet, celui de la connaissance. En résumé, le fait, dans cet axiome d'avoir dans l'antécédent un énoncé soumis à la modalité «être connu» et dans la conclusion le même énoncé nu de toute modalité, établit un lien entre la vérité et le sujet qui a accès à cette vérité.

L'axiome 4 et l'axiome 5, respectivement d'introspection positive et d'introspection négative qui supposent que l'on sait ce que l'on sait et que l'on sait ce que l'on ne sait pas, semblent, à leur façon, dégager la connaissance de son sujet. Cependant à y regarder de plus près ce n'est pas si simple. Tout d'abord cela peut bien être admis pour l'axiome 5 puisque, comme l'a montré

McDermott, un système doté de cet axiome, malgré la règle d'inférence non monotone, ne possède pas de théorème non monotone. Autrement dit, si l'axiome 5, comme nous venons de le dire déconnecte la notion de vérité de celle du sujet d'un énoncé, il fait perdre aussi la non-monotonie. C'est la raison pour laquelle McDermott proposera d'adopter un système axiomatique ne possédant que les axiomes T, K et 4, à l'exclusion du 5. Cependant, comme nous l'avons vu, une des conséquences de la logique de McDermott est, pour reprendre les termes de la critique de Moore, qu'*un agent de la logique non monotone de McDermott est omniscient à propos de tout ce qu'il ne croit pas, mais peut tout ignorer de ce qu'il croit*. Ce qui revient à dire que l'agent de la connaissance est déterminé ou caractérisé par le système d'axiomes choisi. La solution de Moore qui consiste à conserver l'axiome 5 mais à rejeter l'axiome T maintient toutefois la référence à un sujet par le biais de celui de l'introspection positive (axiome 4).

En résumé, la logique non monotone lie la notion de vérité à un moment, un lieu puisqu'elle permet de prendre en compte un ensemble évolutif de connaissances et à un agent. Ce dernier étant tout aussi bien un sujet humain qu'un système artificiel.

## Conclusion

Ce très bref exposé est loin de faire le tour des enjeux épistémologiques des logiques non monotones. Une étude beaucoup plus détaillée des techniques de la non-monotonie permettrait d'approfondir les questions qui viennent d'être soulevées et en ferait certainement surgir bien d'autres. Cependant, j'espère avoir montré deux choses. Premièrement les logiques non monotones présentent un intérêt pratique non négligeable pour modéliser certaines formes de raisonnement qui échappent à la logique classique et s'avèrent commodes pour certaines réalisations informatiques relevant de l'intelligence artificielle. Deuxièmement, elles permettent de poser de manière renouvelée des questions épistémologiques relatives à la possibilité d'une logique de la découverte scientifique au sens où l'entendait, par

exemple, Hanson et relatives   la th orie de la connaissance. Si cet expos  est loin d' tre exhaustif, j'esp re qu'il peut montrer l'existence possible d'un d bat   la fois technique et philosophique.

*Facult  des lettres*  
*Universit  de Bourgogne*  
*Bd Gabriel*  
*F 21000 DIJON*

### Bibliographie

- FROIDEVAUX C. (1986). Axonomic default theory. *Proc. of the 7th European Conf. On A.I. (ECAI-86)*, Brighton, 123-129.
- GREGOIRE E. (1990). *Logiques non monotones et intelligence artificielle*. Paris: Herm s.
- MCDERMOTT D., DOYLE J. (1980). Non-monotonic logic I, *Artificial Intelligence* 13, 41-72.
- MCDERMOTT D. (1982). Non-monotonic logic II: non-monotonic modal theories. *J. of the ACM* 29, 34-57.
- MCDERMOTT D. (1987). A critique of pure reason. *Computational Intelligence* 3, 151-160.
- MOORE R.C. (1984). Possible-world semantics for autoepistemic logic. *Proc. of the AAAI-Workshop on Non-Monotonic Reasoning*. New York: New Paltz, 344-354.
- MOORE R.C. (1988). Autoepistemic logic. In: Smets P. & alii (eds) *Non-Standard Logics for Automated Reasoning*. London: Academic Press, 105-136.
- REITER R. (1980). A logic of default reasoning. *Artificial Intelligence* 13, 81-131.
- REITER R. (1987a). A theory of diagnosis from first principles. *Artificial Intelligence* 32, 57-95.
- REITER R. (1987b). Nonmonotonic reasoning. *Ann. Rev. Comp. Science* 2, 147-186.

TARSKI A. (1974). La conception sémantique de la vérité. In:  
*Logique, sémantique, métamathématique*. Paris: Colin,  
265 sq.