Quand l’adaptation des cas par révision des croyances et l’extrapolation analogique se rencontrent

Jean Lieber, Emmanuel Nauer, Henri Prade

Quand l'adaptation des cas par révision des croyances et l'extrapolation analogique se rencontrent *

Jean Lieber[1] **Emmanuel Nauer**[1] **Henri Prade**[2]
[1] Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
[2] IRIT, CNRS, Université de Toulouse, France
lieber@loria.fr nauer@loria.fr prade@irit.fr

Résumé

Le raisonnement à partir de cas vise à résoudre des problèmes en s'appuyant sur une base de cas, où un cas est un couple problème-solution $(\mathtt{x}, \mathtt{y})$. D'un côté, l'élaboration d'une solution au problème à résoudre (ou problème cible) peut être vue comme une forme de révision des croyances de la solution du cas remémoré (dont la partie problème est similaire au problème cible), contrainte par les connaissances du domaine. D'un autre côté, un mécanisme d'extrapolation fondé sur les proportions analogiques a été proposé récemment. Une proportion analogique est une relation entre 4 termes a, b, c, d qui se lit « a est à b ce que c est à d ». L'extrapolation analogique exploite des triplets de cas dont les parties problèmes forment une proportion analogique avec le problème cible. Puis, l'inférence consiste à résoudre une équation analogique entre les 3 solutions des cas du triplet et l'inconnue de cette équation. Cet article étudie une relation entre ces deux approches de l'adaptation. En plus de constituer une passerelle inattendue entre ces deux domaines, il propose un nouvel éclairage sur le mécanisme d'adaptation en raisonnement à partir de cas. L'article est illustré à l'aide d'un exemple suivi.

Abstract

Case-based reasoning, where cases are described in terms of problem-solution pairs $\mathtt{case} = (\mathtt{x}, \mathtt{y})$, amounts to propose a solution to a new problem on the basis of past experience made of stored cases. On the one hand, the building of the solution to a new problem may be viewed as a form of belief revision of the solution of a retrieved case (whose problem part is similar to the new problem) constrained by domain knowledge. On the other hand, an extrapolation mechanism based on analogical proportions has been proposed. It exploits triplets of cases $(\mathtt{case}^a, \mathtt{case}^b, \mathtt{case}^c)$ whose descriptions of problem parts $\mathtt{x}^a$, $\mathtt{x}^b$, $\mathtt{x}^c$ form an analogical proportion with the new problem $\mathtt{x}^{\mathtt{tgt}}$, in such a way that "$\mathtt{x}^a$ is to $\mathtt{x}^b$ as $\mathtt{x}^c$ is to $\mathtt{x}^{\mathtt{tgt}}$". Then, the analogical inference amounts to compute a solution $\mathtt{y}^{\mathtt{tgt}}$ of $\mathtt{x}^{\mathtt{tgt}}$ by solving (when possible) an equation expressing that "$\mathtt{y}^a$ is to $\mathtt{y}^b$ as $\mathtt{y}^c$ is to $\mathtt{y}^{\mathtt{tgt}}$" (where $\mathtt{y}^a$, $\mathtt{y}^b$ and $\mathtt{y}^c$ are respectively the solution parts of $\mathtt{case}^a$, $\mathtt{case}^b$ and $\mathtt{case}^c$). The paper investigates how the belief revision view and analogical extrapolation relate. Besides that it constitutes an unexpected bridge between areas which ignore each other, it casts some light on the adaptation mechanism in case-based reasoning. The paper is illustrated by a running example.

1 Introduction

La révision de croyances [2] et le raisonnement à partir de cas [1] (RàPC) sont deux domaines de l'intelligence artificielle qui sont généralement considérés comme assez éloignés l'un de l'autre et sans rapport entre eux, puisque le premier se développe principalement dans le cadre de la logique, tandis que le second traite de données et est basé sur la similarité. De plus, la révision des croyances vise à rétablir la cohérence après avoir reçu un nouvel élément d'information qui entre en conflit avec l'état courant des croyances. Le RàPC a un programme tout à fait différent puisqu'il s'agit plutôt de faire face à des informations manquantes en tirant parti de la similarité pour compléter un nouveau problème avec une solution plausible. Cependant, notez que dans les deux approches les conclusions dérivées par la révision des croyances ou par le RàPC ne sont que plausibles.

En dépit de cet état de fait apparent, il existe une vision du RàPC basée sur la révision des croyances [14, 4]. En effet, compléter le nouveau problème par une solution plausible peut être considéré comme une adaptation de la solution d'un cas récupéré afin d'être cohérent avec les spécificités du nouveau cas, ce qui est une forme de révision. Ce qui est révisé n'est clairement pas la base de cas, mais

*Les auteurs tiennent à remercier les relecteurs de cet article pour leurs remarques encourageantes et pour les suggestions qu'ils ont faites et qui seront éclairantes pour leurs travaux futurs.

c'est une copie d'un cas traitant d'un problème similaire. En d'autres termes, ce qui est réellement révisé, c'est l'allégation selon laquelle le cas récupéré peut être *appliqué (directement)* pour résoudre le problème cible.

Un mécanisme d'extrapolation basé sur des proportions analogiques (qui sont des énoncés de la forme « a est à b comme c est à d ») a été proposé pour exploiter les cas en RàPC [16]. Sur la base de 3 cas dont les parties de descriptions de problèmes sont en proportion analogique avec le nouveau problème, on induit une solution pour celui-ci à partir des solutions des 3 cas. Cela signifie qu'un processus d'adaptation a lieu à l'intérieur de l'inférence, ce qui est rendu possible par le fait que les proportions analogiques sont à la fois une question de similarité et de dissimilarité. En effet, lorsque les parties problèmes des cas a, b, c, d sont décrites par des vecteurs de traits caractéristiques, la proportion analogique est valide entre les vecteurs si a diffère de b de la même façon que c diffère de d (et vice-versa) [17, 21]. À cet égard, un couple de cas peut être considéré comme codant une sorte de règle d'adaptation au sens du RàPC [6].

Les points ci-dessus suggèrent qu'il existe une passerelle entre l'approche du RàPC basée sur la révision des croyances et l'extrapolation analogique basée sur les proportions. Cette investigation est le sujet de cet article. Il est organisé comme suit. Dans la section 2, après un bref rappel sur la logique propositionnelle et l'introduction des notations, deux ensembles de rappels sont fournis respectivement sur l'extrapolation analogique des cas et sur la vision du RàPC en termes de révision de croyances. Cette mise en place du cadre nécessaire à l'étude de la problématique traité dans cet article rend la section 2 assez longue, d'autant plus que l'exemple courant y est également introduit. Afin d'établir un pont entre l'extrapolation et l'adaptation basée sur la révision, la section 3 reformule d'abord l'adaptation par extrapolation comme une adaptation d'un seul cas. Ensuite, un opérateur de révision basé sur la compétence des couples de cas est défini, ce qui permet d'établir que l'adaptation par extrapolation peut être obtenue de manière équivalente par une révision constituée de couples de cas compétents exprimant la connaissance de l'adaptation. La section 4 présente des travaux connexes et des remarques finales, ainsi qu'une présentation de travaux futurs.

2 Présentation du problème et de l'exemple suivi

Cette section présente les notions liées au problème que cet article entend résoudre, i.e., comment l'adaptation par extrapolation analogique et l'adaptation par révision peuvent se rencontrer. Dans cette optique, les notions et notations utilisées sont présentées, et l'exemple suivi dans le domaine culinaire est présenté.

2.1 Un rapide rappel sur la logique propositionnelle

Le formalisme de représentation des cas et des connaissances du domaine est, dans cet article, la logique propositionnelle. Soit $\mathcal{V}$ un ensemble fini de symboles, appelés variables. Une formule est soit une variable soit une expression d'une des formes suivantes : $\top$, $\bot$, $\neg\varphi$, $\varphi_1 \wedge \varphi_2$, $\varphi_1 \vee \varphi_2$, $\varphi_1 \rightarrow \varphi_2$ et $\varphi_1 \leftrightarrow \varphi_2$ où φ, φ_1 et φ_2 sont des formules.

Une interprétation $\mathcal{I}$ est une fonction de $\mathcal{V}$ dans $\{0, 1\}$ où 0 et 1 dénotent les valeurs booléennes « faux » et « vrai ». L'ensemble des interprétations est dénoté par Ω. Une interprétation $\mathcal{I} \in \Omega$ est étendue sur l'ensemble de toutes les formules comme suit : $\mathcal{I}(\top) = 1$, $\mathcal{I}(\bot) = 0$, $\mathcal{I}(\neg\varphi) = \textbf{non}\, \mathcal{I}(\varphi)$, $\mathcal{I}(\varphi_1 \wedge \varphi_2) = \mathcal{I}(\varphi_1)$ **et** $\mathcal{I}(\varphi_2)$, $\mathcal{I}(\varphi_1 \vee \varphi_2) = \mathcal{I}(\varphi_1)$ **ou** $\mathcal{I}(\varphi_2)$, $\mathcal{I}(\varphi_1 \rightarrow \varphi_2) = \mathcal{I}(\neg\varphi_1 \vee \varphi_2)$ et $\mathcal{I}(\varphi_1 \leftrightarrow \varphi_2) = \mathcal{I}((\varphi_1 \rightarrow \varphi_2) \wedge (\varphi_2 \rightarrow \varphi_1))$ où **non**, **et** et **ou** sont les opérations classiques sur les booléens. Un modèle d'une formule φ est une interprétation $\mathcal{I}$ telle que $\mathcal{I}(\varphi) = 1$ et l'ensemble des modèles de φ est dénoté par $\mathcal{M}(\varphi)$. Une formule φ_1 entraîne une formule φ_2, noté $\varphi_1 \models \varphi_2$, si $\mathcal{M}(\varphi_1) \subseteq \mathcal{M}(\varphi_2)$. Les formules φ_1 et φ_2 sont équivalentes, noté $\varphi_1 \equiv \varphi_2$, si $\mathcal{M}(\varphi_1) = \mathcal{M}(\varphi_2)$. Une formule φ est *cohérente* si $\mathcal{M}(\varphi) \neq \emptyset$.

