

Adaptation par révision et adaptation différentielle : comparaison de deux approches de l'adaptation

Julien Cojan, Jean Lieber

► **To cite this version:**

Julien Cojan, Jean Lieber. Adaptation par révision et adaptation différentielle : comparaison de deux approches de l'adaptation. 19ème atelier Français de Raisonnement à Partir de Cas - Rapc2011, May 2011, Chambéry, France. 2011. <inria-00595400>

HAL Id: inria-00595400

<https://hal.inria.fr/inria-00595400>

Submitted on 24 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Adaptation par révision et adaptation différentielle : comparaison de deux approches de l'adaptation

Julien Cojan, Jean Lieber

UHP-Nancy 1 – LORIA (UMR 7503 CNRS-INPL-INRIA-Nancy 2-UHP)
BP 239, 54506 Vandœuvre-lès-Nancy, France Prénom.Nom@loria.fr

Résumé

L'objectif de cet article est la comparaison de deux approches de l'adaptation. L'adaptation par révision (APR) s'appuie sur un opérateur de révision des connaissances et consiste à appliquer cet opérateur pour réviser le cas source par le cas cible qui représente le problème à résoudre, en prenant en compte les connaissances du domaine. Les connaissances utiles à l'APR sont l'opérateur de révision et les connaissances du domaine. L'adaptation différentielle (AD) s'inspire du calcul différentiel en le généralisant à des espaces de représentation non nécessairement numériques. Les connaissances d'adaptation de l'AD sont les influences de chaque variation de descripteurs de problème sur les variations des descripteurs de solution. Il est montré que l'APR généralise l'AD au sens où on peut construire un opérateur de révision paramétré par les connaissances de l'AD telles que l'APR muni d'un tel opérateur et sans connaissances du domaine coïncide avec l'AD. Cela indique d'une part comment étendre l'AD pour qu'elle tienne compte de connaissances du domaine et, d'autre part, comment choisir un opérateur de révision pertinent pour l'APR en s'appuyant sur les principes de l'AD.

Mots clés : adaptation par révision, adaptation différentielle, adaptation conservatrice

1 Introduction

L'adaptation est une étape du raisonnement à partir de cas (RÀPC [17]) qui est à l'heure actuelle moins formalisée que, p. ex., la mémorisation, bien que des propositions aient été faites récemment. L'une de ces propositions est l'adaptation par révision qui cherche à réutiliser la théorie de la révision des connaissances pour l'adaptation. Une autre proposition est l'adaptation différentielle, qui généralise le calcul différentiel. L'objectif principal de cet article est d'étudier le lien entre ces deux théories.

En section 2 nous présentons les notions utilisées dans cet article. La section 3 présente l'adaptation par révision et l'adaptation différentielle. Ensuite, nous montrons dans la section 4 que l'adaptation différentielle est un type d'adaptation par révision, et cela entraîne des nouveaux résultats sur ces deux types d'adaptation. En section 5 nous comparons l'adaptation conservatrice (qui est un type de l'adaptation par révision) et l'adaptation différentielle dans une perspective d'acquisition des connaissances d'adaptation. Enfin, la section 6 résume les contributions de cet article et présente quelques perspectives.

2 Préliminaires

Soit \mathcal{U} un ensemble (éventuellement infini) appelé l'espace des cas. Une *instance de cas* u est, par définition, un élément de \mathcal{U} : $u \in \mathcal{U}$. Un *cas* \mathcal{C} est une classe d'instances de cas : $\mathcal{C} \in 2^{\mathcal{U}}$ (dans cet article, un cas représente une classe d'expériences, c'est ce que l'on appelle un cas ossifié dans [17] et un cas généralisé dans [16]). Un *cas source* est noté Source : il s'agit d'un cas de la base de cas. Le *cas cible* est noté Cible : c'est l'entrée du système de RÀPC. Dans beaucoup d'applications, les cas sources sont *spécifiques* : chacun d'entre eux représente une seule instance de cas u ($\text{Source} = \{u\}$). Au contraire, le cas cible ne spécifie que sa « partie

problème » et a besoin d'être complété par une « partie solution » : l'objectif de l'adaptation est d'assurer cette complétion :

$$\text{Adaptation} : (\text{Source}, \text{Cible}) \mapsto \text{CibleComplétée}$$