Soit $\Phi = \{\varphi_1, \varphi_2, \ldots, \varphi_p\}$ un ensemble fini de formules. $\bigvee \Phi$ dénote $\varphi_1 \vee \varphi_2 \vee \ldots \vee \varphi_p$. $\bigwedge \Phi$ dénote $\varphi_1 \wedge \varphi_2 \wedge \ldots \wedge \varphi_p$.

Dans cet article, dans un but de simplification, les notations booléennes et les notations propositionnelles sont parfois utilisées les unes pour les autres (p. ex., 0 et 1 à la place de $\bot$ et $\top$ et utilisation des connecteurs propositionnels entre valeurs booléennes).

Un« flip [1] » est une opération d'édition sur une interprétation consistant à changer exactement l'interprétation d'une variable (on peut effectuer $|\mathcal{V}|$ *flips* sur une interprétation).

2.2 Notions et notations liées au RàPC

Le RàPC a pour objectif de résoudre des problèmes grâce à une base de cas, où un cas est la représentation d'un épisode de résolution de problème. Soit $\mathcal{P}$ et $\mathcal{S}$ les espaces de problèmes et de solutions : un *problème* $\mathtt{x}$ (resp., une *solution* $\mathtt{y}$) est, par définition, un élément de $\mathcal{P}$ (resp., de $\mathcal{S}$). Une relation sur $\mathcal{P} \times \mathcal{S}$ est supposée exister qui se lit « a pour solution », mais qui n'est, pour la majorité des applications, pas complètement connue du système de RàPC. Un *cas* est un couple $\mathtt{cas} = (\mathtt{x}, \mathtt{y}) \in \mathcal{P} \times \mathcal{S}$ tel que $\mathtt{x}$ a pour solution $\mathtt{y}$. Un *cas source* $\mathtt{cas}^s = (\mathtt{x}^s, \mathtt{y}^s)$ est un élément de la base de cas, laquelle est notée $\mathtt{BC}$. Le problème à résoudre est appelé *problème cible*, dénoté par $\mathtt{x}^{\mathtt{cible}}$.

Le modèle de processus classique du RàPC s'appuie sur les étapes de remémoration et d'adaptation [24] (également appelées *retrieve* et *reuse* dans [1], i.e., *retrouver* et *réutiliser*). D'autres étapes suivent l'adaptation mais ne sont pas

1. Le terme anglais a été conservé à défaut de trouver une meilleure traduction que « chiquenaude ».

considérées dans cet article. La remémoration sélectionne un ou plusieurs cas source(s) dans l'optique de la résolution de $\mathtt{x}^{\mathtt{cible}}$. L'adaptation vise à utiliser le(s) cas remémoré(s) pour proposer une solution plausible $\mathtt{y}^{\mathtt{cible}}$ de $\mathtt{x}^{\mathtt{cible}}$. Une *adaptation simple* est une adaptation d'un seul cas remémoré, sinon, on parlera d'*adaptation multiple*.

Les *connaissances du domaine* forment une base de connaissances $\mathtt{CD}$ qui peut être comprise comme un ensemble de contraintes d'intégrité : si un problème $\mathtt{x}$ (resp., une solution $\mathtt{y}$ ou un cas $(\mathtt{x}, \mathtt{y})$) n'est pas cohérent avec $\mathtt{CD}$ alors on peut déduire qu'il n'est pas licite.

Dans cet article, il est supposé qu'une séparation claire entre problèmes et solutions est faite. Cela n'est pas toujours vrai : pour certaines applications du RàPC, chaque cas constitue un tout et le problème cible est considéré comme un cas incomplet. Dans ce cas, la séparation entre partie problème et partie solution d'un cas remémoré peut se faire au moment de l'adaptation. Étant donné que cet article concerne principalement l'adaptation, cette hypothèse de séparation entre problèmes et solutions n'est pas une grande restriction. Par conséquent, dans le cadre propositionnel, $\mathcal{V}$ est partitionné en $\{\mathcal{V}_{\mathcal{P}}, \mathcal{V}_{\mathcal{S}}\}$ et les variables apparaissant dans un problème (resp., dans une solution) sont des éléments de $\mathcal{V}_{\mathcal{P}}$ (resp., de $\mathcal{V}_{\mathcal{S}}$). Cette idée est illustrée dans la section suivante.

De plus, il est supposé que chaque cas source $(\mathtt{x}^s, \mathtt{y}^s)$ et le problème cible sont complètement décrits : étant donné les connaissances du domaine, la valeur de vérité de toute variable $a \in \mathcal{V}_{\mathcal{P}}$ est connue pour $\mathtt{x}^s$ et pour $\mathtt{x}^{\mathtt{cible}}$ (i.e., $\mathtt{CD} \wedge \mathtt{x}^s \models a$ ou $\mathtt{CD} \wedge \mathtt{x}^s \models \neg a$, et de même pour $\mathtt{x}^{\mathtt{cible}}$) et une contrainte similaire est vérifiée par $\mathtt{y}^s$. Un cas $(\mathtt{x}^s, \mathtt{y}^s)$ est représenté en logique propositionnelle par une seule formule $\mathtt{x}^s \wedge \mathtt{y}^s$.

Selon le modèle des connaissances classique du RàPC (voir, p. ex., [23]), la base de connaissances d'un système de RàPC consiste en quatre *conteneurs de connaissances*. Deux d'entre eux ont déjà été mentionnés : la base de cas $\mathtt{BC}$ et les connaissances du domaine $\mathtt{CD}$. Les deux autres sont les connaissances de remémoration et les connaissances d'adaptation, utilisées respectivement durant les étapes de remémoration et d'adaptation.

Une dernière remarque peut être faite ici à propos de la notion de similarité en RàPC : on dit en général que la remémoration vise à trouver des cas sources *similaires* au problème cible. Cette notion de similarité ne doit pas nécessairement être comprise comme une égalité approximative : cela peut aller plus loin. En fait, $(\mathtt{x}^s, \mathtt{y}^s)$ peut être considéré comme similaire à $\mathtt{x}^{\mathtt{cible}}$ même si les descriptions de $\mathtt{x}^s$ et $\mathtt{x}^{\mathtt{cible}}$ sont très différentes, pourvu que l'adaptation de $(\mathtt{x}^s, \mathtt{y}^s)$ retourne une solution plausible de $\mathtt{x}^{\mathtt{cible}}$(2). Cela est lié à ce que Barry Smyth et Mark Keane appellent la remémoration guidée par l'adaptabilité [25].

2. L'exemple suivant illustre cette idée. Supposons qu'un problème soit une représentation d'un réel $\mathtt{x}$ et que sa solution soit une représentation de $\sin(\mathtt{x})$. Soit à présent $\mathtt{x}^{\mathtt{cible}} = -2$ et $(\mathtt{x}^s, \mathtt{y}^s) \in \mathtt{BC}$ tel que $\mathtt{x}^s = 2$. En utilisant la connaissance du domaine indiquant que sin est une fonction impaire ($\sin(-x) = -\sin(x)$ pour tout $x \in \mathbb{R}$), il apparaît que l'adaptation du cas source en $\mathtt{y}^{\mathtt{cible}} = -\mathtt{y}^s$ donne une bonne solution de $\mathtt{x}^{\mathtt{cible}}$. Par conséquent, dans un sens relatif au RàPC et pour cette application, 2 et −2 sont similaires.

2.3 Spécification de l'exemple suivi

L'exemple utilisé tout au long de cet article se situe dans le domaine de la cuisine. Une recette de cuisine est simplement représentée par les types des ingrédients qu'elle contient. Les variables sont des noms d'aliments représentant des classes de recettes. Par exemple, `fruit` représente la classe des recettes ayant au moins un fruit comme ingrédient. Les connaissances du domaine sont représentées par

$$\begin{aligned}\mathtt{CD} = {} & \mathtt{ananas} \to \mathtt{fruit} \;\wedge\; \mathtt{crème} \to \mathtt{sauce} \\ & \wedge \mathtt{pesto} \to \mathtt{sauce} \;\wedge\; \mathtt{Saint\text{-}Pierre} \to \mathtt{poisson} \\ & \wedge \mathtt{saumon} \to \mathtt{poisson} \;\wedge\; \neg(\mathtt{fruit} \wedge \mathtt{pesto})\end{aligned}$$

Le dernier terme de cette conjonction signifie qu'une recette ne devrait pas contenir à la fois du pesto (qui, lui-même, contient de l'ail) et un fruit (où la notion de fruit est considérée dans un sens culinaire, pas botanique).

Les cas sources considérés dans cet exemple sont :

$$\begin{aligned}\mathtt{cas}^a = {} & \mathtt{Saint\text{-}Pierre} \wedge \mathtt{crème} \wedge \mathtt{vanille} \wedge \mathtt{curry} \\ & \wedge \textit{Rien d'autre} \qquad (1)\\ \mathtt{cas}^b = {} & \mathtt{saumon} \wedge \mathtt{crème} \wedge \mathtt{vanille} \wedge \mathtt{pesto} \wedge \\ & \wedge \textit{Rien d'autre} \qquad (2)\\ \mathtt{cas}^s = {} & \mathtt{Saint\text{-}Pierre} \wedge \mathtt{crème} \wedge \mathtt{ananas} \wedge \\ & \wedge \textit{Rien d'autre} \qquad (3)\end{aligned}$$

où, pour une formule φ, $\varphi \wedge$ *Rien d'autre* dénote la formule $\varphi \wedge \Gamma$, avec Γ la conjonction des littéraux négatifs $\neg a$ tels que $a \in \mathcal{V}$ et $\mathtt{CD} \wedge \varphi \not\models a$ (on peut voir cela comme une application de l'hypothèse du monde clos dans la représentation des cas sources). Par exemple, $\mathtt{cas}^s$ contient le littéral $\neg\mathtt{saumon}$.

Le problème cible représente la requête d'une recette avec du poisson, du pesto, mais pas de vanille :

$$\mathtt{x}^{\mathtt{cible}} = \mathtt{poisson} \wedge \mathtt{pesto} \wedge \neg\mathtt{vanille} \qquad (4)$$

L'ensemble des variables pour décrire les problèmes est donc $\mathcal{V}_{\mathcal{P}} = \{\mathtt{poisson}, \mathtt{pesto}, \mathtt{vanille}\}$. Les autres variables sont utilisées pour décrire les solutions.

2.4 Proportions analogiques et RàPC

Les proportions analogiques sont des énoncés de la forme « a est à b ce que c est à d », notée $a\!:\!b\!::\!c\!:\!d$. Leur origine remonte (au moins) à Aristote [3], et a été inspirée par un parallèle avec la proportion numérique (géométrique) : $\frac{a}{b} = \frac{c}{d}$; voir [20].

En accord avec ce parallèle, elles sont supposées obéir aux postulats suivants (e.g., [13]), pour a, b, c et d, des éléments de l'univers considéré $\mathcal{U}$:

1. $a:b::a:b$ (*réflexivité*) ;
2. si $a:b::c:d$ alors $c:d::a:b$ (*symétrie*) ;
3. si $a:b::c:d$ alors $a:c::b:d$ (*permutation centrale*).

L'unique modèle booléen minimal [22] obéissant à ces 3 postulats est un connecteur quaternaire de la logique propositionnelle, lorsque $\mathcal{U}$ est l'ensemble des booléens $\mathbb{B} = \{0, 1\}$ [17] :

$$a:b::c:d = ((a \wedge \neg b) \leftrightarrow (c \wedge \neg d)) \wedge ((\neg a \wedge b) \leftrightarrow (\neg c \wedge d))$$

Cela rend explicite que « a diffère de b comme c diffère de d (et vice-versa) ». Il est facile de vérifier que cette formule n'est valable que pour les 6 valuations suivantes $0:0::0:0$, $1:1::1:1$, $0:1::0:1$, $1:0::1:0$, $0:0::1:1$, et $1:1::0:0$.

On peut voir que 1 et 0 jouent un rôle symétrique, ce qui rend la définition *indépendante du code* : ceci s'exprime formellement avec l'opérateur de négation comme suit si $a:b::c:d$ alors $\neg a:\neg b::\neg c:\neg d$([3]). Pour traiter des éléments, par exemple des cas, représentés par des *vecteurs* de valeurs booléennes, la définition de la proportion analogique est étendue composante par composante de $\mathcal{U}$ à $\mathcal{U}^n$:

$$a:b::c:d \quad \text{ssi} \quad \text{pour tout } i \in \{1, \ldots, n\}, a_i:b_i::c_i:d_i$$

C'est la base d'un principe d'inférence, proposé pour la première fois dans [18] pour des valeurs nominales, qui peut s'énoncer comme suit :

$$\frac{\forall i \in \{1, ..., m\}, \quad a_i:b_i::c_i:d_i}{\forall j \in \{m+1, ..., n\}, \quad a_j:b_j::c_j:d_j}$$

Comme on peut le voir, la connaissance de certaines composantes des vecteurs sources est transférée à leurs autres composantes, en supposant implicitement que les valeurs des premières composantes déterminent les valeurs des autres.

Cela requiert de trouver ? tel que $a:b::c:?$ soit valide. La solution peut ne pas exister (e.g., pour $0:1::1:?$). Il y a une solution si et seulement si $a = b$ ou $a = c$, dans le cas booléen. L'*unique* solution est alors donnée par $? = c$ si $a = b$ et $? = b$ si $a = c$. Ainsi, nous avons la propriété suivante dans le cas booléen

$$a:a::b:? \text{ si et seulement si } ? = b \tag{5}$$

3. Cette propriété est importante en pratique car elle indique qu'entre deux propriétés contraires, par exemple `marié` et `célibataire`, peu importe celle qui est codée par 1 ou par 0.

qui est quelquefois prise comme un postulat supplémentaire, et qui n'est pas une conséquence des 3 postulats. Ceci est la base de l'*extrapolation* analogique entre cas proposée dans [16].

Étant donné un couple de vecteurs décrivant des problèmes $(\mathtt{x}^a, \mathtt{x}^b)$, avec n composantes, leur comparaison donne une partition des n caractéristiques en deux sous-ensembles : le sous-ensemble des caractéristiques $\mathcal{E}_{(\mathtt{x}^a,\mathtt{x}^b)}$ pour lesquelles les valeurs de $\mathtt{x}^a$ et $\mathtt{x}^b$ sont égales et le sous-ensemble $\mathcal{D}_{(\mathtt{x}^a,\mathtt{x}^b)}$ pour lesquelles elles sont différentes. Considérons deux couples de vecteurs décrivant des problèmes $(\mathtt{x}^a, \mathtt{x}^b)$ et $(\mathtt{x}^c, \mathtt{x}^d)$ tels que $\mathcal{E}_{(\mathtt{x}^a,\mathtt{x}^b)} = \mathcal{E}_{(\mathtt{x}^c,\mathtt{x}^d)}$ et tels que $\forall i \in \mathcal{D}_{(\mathtt{x}^a,\mathtt{x}^b)}, \mathtt{x}^{a_i} = \mathtt{x}^{c_i}$ et $\mathtt{x}^{b_i} = \mathtt{x}^{d_i}$. Il est clair alors $\mathtt{x}^a:\mathtt{x}^b::\mathtt{x}^c:\mathtt{x}^d$ est valide, puisque

(i) pour chaque $j \in \mathcal{E}_{(\mathtt{x}^a,\mathtt{x}^b)} = \mathcal{E}_{(\mathtt{x}^c,\mathtt{x}^d)}$, $(\mathtt{x}^{a_j}, \mathtt{x}^{b_j}, \mathtt{x}^{c_j}, \mathtt{x}^{d_j})$ est de la forme (u, u, v, v) ou (u, u, u, u) pour des valeurs u et v où $u \neq v$ et $u, v \in \{0, 1\}$,

(ii) pour chaque $k \in \mathcal{D}_{(\mathtt{x}^a,\mathtt{x}^b)} = \mathcal{D}_{(\mathtt{x}^c,\mathtt{x}^d)}$, $(\mathtt{x}^{a_k}, \mathtt{x}^{b_k}, \mathtt{x}^{c_k}, \mathtt{x}^{d_k})$ est de la forme (u, v, u, v).

L'idée de considérer des couples de cas peut être reliée à la lecture d'un couple $((\mathtt{x}^a, \mathtt{y}^a), (\mathtt{x}^b, \mathtt{y}^b))$ comme une règle virtuelle exprimant soit que le changement de $\mathtt{x}^a$ à $\mathtt{x}^b$ induit le changement de $\mathtt{y}^a$ à $\mathtt{y}^b$, quel que soit le contexte du problème (encodé par les caractéristiques où $\mathtt{x}^a$ et $\mathtt{x}^b$ sont égaux), soit que le changement de $\mathtt{x}^a$ à $\mathtt{x}^b$ ne modifie pas la solution (dans le cas $\mathtt{y}^a = \mathtt{y}^b$).