avec $\text{CibleComplétée} \subseteq \text{Cible}$

Dans beaucoup d'applications du RÀPC, une instance de cas u peut être décomposée en une partie problème x et une partie solution $y : u = (x, y)$. Soient \mathcal{U}_{pb} et \mathcal{U}_{sol} les espaces d'instances de problèmes et de solutions : $x \in \mathcal{U}_{\text{pb}}$, $y \in \mathcal{U}_{\text{sol}}$, $\mathcal{U} = \mathcal{U}_{\text{pb}} \times \mathcal{U}_{\text{sol}}$. Un cas source Source est décomposé en un *problème source* $\text{srce} \in 2^{\mathcal{U}_{\text{pb}}}$ et sa solution $\text{Sol}(\text{srce}) \in 2^{\mathcal{U}_{\text{sol}}}$, ainsi $\text{Source} = \text{srce} \times \text{Sol}(\text{srce})$ avec comme interprétation : pour tout $x \in \text{srce}$ il existe $y \in \text{Sol}(\text{srce})$ tel que $u = (x, y)$ est un cas licite (c.-à-d., y résout x). La partie solution du cas cible est inconnue, ainsi $\text{Cible} = \text{cible} \times \mathcal{U}_{\text{sol}}$, où cible est appelé le *problème cible*. Cette décomposition problème-solution est inutile pour présenter l'adaptation par révision mais elle est un pré-requis pour l'adaptation différentielle.

Les connaissances du domaine CD forment une base de connaissances qui donne la condition nécessaire pour qu'une instance de cas soit licite¹. Ainsi, les connaissances du domaine peuvent être représentées par un sous-ensemble CD de \mathcal{U} et pour chaque $u \in \mathcal{U}$, $u \notin \text{CD}$ implique que u n'est pas licite. Quand l'espace des cas est décomposé en $\mathcal{U}_{\text{pb}} \times \mathcal{U}_{\text{sol}}$, $u = (x, y) \notin \text{CD}$ veut dire que y n'est pas une solution de x ou que x et/ou y n'ont pas de sens (c.-à-d., que ce sont des objets représentés dans un langage qui n'a pas de correspondance dans le monde réel, p. ex., dans le domaine de la zoologie, un chat n'est pas un mammifère). Ne pas avoir de connaissances du domaine (ou ne pas les prendre en compte) équivaut à $\text{CD} = \mathcal{U}$.

Une distance dans l'ensemble \mathcal{U} est définie dans cet article comme une fonction $\text{dist} : \mathcal{U} \times \mathcal{U} \rightarrow [0; +\infty[$ telle que $\text{dist}(u^1, u^2) = 0$ ssi $u^1 = u^2$ (la propriété de symétrie et l'inégalité triangulaire ne sont pas nécessaires dans cet article). Si $\mathcal{U} = \mathbb{R}^n$ (où \mathbb{R} est l'ensemble de nombres réels), une distance L_1 est une distance dist définie dans une base \mathcal{B} de l'espace vectoriel \mathcal{U} par

$$\text{dist}(x, y) = \sum_{i=1}^n |b_i - a_i|$$

où (a_1, a_2, \dots, a_n) (resp., (b_1, b_2, \dots, b_n)) est la représentation de x (resp., de y) dans \mathcal{B} . Quand \mathcal{B} est une base canonique, $x_i = a_i$ et $y_i = b_i$ pour chaque $i \in \{1, 2, \dots, n\}$.

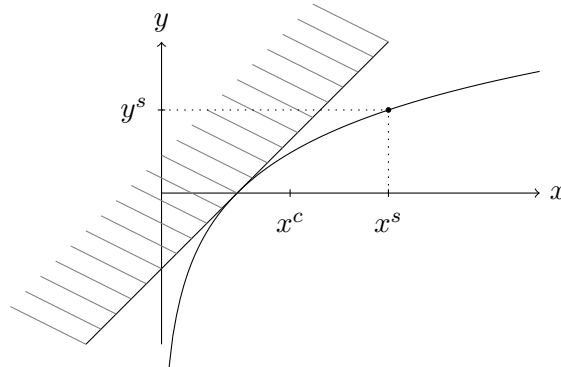


FIG. 1 – La fonction logarithme naturel est approchée ici par RÀPC. Le demi-plan sous la droite est CD (la partie hachurée est $\mathcal{U} \setminus \text{CD}$). $\text{Source} = \{(x^s, y^s)\}$ est un cas source spécifique. $\text{cible} = \{x^c\}$ est un problème cible spécifique.