Cependant, de telles règles virtuelles peuvent avoir des exceptions dans la base de cas. En effet, il peut exister $(\mathtt{x}^a, \mathtt{x}^b)$, $(\mathtt{x}^c, \mathtt{x}^d)$ et $(\mathtt{x}^{a'}, \mathtt{x}^{b'})$ tels que

— $\mathtt{x}^a:\mathtt{x}^b::\mathtt{x}^c:\mathtt{x}^d$ et $\mathtt{x}^{a'}:\mathtt{x}^{b'}::\mathtt{x}^c:\mathtt{x}^d$ sont vraies. Cela signifie que $\mathcal{D}_{(\mathtt{x}^a,\mathtt{x}^b)} = \mathcal{D}_{(\mathtt{x}^c,\mathtt{x}^d)} = \mathcal{D}_{(\mathtt{x}^{a'},\mathtt{x}^{b'})}$ et que les changements de $\mathtt{x}^a$ à$\mathtt{x}^b$, de $\mathtt{x}^c$ à $\mathtt{x}^d$, de $\mathtt{x}^{a'}$ à $\mathtt{x}^{b'}$ sont les mêmes.
— $\mathtt{y}^a:\mathtt{y}^b::\mathtt{y}^c:\mathtt{y}^d$ est vraie.
— $\mathtt{y}^{a'}:\mathtt{y}^{b'}::\mathtt{y}^{c'}:\mathtt{y}^{d'}$ *ne tient pas* pour une caractéristique i.

Cela peut arriver par exemple quand $\mathtt{y}^{a_i} \neq \mathtt{y}^{b_i}$ tandis que $\mathtt{y}^{a'_i} = \mathtt{y}^{b'_i}$. Dans une telle situation, les deux règles virtuelles associées à $((\mathtt{x}^a, \mathtt{y}^{a_i}), (\mathtt{x}^b, \mathtt{y}^{b_i}))$ et à $((\mathtt{x}^{a'}, \mathtt{y}^{a'_i}), (\mathtt{x}^{b'}, \mathtt{y}^{b'_i}))$ sont en désaccord. On a alors $\mathtt{y}^{a_i} = \mathtt{y}^{c_i}$ et $\mathtt{y}^{b_i} = \mathtt{y}^{d_i}$, mais la solution de l'équation $\mathtt{y}^{a'_i}:\mathtt{y}^{b'_i}::\mathtt{y}^{c_i}:?\mathtt{y}_i$ est $?\mathtt{y}_i = \mathtt{y}^{c_i} \neq \mathtt{y}^{d_i}$.

Ainsi le taux d'exceptions de la règle virtuelle associée à un couple $((\mathtt{x}^a, \mathtt{y}^{a_i}), (\mathtt{x}^b, \mathtt{y}^{b_i}))$ est une indication de l'intérêt du couple pour l'inférence analogique. C'est ce qu'on appelle la compétence du couple de cas [15]. Notez que chaque règle se rapporte à une caractéristique particulière utilisée dans la description des solutions. En effet, il n'y a pas toujours une règle unique qui calcule l'adaptation de $\mathtt{y}^a$ en $\mathtt{y}^b$ à partir du même contexte de problème. Appliquons l'inférence analogique à l'exemple suivi. Il est facile de vérifier que son expression propositionnelle donne naissance au tableau de la figure 1. En considérant les 3 cas $(\mathtt{cas}^a, \mathtt{cas}^b, \mathtt{cas}^s)$, il apparaît que $\mathtt{x}^a:\mathtt{x}^b::\mathtt{x}^s:\mathtt{x}^{\mathtt{cible}}$ est valide pour les caractéristiques `poisson`, `pesto` et `vanille`.

	poisson	pesto	vanille	sauce	fruit	ananas	Saint-Pierre	saumon	crème	curry
$\mathtt{cas}^a$	1	0	1	1	0	0	1	0	1	1
$\mathtt{cas}^b$	1	1	1	1	0	0	0	1	1	0
$\mathtt{cas}^s$	1	0	0	1	1	1	1	0	1	0
$\mathtt{x}^{\mathtt{cible}}$	1	1	0	?	?	?	?	?	?	?
$\mathtt{cas}^{\mathtt{cible}}_{\text{extrap}}$	1	1	0	1	1	1	0	1	1	?

FIGURE 1 – Inférence analogique dans l'exemple suivi.

De plus, l'équation analogique $\mathtt{x}^a : \mathtt{x}^b :: \mathtt{x}^s : \mathtt{x}^{\mathtt{cible}}$ peut être résolue pour les caractéristiques `sauce`, `fruit`, `ananas`, `Saint-Pierre`, `saumon`, et `crème`. On obtient ainsi

$$\begin{array}{rcl} \mathtt{cas}^{\mathtt{cible}}_{\text{extrap}} & = & \mathtt{x}^{\mathtt{cible}} \wedge \mathtt{y}^{\mathtt{cible}}_{\text{extrap}} \\ \text{avec } \mathtt{y}^{\mathtt{cible}}_{\text{extrap}} & \equiv & \mathtt{sauce} \wedge \mathtt{fruit} \wedge \mathtt{ananas} \\ & \wedge & \neg\mathtt{Saint\text{-}Pierre} \wedge \mathtt{saumon} \\ & \wedge & \mathtt{crème} \end{array} \quad (6)$$

Quelques remarques méritent d'être mentionnées :

1. L'équation analogique ne peut pas être résolue pour la caractéristique `curry`, donc rien n'est inféré concernant sa présence ou son absence à partir des cas ($\mathtt{cas}^a, \mathtt{cas}^b, \mathtt{cas}^s$). Dans un exemple plus réaliste (où la base de cas serait plus riche), il pourrait exister un triplet de cas nous permettant de conclure sur `curry`.
2. Cependant, même si une résolution exacte pour `curry` n'existe pas, il serait possible de minimiser une mesure de dissimilarité analogique $AD(a, b, c, d)$ pour calculer une solution « approchée ». $AD(a, b, c, d)$ est égal au nombre minimal de *flips* pour passer de (a, b, c, d) à un quadruplet correspondant à une proportion analogique : $AD(a, b, c, d)$ est maximal (et égal à 2) pour $(1, 0, 0, 1)$ (ou $(0, 1, 1, 0)$) [17]. Ainsi ici, la solution approchée serait 0 (c'est-à-dire pas de `curry`) puisque $AD(1, 0, 0, 0) = 1$.
3. En général, de nombreux triplets de cas peuvent être appliqués à un même $\mathtt{x}^{\mathtt{cible}}$. Rappelons que l'on ne doit utiliser dans de telles situations que les triplets construits sur les couples les plus compétents (dans le cas où ils mènent à des conclusions différentes, un vote doit avoir lieu).
4. Comme on peut le voir dans le calcul de $\mathtt{cas}^{\mathtt{cible}}$, une modification de $\mathtt{cas}^s$ a lieu : à savoir, dans le contexte `poisson` et `¬vanille` avec l'ajout de `pesto`, $\mathtt{cas}^s$ est adapté en changeant `Saint-Pierre` en `saumon`.
5. Dans cet exemple, certaines caractéristiques sont liées par des implications. Il convient de noter que si pour deux caractéristiques mutuellement exclusives i et j, on a $\mathtt{x}^{a_i} : \mathtt{x}^{b_i} :: \mathtt{x}^{c_i} : \mathtt{x}^{d_i}$ et $\mathtt{x}^{a_j} : \mathtt{x}^{b_j} :: \mathtt{x}^{c_j} : \mathtt{x}^{d_j}$, cela implique que c'est également vrai pour une caractéristique k pour laquelle i et j sont des sous-classes, comme on peut le voir sur l'exemple.
6. On peut observer que les rôles de $\mathtt{cas}^b$ et $\mathtt{cas}^s$ pourraient être échangés, puisque les proportions analogiques sont stables par permutation centrale.

2.5 La révision des croyances et le RàPC

La révision des croyances. Considérons un agent ayant un ensemble de croyances ψ et qui est confronté à un autre ensemble de croyances μ, lequel est supposé être prioritaire par rapport à ψ. Dans cet article, ψ et μ sont supposés être représentés en logique propositionnelle. La question soulevée par la révision des croyances est « Comment les croyances de l'agent évoluent-elles par incorporation de μ ? » Quand le nouvel ensemble de croyances n'est pas en contradiction avec l'ancien — i.e., $\psi \wedge \mu$ est cohérent — alors la révision donne simplement la conjonction $\psi \wedge \mu$. Sinon, selon le principe de changement minimal de la théorie AGM (nommée d'après les initiales des auteurs de [2]), la révision des croyances consiste à faire un « changement minimal » de ψ en ψ' de telle sorte que $\psi' \wedge \mu$ soit cohérent, le résultat de la révision est alors $\psi' \wedge \mu$, noté $\psi \dotplus \mu$ dans la suite. Cependant, la notion de minimalité du changement n'est pas définie de façon unique et dépend de la manière dont le changement est évalué. Par conséquent, de nombreux *opérateurs de révision* $\dotplus$ existent. Les postulats AGM proposent un ensemble de postulats qu'un tel opérateur $\dotplus$ devrait respecter. Ces postulats ont été formulés en logique propositionnelle par [12] de la façon suivante (pour toutes formules φ, ψ, φ_1, φ_2, ψ_1, ψ_2 et χ) :

($\dotplus$1) Si μ est cohérent alors $\psi \dotplus \mu$ est cohérent.

($\dotplus$2) Si $\psi \wedge \mu$ est cohérent alors $\psi \dotplus \mu \equiv \psi \wedge \mu$.

($\dotplus$3) $\psi \dotplus \mu \models \mu$.

($\dotplus$4) Si $\psi_1 \equiv \psi_2$ et $\mu_1 \equiv \mu_2$ alors $\psi_1 \dotplus \mu_1 \equiv \psi_2 \dotplus \mu_2$.

($\dotplus$5) Si $(\psi \dotplus \mu) \wedge \chi$ est cohérent
alors $\psi \dotplus (\mu \wedge \chi) \models (\psi \dotplus \mu) \wedge \chi$.

($\dotplus$6) $(\psi \dotplus \mu) \wedge \chi \models \psi \dotplus (\mu \wedge \chi)$.

($\dotplus$1) affirme qu'un agent vise à avoir des croyances cohérentes (à moins d'accepter un ensemble de croyances μ incohérent). ($\dotplus$2) est lié au principe du changement minimal : si ψ est cohérent avec μ alors le changement minimal $\psi \mapsto \psi'$ est $\psi' = \psi$ (i.e., pas de changement). ($\dotplus$3) est lié au fait que μ est prioritaire par rapport à ψ, i.e., les seuls changements de croyances effectuées le sont sur ψ : μ est inchangé. ($\dotplus$4) affirme que la révision doit respecter le principe d'indépendance à la syntaxe (substituer une formule par une formule équivalente doit conserver le résultat de l'inférence, à l'équivalence logique près). On peut montrer que la conjonction des postulats ($\dotplus$5) et ($\dotplus$6) est équivalente à l'assertion suivante : si $(\psi \dotplus \mu) \wedge \chi$ est cohérent alors $\psi \dotplus (\mu \wedge \chi) \equiv (\psi \dotplus \mu) \wedge \chi$. En d'autres termes, s'il n'y a pas de nécessité de modifier les croyances après la révision de ψ par μ pour incorporer χ, alors aucune modification additionnelle n'est effectuée : les nouvelles croyances χ sont simplement ajoutées aux croyances $\psi \dotplus \mu$. C'est également lié au principe du changement minimal : quand aucun changement n'est nécessaire pour restaurer la cohérence, alors aucun changement n'est effectué.

Malgré ces postulats, l'ensemble des opérateurs de révision des croyances reste très grand et dépend de la façon dont le changement est évalué. En particulier, ce changement peut être évalué grâce à une mesure de similarité entre interprétations, ce qui permet de définir la famille d'opérateurs de révision des croyances présentée ci-dessous.

Opérateurs de révision fondée sur des mesures de similarité. Une mesure de similarité sur un ensemble E est définie dans cet article comme étant une fonction $\mathtt{sim} : E \times E \to [0,1]$ telle que $\mathtt{sim}(a,b) = 1$ ssi $a = b$ (pour $a, b \in E$). Un opérateur de révision des croyances $\dotplus^{\mathtt{sim}}$ satisfaisant les postulats AGM peut être défini pour toute mesure de similarité $\mathtt{sim}$ sur Ω, l'ensemble des interprétations, de la façon suivante :

$$\text{avec } \mathtt{sim}^* = \max \left\{ \mathtt{sim}(\mathcal{I},\mathcal{J}) \;\middle|\; \begin{array}{l} \mathcal{I} \in \mathcal{M}(\psi) \text{ et} \\ \mathcal{J} \in \mathcal{M}(\mu) \end{array} \right\}$$

$$\mathcal{M}(\psi \dotplus^{\mathtt{sim}} \mu) = \left\{ \mathcal{J} \in \mathcal{M}(\mu) \;\middle|\; \max_{\mathcal{I} \in \mathcal{M}(\psi)} \mathtt{sim}(\mathcal{I},\mathcal{J}) = \mathtt{sim}^* \right\} \tag{7}$$

En d'autres termes, les modèles de $\psi \dotplus^{\mathtt{sim}} \mu$ sont les modèles de μ les plus similaires aux modèles de ψ, selon $\mathtt{sim}$. Notons que cela ne permet de définir $\dotplus^{\mathtt{sim}}$ que modulo l'équivalence logique, ce qui ne soulève pas de problème : toute formule ϱ telle que $\varrho \equiv \psi \dotplus^{\mathtt{sim}} \mu$ constitue une $\dotplus^{\mathtt{sim}}$-révision de ψ par μ selon le principe d'indépendance à la syntaxe.

Étant donné une fonction $\mathtt{dist}$ sur Ω, une mesure de similarité $\mathtt{sim}$ peut être définie par $\mathtt{sim}(\mathcal{I},\mathcal{J}) = 1/(1 + \mathtt{dist}(\mathcal{I},\mathcal{J}))$ pour $\mathcal{I},\mathcal{J} \in \Omega$. En particulier, soit H la distance de Hamming entre interprétations : $H(\mathcal{I},\mathcal{J})$ est le nombre de variables a telles que $\mathcal{I}(a) \neq \mathcal{J}(a)$. Soit à présent $\mathtt{sim}_H$ la mesure de similarité associée à H. L'opérateur de révision des croyances $\dotplus^{\mathtt{sim}_H}$ coincide avec l'opérateur de Dalal [5] et sera noté $\dotplus_{\text{Dalal}}$ dans la suite. La distance de Hamming pondère chaque variable de la même façon et considère les variables indépendamment les unes des autres : elle peut être vue comme une distance d'édition sur des interprétations fondée sur l'opérateur *flip*. Pour cette raison, quand il n'y a aucune connaissance explicitant comment les changements doivent être évalués, la distance de Hamming étant « neutre » est utilisée pour mesurer ces changements. C'est pourquoi $\dotplus_{\text{Dalal}}$ est utilisé comme un opérateur « non informé », avec une connaissance sur le changement vide.

L'adaptation par révision est une approche d'adaptation simple fondée sur un opérateur de révision $\dotplus$. L'intuition est que la modification du cas remémoré $\mathtt{x}^s \wedge \mathtt{y}^s$ afin de proposer une solution au problème cible $\mathtt{x}^{\mathtt{cible}}$ est effectuée par $\dotplus$. Le cas remémoré et le problème cible sont interprétés en accord avec les connaissances du domaine, la révision qui doit être effectuée est donc celle de $\mathtt{CD} \wedge \mathtt{x}^s \wedge \mathtt{y}^s$ par $\mathtt{CD} \wedge \mathtt{x}^{\mathtt{cible}}$. Le résultat de cette révision est une formule ϱ qui entraîne $\mathtt{x}^{\mathtt{cible}}$ (selon le postulat ($\dotplus$3)) et donc, ϱ est équivalent à une formule $\mathtt{x}^{\mathtt{cible}} \wedge \mathtt{y}^{\mathtt{cible}}$ où toutes les variables de $\mathtt{y}^{\mathtt{cible}}$ appartiennent à $\mathcal{V}_{\mathcal{S}}$. $\mathtt{y}^{\mathtt{cible}}$ est la solution proposée pour $\mathtt{x}^{\mathtt{cible}}$. Formellement, cela s'écrit comme suit :

$$(\mathtt{CD} \wedge \mathtt{x}^s \wedge \mathtt{y}^s) \;\dotplus\; (\mathtt{CD} \wedge \mathtt{x}^{\mathtt{cible}}) \quad \equiv \quad \mathtt{x}^{\mathtt{cible}} \wedge \mathtt{y}^{\mathtt{cible}} \tag{8}$$

Afin d'appliquer l'adaptation par révision, un opérateur de révision doit être choisi. Ce choix est lié à la façon dont le changement est évalué, c'est-à-dire, en terme de RàPC, à la connaissance d'adaptation. En particulier, quand aucune connaissance d'adaptation n'est disponible, l'opérateur de révision de Dalal est utilisé.

On peut noter que la solution $\mathtt{y}^{\mathtt{cible}}$ donnée par adaptation par révision est nécessairement cohérente pour peu que $\mathtt{CD} \wedge \mathtt{x}^{\mathtt{cible}}$ le soit, mais elle n'est pas nécessairement complètement décrite : $\mathcal{M}\left(\mathtt{CD} \wedge \mathtt{x}^{\mathtt{cible}} \wedge \mathtt{y}^{\mathtt{cible}}\right)$ peut contenir plusieurs interprétations. Dans une telle situation, cela signifie que l'adaptation par révision affirme qu'il existe une solution plausible $\mathtt{y}$ de $\mathtt{x}^{\mathtt{cible}}$ vérifiant $\mathtt{y} \models \mathtt{y}^{\mathtt{cible}}$. Dans le cas extrême, $\mathtt{y}^{\mathtt{cible}} \equiv \top$, ce qui signifie que l'adaptation par révision ne donne aucune information sur une solution potentielle de $\mathtt{x}^{\mathtt{cible}}$.

L'exemple suivi peut être résolu en utilisant l'adaptation par révision. On suppose que le cas remémoré est $\mathtt{cas}^s$ et que le problème cible est $\mathtt{x}^{\mathtt{cible}}$, définis par (3) et (4). On considérera dans cette section que l'adaptation utilise $\dotplus_{\text{Dalal}}$. On peut alors montrer que le résultat $\mathtt{y}^{\mathtt{cible}}$ de cette révision est

$$y^{\texttt{cible}}_{\text{Dalal}} \equiv \texttt{Saint-Pierre} \wedge \texttt{crème} \wedge \texttt{sauce} \wedge \neg\texttt{fruit} \wedge \neg\texttt{ananas} \wedge \neg\texttt{saumon} \wedge \neg\texttt{curry} \quad (9)$$

Ainsi, cette adaptation consiste à enlever de $\texttt{x}^s \wedge \texttt{y}^s$ les fruits, puisque leur présence serait incohérente avec $\texttt{CD} \wedge \texttt{x}^{\texttt{cible}}$ (à cause du pesto).