¹En général, elles ne fournissent pas une condition *suffisante* : si c'était le cas, cela voudrait dire que les connaissances du domaine sont complètes, ce qui est une hypothèse rarement faite en RÀPC [14].

Présentation de l'exemple suivi au cours de cet article. Les deux approches de l'adaptation étudiées dans cet article ont été appliquées à des domaines avec des représentations symboliques complexes. Mais, pour des raisons de clarté, un exemple numérique simple est considéré : les tables de logarithmes. Plus précisément, $\mathcal{U}_{pb} =]0; +\infty[$, $\mathcal{U}_{sol} = \mathbb{R}$, $(x, y) \in \mathcal{U}$ est une instance de cas licite si $y \simeq \ln x$ (où \ln est un logarithme naturel, avec base e), et la base de cas est une table logarithmique : chaque cas source `Source` est spécifique : `Source` = $\{(x, y)\}$ où $y \simeq \ln x$. Les connaissances du domaine prises en compte sont $\ln x \leq x - 1$, ce qui veut dire que :

$$CD = \{(x, y) \in \mathcal{U} \mid y \leq x - 1\} \quad (1)$$

La figure 1 illustre ces notions.

3 Deux approches de l'adaptation

3.1 L'adaptation par révision

Dans un formalisme de représentation donné, un opérateur de révision \circ associe à deux bases de connaissances ψ et μ une base de connaissances $\psi \circ \mu$, la révision de ψ par μ . Intuitivement, $\psi \circ \mu$ est obtenue en changeant ψ au minimum pour obtenir ψ' telle que la conjonction de ψ' et μ , $\psi' \cup \mu$, soit cohérente. $\psi \circ \mu$ est le résultat de cette conjonction. La notion de changement minimal peut être modélisée de plusieurs façons, ainsi il y a plusieurs opérateurs. Cependant, quelques postulats ont été proposés pour cet opérateur, comme les postulats AGM [1]. Ces postulats ont été appliqués à la logique propositionnelle [13] et beaucoup étudiés dans ce formalisme (voir, p. ex., [11]). En particulier, étant donné une distance `dist` sur l'ensemble \mathcal{U} des interprétations, un opérateur \circ_{dist} unique (à l'équivalence près) peut être défini en établissant que l'ensemble des modèles de $\psi \circ_{\text{dist}} \mu$ est l'ensemble des modèles de μ à distance minimale de l'ensemble des modèles de ψ . Cette idée peut être généralisée en définissant un opérateur de révision \circ_{dist} sur les sous-ensembles d'un espace métrique $(\mathcal{U}, \text{dist})$:

$$C \circ_{\text{dist}} D = \{d \in D \mid \text{dist}(C, d) = \text{dist}(C, D)\}$$

avec $\text{dist}(C, d) = \inf_{c \in C} \text{dist}(c, d)$ et $\text{dist}(C, D) = \inf_{c \in C, d \in D} \text{dist}(c, d)$

Cet opérateur de révision est une généralisation de l'opérateur de révision défini en [11]. Le seul postulat requis pour `dist` est : $\text{dist}(c, d) = 0$ ssi $c = d$.

Pour un opérateur de révision \circ donné, la \circ -adaptation consiste simplement à utiliser l'opérateur de révision pour effectuer l'adaptation, tout en prenant en compte les connaissances du domaine :

$$\text{CibleComplétée} = (CD \cap \text{Source}) \circ (CD \cap \text{Cible}) \quad (2)$$

L'idée est de réutiliser au maximum les connaissances du cas source (c.-à-d. de le changer au minimum), pour prendre en compte les contraintes données par le cas cible. La première propriété attendue d'une adaptation est qu'elle soit cohérente avec l'information associée avec les connaissances du domaine et le cas cible :

$$\text{CibleComplétée} \subseteq CD \cap \text{Cible}$$

Cette propriété est entraînée par la \circ -adaptation en raison de la propriété de \circ suivante : $C \circ D \subseteq D$. Cette propriété est satisfaite par un opérateur vérifiant les postulats AGM et donc par un opérateur \circ_{dist} .

Dans [15], cette approche de l'adaptation a été étudiée dans le cadre de la logique propositionnelle. Dans [3, 4], la \circ -adaptation est étudiée dans un cadre formel similaire à celui de cet article. Dans [5], elle est étendue à la combinaison de cas (aussi connue comme adaptation de plusieurs cas). Dans [7, 8], une approche similaire de l'adaptation est décrite pour la logique de descriptions \mathcal{ALC} .