3 Un pont entre l'extrapolation et l'adaptation par révision

Les deux approches de l'adaptation des cas présentées ci-dessus — celle fondée sur l'extrapolation analogique et celle fondée sur la révision des croyances — apparaissent de façon assez différente : la première est une adaptation multiple de cas alors que la seconde est une application simple de cas (elles s'appuient respectivement sur la remémoration de cas sources par triplets et par singletons). Néanmoins, le but de cette section est de montrer comment ces deux approches peuvent se rencontrer. Premièrement, l'approche fondée sur l'extrapolation est reformulée comme une adaptation simple. Deuxièmement, un opérateur de révision fondé sur des couples de cas est défini et il est montré comment, sous certaines hypothèses, les deux approches de l'adaptation coïncident. Cela permet de définir une approche de l'adaptation fondée à la fois sur l'extrapolation et la révision qui prend en compte, d'un côté, la base de cas et les compétences des couples de cas et, d'un autre côté, les connaissances du domaine.

3.1 Reformuler l'adaptation par extrapolation comme une adaptation simple

Dans la présentation faite ci-dessus de l'extrapolation analogique, on considérait qu'un triplet de cas $(\texttt{cas}^a, \texttt{cas}^b, \texttt{cas}^c)$ était remémoré puis réutilisé pour résoudre $\texttt{x}^{\texttt{cible}}$. À présent, cela va être reformulé en considérant que seul le cas $\texttt{cas}^c$ est remémoré, et que les autres, $\texttt{cas}^a$ et $\texttt{cas}^b$, sont sélectionnés durant le processus d'adaptation lui-même. Cette « brisure de symétrie » a deux avantages. D'abord, elle peut être utilisée pour une implantation efficace de l'extrapolation analogique (ce point a été détaillé dans les algorithmes d'extrapolation décrits dans [16]). Ensuite, cela facilite l'association des deux approches de l'adaptation. Pour cette raison, le cas remémoiré $\texttt{cas}^c$ est renommé $\texttt{cas}^s = (\texttt{x}^s, \texttt{y}^s)$, pour mieux correspondre avec les notations de l'adaptation simple.

Par conséquent, la reformulation de l'extrapolation analogique en une adaptation simple se fait comme suit :

Entrée : un cas à adapter $(\texttt{x}^s, \texttt{y}^s)$, le problème cible $\texttt{x}^{\texttt{cible}}$, la base de cas BC, la relation de préférence entre couples de cas ;

Sortie : un ensemble Y de solutions proposées pour $\texttt{x}^{\texttt{cible}}$

1. Soit CouplesCandidats l'ensemble des $(\texttt{cas}^a, \texttt{cas}^b) \in \texttt{BC} \times \texttt{BC}$ tels que $\texttt{x}^a : \texttt{x}^b :: \texttt{x}^s : \texttt{x}^{\texttt{cible}}$ et tels que l'équation analogique $\texttt{y}^a : \texttt{y}^b :: \texttt{y}^s : ?\texttt{y}$ admet une solution.
2. Soit MeilleursCouplesCandidats l'ensemble des couples de cas les plus compétents de CouplesCandidats.
3. Soit $\texttt{Y} = \left\{ \texttt{y} \;\middle|\; \begin{array}{l} \texttt{y} \text{ est la solution de } \texttt{y}^a : \texttt{y}^b :: \texttt{y}^s : ?\texttt{y} \\ \text{pour } (\texttt{cas}^a, \texttt{cas}^b) \in \\ \quad \texttt{MeilleursCouplesCandidats} \end{array} \right\}$.
4. Y est retourné comme ensemble de solutions candidates pour $\texttt{x}^{\texttt{cible}}$.

3.2 Un opérateur de révision fondé sur la compétence de couples de cas

En s'appuyant sur la compétence des couples de cas et sous certaines hypothèses, on peut définir une mesure de similarité $\texttt{sim}_{\texttt{comp}}$ sur Ω, ce qui permet d'introduire l'opérateur de révision $\dotplus^{\texttt{sim}_{\texttt{comp}}}$. Il est notable que pour les préordres de compétences présentés dans [15], ces hypothèses sont vérifiées.

La première hypothèse est que le préordre de compétence entre couples de cas peut être caractérisé grâce à un *niveau de compétence*, i.e., une fonction nivComp qui associe à un couple de cas sources une valeur de [0, 1] telle que $(\texttt{cas}^1, \texttt{cas}^2)$ est considéré comme étant strictement plus compétent que $(\texttt{cas}^3, \texttt{cas}^4)$ ssi $\texttt{nivComp}(\texttt{cas}^1, \texttt{cas}^2) > \texttt{nivComp}(\texttt{cas}^3, \texttt{cas}^4)$.

La deuxième hypothèse traduit simplement le fait que deux couples de cas reliés par des proportions analogiques (dans l'espace des problèmes et l'espace des solutions) ont la même compétence et donc le même niveau de compétence :

$$\begin{array}{rl} \textbf{si} & \texttt{x}^a : \texttt{x}^b :: \texttt{x}^c : \texttt{x}^d \textbf{ et } \texttt{y}^a : \texttt{y}^b :: \texttt{y}^c : \texttt{y}^d \\ \textbf{alors} & \texttt{nivComp}(\texttt{cas}^a, \texttt{cas}^b) = \texttt{nivComp}(\texttt{cas}^c, \texttt{cas}^d) \end{array} \quad (10)$$

La troisième hypothèse est que le niveau maximum de compétence est 1 et est atteint seulement par les couples $(\texttt{cas}^a, \texttt{cas}^a)$ pour $\texttt{cas}^a \in \texttt{BC}$. Cette troisième hypothèse peut être justifiée par le fait que si $\texttt{x}^a : \texttt{x}^a :: \texttt{x}^s : \texttt{x}^{\texttt{cible}}$ alors $\texttt{x}^s = \texttt{x}^{\texttt{cible}}$ (d'après (5)) et donc que l'équation analogique $\texttt{y}^a : \texttt{y}^a :: \texttt{y}^s : ?\texttt{y}$ a exactement une solution $?\texttt{y} = \texttt{y}^s$ qui résout $\texttt{x}^{\texttt{cible}} = \texttt{x}^s$.

La quatrième hypothèse est que le niveau de compétence est minimal pour $(\texttt{cas}^a, \texttt{cas}^b)$ si et seulement si ces deux cas ne sont en analogie avec aucun couple de cas $(\texttt{cas}^c, \texttt{cas}^d)$ de la base de cas (en particulier, au moins un cas parmi $\texttt{cas}^a$ et $\texttt{cas}^b$ n'est pas un cas de la base de cas).

Une bijection casDe entre Ω et l'ensemble des cas complètement décrits peut être définie, pour $\mathcal{I} \in \Omega$, par $\texttt{casDe}(\mathcal{I}) = (\texttt{x}, \texttt{y})$ est tel que $\mathcal{M}(\texttt{x} \wedge \texttt{y}) = \{\mathcal{I}\}$. On peut alors définir une mesure de similarité $\mathtt{sim_{comp}}$ utilisant cette bijection et la fonction donnant le niveau de compétence :

$$\mathtt{sim_{comp}}(\mathcal{I},\mathcal{J}) = \mathtt{nivComp}(\mathtt{casDe}(\mathcal{I}),\mathtt{casDe}(\mathcal{J})) \quad (\text{pour } \mathcal{I},\mathcal{J}\in\Omega)$$

La rencontre entre les deux approches de l'adaptation peut être exprimée par le résultat suivant (étant donné un cas source $(\mathtt{x}^s,\mathtt{y}^s)$ et un problème cible $\mathtt{x}^{\mathtt{cible}}$) :

$$\left.\begin{array}{ll}\textbf{si} & \mathtt{y}^{\mathtt{cible}} \text{ est le résultat de l'adaptation}\\ & \text{par révision avec } \dotplus = \dotplus^{\mathtt{sim_{comp}}},\\ & \text{les connaissances du domaine sont vides}\\ & (\mathtt{CD} = \top),\\ & \mathtt{Y} \text{ est l'ensemble des solutions obtenues}\\ & \text{par extrapolation analogique}\\ & \text{et } \mathtt{Y}\neq\emptyset\\ \textbf{alors} & \mathtt{y}^{\mathtt{cible}} \equiv \bigvee \mathtt{Y}\end{array}\right| \qquad (11)$$

En d'autres termes, en l'absence de connaissances du domaine, l'adaptation par révision utilisant l'opérateur de révision $\dotplus^{\mathtt{sim_{comp}}}$ qui s'appuie sur la compétence de couples de cas donne le même résultat que l'extrapolation analogique, à moins que cette dernière ne donne aucune solution. De plus, on peut montrer que si cet ensemble est vide, alors $\mathtt{y}^{\mathtt{cible}} \equiv \top$, i.e., dans ce cas, l'adaptation par révision ne donne aucune information sur la solution.