Application à l'exemple. Soit `dist` la distance sur $\mathcal{U} =]0; +\infty[\times \mathbb{R}$ définie pour $u^1, u^2 \in \mathcal{U}$ avec $u^i = (x^i, y^i)$ par $\text{dist}(u^1, u^2) = |x^2 - x^1| + |y^2 - y^1|$ (`dist` est la distance L_1 dans la base canonique). Considérons

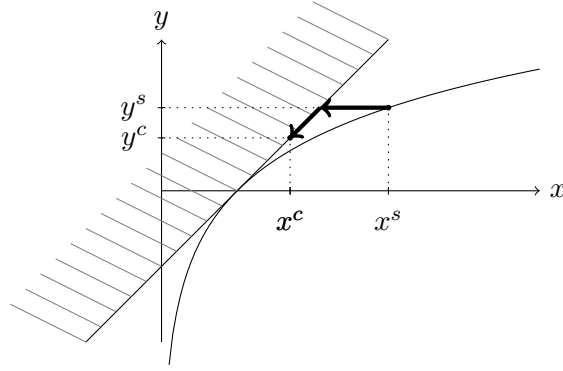


FIG. 2 – Illustration sur l'exemple de l'adaptation par révision suivant un opérateur défini à l'aide de la distance L_1 de la base canonique. Cette adaptation correspond à une adaptation par copie et correction : tant que la copie est cohérente avec les connaissances du domaine elle est appliquée (cf. la flèche horizontale à droite), quand elle ne l'est plus, elle est corrigée suivant les connaissances du domaine (cf. l'autre flèche).

l'adaptation d'un cas source $\text{Source} = \{u^s\}$ avec $u^s = (x^s, y^s)$ pour résoudre un problème cible $\text{cible} = \{x^c\}$. Le cas cible est $\text{Cible} = \text{cible} \times \mathbb{R}$.

On suppose d'abord qu'il n'y a pas de connaissances du domaine : $\text{CD} = \mathcal{U}$. CibleComplétée est l'ensemble des (x^c, y^c) qui minimisent $\text{dist}((x^s, y^s), (x^c, y^c))$. Puisque x^s, y^s , et x^c sont fixés, la minimisation de $\text{dist}((x^s, y^s), (x^c, y^c))$ consiste à minimiser $|y^c - y^s|$. Le minimum est atteint pour $y^c = y^s$, ainsi $\text{CibleComplétée} = \{(x^c, y^s)\}$. Par conséquent, avec cette distance et sans connaissances du domaine, la \circ_{dist} -adaptation est une adaptation par copie (aussi appelée « adaptation nulle » [17]) : la solution du cas source est simplement réutilisée par le problème cible).

On prend maintenant en compte les connaissances du domaine définies par l'équation (1), si le résultat de l'adaptation par copie est cohérent avec CD alors l'adaptation par copie est appliquée, sinon, elle fournit une solution à la frontière de CD :

$$y^c = \min(y^s, x^c - 1) \quad (3)$$

Cette adaptation est une « adaptation par copie et correction » : la solution du cas source est réutilisée et, quand elle est incohérente avec les connaissances du domaine et les contraintes du problème cible, elle est corrigée pour résoudre cette incohérence.

La figure 2 illustre cet exemple.

On montre plus loin qu'une \circ_{dist} -adaptation peut être plus sophistiquée et incorporer des connaissances d'adaptation spécifiques au domaine.

3.2 L'adaptation différentielle

L'adaptation différentielle s'inspire de l'équation de calcul différentiel suivante :

$$dy_j = \sum_i \frac{\partial y_j}{\partial x_i} \cdot dx_i \quad (4)$$

appliquée à une fonction différentiable $f : x \in \mathbb{R}^m \mapsto y \in \mathbb{R}^n$. Cette équation peut servir à calculer une approximation du premier ordre de f en interprétant les différentielles dx_i et dy_j comme des petites différences. Ce calcul peut être effectué grâce à la connaissance $\text{CA} = \left\{ \frac{\partial y_j}{\partial x_i} \right\}_{ij}$. CA est l'abréviation de *connaissances d'adaptation* : cette approximation du premier ordre peut être vue comme une adaptation. En effet, si $\mathcal{U}_{\text{pb}} = \mathbb{R}^m$, $\mathcal{U}_{\text{so1}} = \mathbb{R}^n$, $\text{Source} = \{u^s\}$ est un cas source spécifique avec $u^s = (x^s, y^s)$ et $y^s = f(x^s)$, alors (4) permet de calculer

$$\text{CibleComplétée} = \{(x^c, y^c)\} \quad \text{avec} \quad \{x^c\} = \text{cible} \quad \text{et} \quad y_j^c = y_j^s + \sum_i \frac{\partial y_j}{\partial x_i} \cdot (x_i^c - x_i^s)$$