3.3 Une approche de l'adaptation fondée sur l'extrapolation et la révision

À présent, considérons l'application de l'adaptation par $\dotplus^{\mathtt{sim_{comp}}}$ pour l'exemple suivi et, dans un premier temps, avec $\mathtt{CD}$ vide. À cause de la variable $\mathtt{curry}$, l'extrapolation analogique ne donne aucune solution [4] ($\mathtt{Y}=\emptyset$) et, dans cette situation, $\mathtt{y}^{\mathtt{cible}} \equiv \top$ (aucune information sur la solution n'est proposée). Si la variable $\mathtt{curry}$ est supprimée de $\mathcal{V}$ alors, selon (11), la solution proposée vérifie

$$\mathtt{y}^{\mathtt{cible}} = \mathtt{y}^{\mathtt{cible}}_{\dotplus^{\mathtt{sim_{comp}}}\text{-adaptation sans curry}} \equiv \mathtt{y}^{\mathtt{cible}}_{\text{extrap}} \qquad (12)$$

À présent, afin de définir une approche qui considère toutes les variables, même celles telles que $\mathtt{curry}$ pour laquelle l'extrapolation ne s'applique pas exactement, une mesure de similarité $\mathtt{sim_{E\&D}}$ peut être définie qui combine $\mathtt{sim_{comp}}$ et $\mathtt{sim}_H$ (pour un α donné, $0<\alpha<1$, et $\mathcal{I},\mathcal{J}\in\Omega$) :

$$\mathtt{sim_{E\&D}}(\mathcal{I},\mathcal{J}) = (1-\alpha)\mathtt{sim_{comp}}(\mathcal{I},\mathcal{J}) + \alpha\mathtt{sim}_H(\mathcal{I},\mathcal{J}) \qquad (13)$$

d'où l'opérateur de révision $\dotplus^{\mathtt{sim_{E\&D}}}$ (E&D pour « extrapolation et Dalal »). Donc, pour un α suffisamment petit, l'adaptation par $\dotplus^{\mathtt{sim_{E\&D}}}$ consiste à faire un nombre minimal de *flips* de variables sur $\mathtt{cas}^s$ pour rendre l'extrapolation possible puis l'appliquer [5].

4. Plus précisément, l'équation analogique $\mathtt{y}^a:\mathtt{y}^b::\mathtt{y}^s:\mathtt{y}$ n'a pas de solution : cette équation est résoluble attribut par attribut sur tous les attributs à l'exception de $\mathtt{curry}$, donc une solution peut être proposé pour tous les autres attributs.

5. On peut montrer qu'il est suffisant d'avoir $\alpha < (2\,|\mathcal{V}|)^{-1}$.

L'application de l'adaptation par $\dotplus^{\mathtt{sim_{E\&D}}}$ sur l'exemple suivi donne :

$$\begin{array}{rcl}\mathtt{y}^{\mathtt{cible}}_{\text{E\&D}} & \equiv & \mathtt{sauce} \wedge \mathtt{fruit} \wedge \mathtt{ananas}\\ & \wedge & \neg\mathtt{Saint\text{-}Pierre} \wedge \mathtt{saumon}\\ & \wedge & \mathtt{crème} \wedge \neg\mathtt{curry}\end{array} \qquad (14)$$

On peut expliquer cela comme suit. L'équation analogique $1:0::0:?\mathtt{y_{curry}}$ pour trouver la valeur de l'attribut $\mathtt{curry}$ dans $\mathtt{y}^{\mathtt{cible}}$ par extrapolation de $(\mathtt{cas}^a,\mathtt{cas}^b,\mathtt{cas}^s)$ n'a pas de solution, donc un *flip* de cet attribut pour $\mathtt{cas}^s$ donne $\mathtt{cas}^{s\prime}$ et le triplet $(\mathtt{cas}^a,\mathtt{cas}^b,\mathtt{cas}^{s\prime})$ peut être utilisé par extrapolation pour résoudre $\mathtt{x}^{\mathtt{cible}}$ en $\mathtt{y}^{\mathtt{cible}} = \mathtt{y}^{\mathtt{cible}}_{\text{E\&D}}$, puisque $1:0::1:?\mathtt{y_{curry}}$ a une solution unique $?\mathtt{y_{curry}}=0$ (d'où le $\neg\mathtt{curry}$ dans la solution proposée de $\mathtt{x}^{\mathtt{cible}}$).

Cette adaptation n'a pas tenu compte des connaissances du domaine $\mathtt{CD}$ définies dans la section 2.3 et, de fait, $\mathtt{CD} \wedge \mathtt{x}^{\mathtt{cible}} \wedge \mathtt{y}^{\mathtt{cible}}_{\text{E\&D}}$ est incohérent avec ces connaissances du domaine (à cause du conflit pesto-fruit). Par conséquent, l'approche de l'adaptation proposée consiste à faire une adaptation par $\dotplus^{\mathtt{sim_{E\&D}}}$ prenant en compte $\mathtt{CD}$, selon (8). Dans le cadre de l'exemple suivi, cela donne :

$$\begin{array}{rcl}\mathtt{y}^{\mathtt{cible}}_{\text{E\&D avec CD}} & \equiv & \mathtt{sauce} \wedge \neg\mathtt{fruit} \wedge \neg\mathtt{ananas}\\ & \wedge & \neg\mathtt{Saint\text{-}Pierre} \wedge \mathtt{saumon}\\ & \wedge & \mathtt{crème} \wedge \neg\mathtt{curry}\end{array} \qquad (15)$$

qui consiste à enlever les fruits de $\mathtt{y}^{\mathtt{cible}}_{\text{E\&D}}$.

3.4 Synthèse

La figure 2 décrit les cas sources et le problème cible de l'exemple suivi, ainsi que les cas proposés $\mathtt{cas}^s = \mathtt{x}^{\mathtt{cible}} \wedge \mathtt{y}^{\mathtt{cible}}$ après les différents processus d'adaptation.

Cet exemple illustre comment les atouts des deux approches d'adaptation peuvent être combinés :

— La force de l'extrapolation analogique est d'exploiter les variations dans la base de cas (variations pouvant être considérées comme des règles d'adaptation spécifiques) : la variation « Saint-Pierre vers saumon » de $\mathtt{cas}^a$ à $\mathtt{cas}^b$ est appliquée à $\mathtt{cas}^s$ dans un contexte similaire.

— La force de l'adaptation par $\dotplus_{\text{Dalal}}$ est de prendre en compte les connaissances du domaine afin d'ajuster le cas remémoré pour proposer une solution au problème cible : le pesto étant incompatible avec les fruits, l'ananas est retiré de la recette [6].

— L'adaptation par $\dotplus^{\mathtt{sim_{comp}}}$ combine les deux approches d'adaptation : elle applique une extrapolation analogique sur chaque caractéristique pour laquelle c'est à la fois possible et cohérent avec les connaissances du domaine, et ajuste les autres caractéristiques avec l'adaptation par $\dotplus_{\text{Dalal}}$.

6. Il est à noter que l'adaptation $\dotplus_{\text{Dalal}}$ peut faire plus que supprimer les faits positifs car elle peut remplacer une classe par une classe sœur dans la taxonomie (voir, par exemple, [14]).

n°§		poisson	pesto	vanille	sauce	fruit	ananas	Saint-Pierre	saumon	crème	curry
2.3	$\mathtt{cas}^a$	1	0	1	1	0	0	1	0	1	1
	$\mathtt{cas}^b$	1	1	1	1	0	0	0	1	1	0
	$\mathtt{cas}^s$	1	0	0	1	1	1	1	0	1	0
	$\mathtt{x}^{\mathtt{cible}}$	1	1	0	?	?	?	?	?	?	?
2.4	$\mathtt{cas}^{\mathtt{cible}}_{\text{extrap}}$	1	1	0	1	1	1	0	1	1	?
2.5	$\mathtt{cas}^{\mathtt{cible}}_{\text{Dalal}}$	1	1	0	1	0	0	1	0	1	0
3.3	$\mathtt{cas}^{\mathtt{cible}}_{\text{E\&D}}$	1	1	0	1	1	1	0	1	1	0
3.3	$\mathtt{cas}^{\mathtt{cible}}_{\text{E\&D avec CD}}$	1	1	0	1	0	0	0	1	1	0

FIGURE 2 – L'exemple suivi : spécification du problème et résultats $\mathtt{cas}^{\mathtt{cible}} = \mathtt{x}^{\mathtt{cible}} \wedge \mathtt{y}^{\mathtt{cible}}$ des processus d'adaptation présentés dans l'article. La colonne n°§ indique le numéro de la section pertinente.

4 Travaux proches et remarques finales

Cet article a examiné deux approches très différentes pour l'adaptation des cas — l'extrapolation analogique et l'adaptation fondée sur la révision — et a étudié la question de savoir comment elles peuvent se rencontrer. Il a été montré que, dans certaines circonstances (dans le cadre de la logique propositionnelle, sans connaissance du domaine, etc.), elles coïncident (cf. (11)) et que l'approche peut être étendue lorsque des connaissances du domaine sont ajoutées et/ou lorsque les proportions analogiques peuvent être trouvées seulement pour certaines composantes de la solution. L'idée est que les couples de cas et leurs compétences — utilisées dans extrapolation analogique — peuvent être utilisés pour « remodeler l'espace d'adaptation » en rendant plus similaire le cas source et le problème cible, et que cette similarité est utilisée par l'opérateur de révision.

Travaux connexes. Il existe une riche littérature sur la révision des croyances suite aux travaux fondateurs d'Alchourrón, Gärdenfors et Makinson [2], et son expression dans une logique propositionnelle [12]. Cependant, l'idée d'appliquer la révision des croyances au RàPC, telle que réaffirmée dans les préliminaires de cet article, ne peut être trouvée que dans quelques travaux (voir en particulier [4]).

L'idée d'un modèle d'inférence analogique fondé sur des proportions analogiques remonte à [18]. Son application au RàPC est suggérée dans [21], et étudiée plus en profondeur dans [16].

Il n'y a eu jusqu'à présent aucun travail reliant l'extrapolation analogique et la révision des croyances. En gardant à l'esprit que la révision des croyances et la logique non monotone sont, dans un certain sens, les deux faces d'une même pièce [10], nous pouvons cependant mentionner une discussion [19] opposant raisonnement non monotone et raisonnement analogique, mais aussi donnant des passerelles entre les deux.

Perspectives. L'étude présentée dans cet article montre que deux modèles d'adaptation peuvent se rencontrer, mais une question restante concerne l'utilité pratique de cette rencontre, au-delà de l'exemple courant. Dans ce but, une première perspective consiste à mener une expérimentation pour comparer les différentes approches présentées.

Les opérateurs de révision $\dotplus^{\mathtt{sim}_{\text{comp}}}$ et $\dotplus^{\mathtt{sim}_{\text{E\&D}}}$ peuvent changer avec l'ajout d'un nouveau cas dans la base de cas, par exemple, le cas $(\mathtt{x}^{\mathtt{cible}}, \mathtt{y}^{\mathtt{cible}})$ lorsque ce nouveau cas est validé. En effet, le niveau de compétence des couples de cas est calculé sur la base de $\mathtt{BC}$. L'évolution des opérateurs de révision au fil du temps est un problème lié à une révision itérée (voir, par exemple, [7]) ; il serait alors intéressant d'étudier $\dotplus^{\mathtt{sim}_{\text{comp}}}$ et $\dotplus^{\mathtt{sim}_{\text{E\&D}}}$ à la lumière des postulats de la révision itérée.

Cette rencontre entre adaptation par proportions analogiques et adaptation fondée sur la révision des croyances peut également être étudiée plus globalement pour la résolution de problèmes par RàPC (i.e., remémoration et adaptation) : cela constitue la troisième direction de travail.

Ce travail contribue à l'idée selon laquelle l'adaptation par révision recouvre une grande variété d'approches de l'adaptation : en l'occurrence, on peut choisir un opérateur de révision des croyances pour que l'adaptation par extrapolation analogique rencontre l'adaptation par révision. D'autres travaux en cours montrent que d'autres approches de l'adaptation sont également couvertes par l'adaptation par révision. Cela ne rend pas ces travaux caduques : l'approche de l'adaptation par révision ne dit pas grand chose sur le choix de l'opérateur de révision.

Enfin, on sait que la révision des croyances peut être encodée en logique possibiliste [9], puisque la révision des croyances repose au niveau sémantique sur des relations d'enracinement épistémiques [11], qui ne sont que des relations de nécessité qualitative au sens de la théorie des possibilités [8]. Comment traiter les différents types de révision/adaptation envisagés dans cet article dans le cadre possibiliste est un autre sujet de recherche future.

Références

[1] Aamodt, A. et E. Plaza: *Case-based Reasoning : Foundational Issues, Methodological Variations, and System Approaches*. AI Communications, 7(1) :39–59, 1994.

[2] Alchourrón, C. E., P. Gärdenfors et D. Makinson: *On the logic of theory change : partial meet functions for contraction and revision*. J. Symbolic Logic, 50 :510–530, 1985.

[3] Aristotle: *Nicomachean Ethics*. Univ. of Chicago Press, 2011. Trans. by R. C. Bartlett and S. D. Collins.

[4] Cojan, J. et J. Lieber: *Applying belief revision to case-based reasoning*. Dans Prade, H. et G. Richard (rédacteurs) : *Computational Approaches to Analogical Reasoning : Current Trends*, tome 548 de *Studies in Computational Intelligence*, pages 133–161. Springer, 2014.

[5] Dalal, M.: *Investigations into a theory of Knowledge Base Revision : Preliminary Report*. Dans *Proceedings of the Seventh National Conference on Artificial Intelligence (AAAI)*, pages 475–479, 1988.

[6] d'Aquin, M., F. Badra, S. Lafrogne, J. Lieber, A. Napoli et L. Szathmary: *Case base mining for adaptation knowledge acquisition*. Dans Veloso, M. M. (rédacteur) : *IJCAI 2007, Proc. of the 20th Int. Joint Conf. on Artificial Intelligence, Hyderabad, Jan. 6-12*, pages 750–755, 2007.

[7] Darwiche, A. et J. Pearl: *On the logic of iterated belief revision*. Artificial intelligence, 89(1-2) :1–29, 1997.

[8] Dubois, D. et H. Prade: *Epistemic entrenchment and possibilistic logic*. Artif. Intell., 50(2) :223–239, 1991.

[9] Dubois, D. et H. Prade: *Possibilistic logic - An overview*. Dans Siekmann, J. H. (rédacteur) : *Computational Logic*, tome 9 de *Handbook of the History of Logic*, pages 283–342. Elsevier, 2014.

[10] Gärdenfors, P.: *Belief revision and nonmonotonic logic : Two sides of the same coin ?* Dans *Proc. 9th Europ. Conf. on Artificial Intelligence (ECAI'90), Stockholm*, pages 768–773, 1990.

[11] Gärdenfors, P. et D. Makinson: *Revisions of knowledge systems using epistemic entrenchment*. Dans Vardi, M. Y. (rédacteur) : *Proc. 2nd Conf. on Theoretical Aspects of Reasoning about Knowledge, Pacific Grove*, pages 83–95. Morgan Kaufmann, 1988.

[12] Katsuno, H. et A. Mendelzon: *Propositional knowledge base revision and minimal change*. Artificial Intelligence, 52(3) :263–294, 1991.

[13] Lepage, Y.: *Analogy and formal languages*. Electr. Notes Theor. Comput. Sci., 53, 2001.

[14] Lieber, J.: *Application of the revision theory to adaptation in case-based reasoning : The conservative adaptation*. Dans Weber, R. et M. M. Richter (rédacteurs) : *Case-Based Reasoning Research and Development, 7th Int. Conf. on Case-Based Reasoning, ICCBR 2007, Belfast, Aug. 13-16, Proceedings*, tome 4626 de *LNCS*, pages 239–253. Springer, 2007.

[15] Lieber, J., E. Nauer et H. Prade: *Improving analogical extrapolation using case pair competence*. Dans *Case-Based Reasoning Research and Development, 27th International Conference (ICCBR-2019)*, Otzenhausen, France, septembre 2019.

[16] Lieber, J., E. Nauer, H. Prade et G. Richard: *Making the best of cases by approximation, interpolation and extrapolation*. Dans *ICCBR 2018 - 26th International Conference on Case-Based Reasoning*, Stockholm, Sweden, juillet 2018. Springer.

[17] Miclet, L. et H. Prade: *Handling analogical proportions in classical logic and fuzzy logics settings*. Dans *Proc. 10th Eur. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'09)*, pages 638–650. Springer, LNCS 5590, 2009.

[18] Pirrelli, V. et F. Yvon: *Analogy in the lexicon : a probe into analogy-based machine learning of language*. Dans *Proc. 6th Int. Symp. on Human Communication*, Santiago de Cuba, 6 p.,1999.

[19] Prade, H. et G. Richard: *Cataloguing/analogizing : A nonmonotonic view*. Int. J. Intell. Syst., 26(12) :1176–1195, 2011.

[20] Prade, H. et G. Richard: *From Analogical Proportion to Logical Proportions*. Logica Universalis, 7(4) :441–505, 2013.

[21] Prade, H. et G. Richard: *Analogical proportions and analogical reasoning - An introduction*. Dans Aha, D. W. et J. Lieber (rédacteurs) : *Proc. 25th Int. Conf. on Case-Based Reasoning Research and Development (ICCBR'17), Trondheim, Norway, June 26-28*, tome 10339 de *LNAI*, pages 16–32. Springer, 2017.

[22] Prade, H. et G. Richard: *Analogical proportions : From equality to inequality*. Int. J. of Approximate Reasoning, 101 :234 – 254, 2018.

[23] Richter, M. M. et R. O. Weber: *Case-based reasoning, a textbook*. Springer, 2013.

[24] Riesbeck, C. K. et R. C. Schank: *Inside Case-Based Reasoning*. Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey, 1989. Available on line.

[25] Smyth, B. et M. T. Keane: *Using adaptation knowledge to retrieve and adapt design cases*. Knowledge-Based Systems, 9(2) :127–135, 1996.