L'adaptation différentielle est présentée en détail dans [12], nous proposons dans cet article une version simplifiée. Cette simplification consiste à généraliser le calcul différentiel en substituant \mathbb{R} par (éventuellement) d'autres ensembles de valeurs : $\mathcal{U}_{pb} = V_1 \times \dots \times V_m$, $\mathcal{U}_{so1} = W_1 \times \dots \times W_n$, où les V_i et les W_j sont différents types d'ensembles de valeurs. Un type \mathcal{T} peut être numérique (p. ex., $\mathcal{T} = \mathbb{R}$ or $\mathcal{T} = \mathbb{Z}$) ou non (p. ex., types énumérés, ensembles des parties d'un ensemble fini, etc.), sous condition de disposer d'une opération, notée $+$, telle que $(\mathcal{T}, +)$ soit un groupe commutatif². On considère donc $m + n$ groupes commutatifs $(V_i, +)$ et $(W_j, +)$ qui ne sont pas nécessairement différents. Soit $+$ définie sur \mathcal{U}_{pb} (resp., sur \mathcal{U}_{so1}) par $a + b = (a_1 + b_1, \dots, a_m + b_m)$ (resp., $a + b = (a_1 + b_1, \dots, a_n + b_n)$). $(\mathcal{U}_{pb}, +)$ et $(\mathcal{U}_{so1}, +)$ sont des groupes commutatifs. Étant donné un cas spécifique $\text{Source} = \{u\}$ avec $u = (x^s, y^s)$ et un problème cible spécifique $\text{cible} = \{x^c\}$, dx et dy dénotant respectivement $x^c - x^s$ et $y^c - y^s$.

L'adaptation différentielle consiste à ① calculer dy à partir de dx et ② calculer $y^c = y^s + dy$. L'étape ② est simple. Soit D_{x^s} la fonction qui calcule l'étape ① : $D_{x^s} : dx \mapsto dy$. Dans le cadre du calcul différentiel numérique, D_{x^s} est une transformation linéaire de matrice $\left[\frac{\partial y_j}{\partial x_i}(x^s) \right]_{ij}$. Dans le cadre de l'adaptation différentielle générale, D_{x^s} correspond aux fonctions $\frac{\partial y_j}{\partial x_i}(x^s)$ qui associent à chaque $dx_i \in V_i$ une contribution $d_i y_j \in W_j$ à dy_j :

$$\begin{aligned} \frac{\partial y_j}{\partial x_i}(x^s) : dx_i &\mapsto d_i y_j & dy_j &= \sum_i d_i y_j \\ D_{x^s}((dx_1, \dots, dx_m)) &= (dy_1, \dots, dy_n) \\ y^c &= y^s + D_{x^s}(x^c - x^s) \end{aligned} \quad (5)$$

On suppose que $\left(\frac{\partial y_j}{\partial x_i}(x^s) \right) (0) = 0$: si le descripteur du problème x_i ne change pas ($dx_i = 0$) alors il ne contribue pas au changement du descripteur de la solution dy_j ($d_i y_j = 0$). Cela entraîne que $D_{x^s}(0) = 0$.

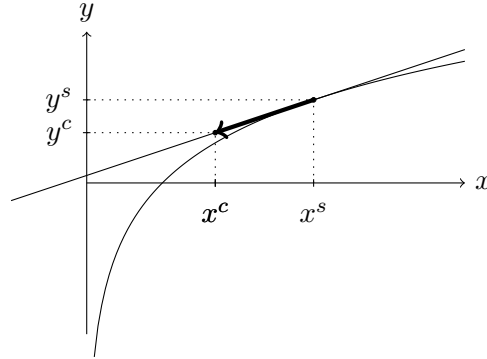


FIG. 3 – Illustration de l'adaptation différentielle sur l'exemple. Elle consiste simplement à suivre la droite de pente ϱ^s (le choix optimal de ϱ^s est choisi ici : la droite est la tangente à la courbe en x^c). Aucune connaissance du domaine n'est considérée ($CD = \mathcal{U}$).

Application à l'exemple. Pour cet exemple, $D_{x^s} : dx \in \mathbb{R} \mapsto \varrho^s \cdot dx \in \mathbb{R}$, avec $\varrho^s \in \mathbb{R}$. Ainsi, l'adaptation est simplement définie par

$$y^c = y^s + \varrho^s \cdot (x^c - x^s) \quad (6)$$

(l'adaptation se fait en suivant une droite et ϱ^s est la pente de cette droite). Le choix « optimal » de ϱ^s est $\varrho^s = \frac{d \ln x^s}{dx^s} = \frac{1}{x^s}$ (³). Le choix $\varrho^s = 0$ correspond à l'adaptation par copie : $y^c = y^s$.

²Un groupe commutatif $(\mathcal{T}, +)$ est tel que $+$ est une opération interne associative et commutative, $(\mathcal{T}, +)$ a un élément neutre, noté 0, et chaque $a \in \mathcal{T}$ a un élément inverse noté $-a$ ($a + (-a) = 0$). $b - a$ est une notation pour $b + (-a)$. Dans [12] les hypothèses sur $(\mathcal{T}, +)$ sont plus faibles.

³C'est le choix optimal dans le sens où $y^c - (y^s + \varrho^s \cdot (x^c - x^s)) =_{x^c \rightarrow x^s} o(x^c - x^s)$ ssi $\varrho^s = \frac{1}{x^s}$, ce qui est une conséquence directe de la définition des dérivées.

La figure 3 illustre cet exemple.

4 L'adaptation différentielle vue comme une adaptation par révision

La proposition suivante montre que l'adaptation différentielle peut être simulée par l'adaptation par révision :

Proposition 1 *Étant donné les connaissances d'adaptation différentielles exprimées par la fonction D_{x^s} paramétrée par $x^s \in \mathcal{U}_{pb}$, il existe un opérateur de révision \circ tel que le résultat de l'adaptation différentielle coïncide avec la \circ -adaptation sans connaissances du domaine.*

Plus précisément, soient $dist_{pb}$ et $dist_{sol}$ respectivement des distances sur \mathcal{U}_{pb} et \mathcal{U}_{sol} , et soit $dist$ la distance telle que pour $u^s, u^c \in \mathcal{U}$, avec $u^s = (x^s, y^s)$ et $u^c = (x^c, y^c)$,

$$dist(u^s, u^c) = dist_{pb}(x^s, x^c) + dist_{sol}(y^s, y^c - D_{x^s}(x^c - x^s)) \quad (7)$$

La \circ_{dist} -adaptation de Source = $\{(x^s, y^s)\}$ pour résoudre Cible = $\{x^c\} \times \mathcal{U}_{sol}$ donne donc $y^c = y^s + D_{x^s}(x^c - x^s)$ qui est le résultat de l'adaptation différentielle.

Preuve. Montrons d'abord que $dist$ est une distance. Si $dist(u^s, u^c) = 0$ alors (a) $dist_{pb}(x^s, x^c) = 0$ et (b) $dist_{sol}(y^s, y^c - D_{x^s}(x^c - x^s)) = 0$. Puisque $dist_{pb}$ est une distance, (a) entraîne que $x^s = x^c$. Donc $D_{x^s}(x^c - x^s) = 0$, et (b) entraîne que $dist_{sol}(y^s, y^c - 0) = 0$. Puisque $dist_{sol}$ est une distance, $y^s = y^c$. Finalement, $u^s = u^c$ et $dist$ est une distance (suivant la définition de la section 2).

Soit y^c le résultat de la \circ_{dist} -adaptation de $\{(x^s, y^s)\}$ par l'instance du problème cible x^c . y^c minimise donc $\Delta = dist((x^s, y^s), (x^c, y^c))$, où x^s, y^s et x^c sont fixés. (7) entraîne que $\Delta \geq dist_{pb}(x^s, x^c)$. $\Delta = dist_{pb}(x^s, x^c)$ n'est atteint que si $dist_{sol}(y^s, y^c - D_{x^s}(x^c - x^s)) = 0$, ce qui équivaut, puisque $dist_{sol}$ est une distance, à $y^s = y^c - D_{x^s}(x^c - x^s)$, soit $y^c = y^s + D_{x^s}(x^c - x^s)$. Donc, si y^c est le résultat de la \circ_{dist} -adaptation avec $dist$ définie par (7) alors y^c est aussi le résultat de l'adaptation différentielle. La preuve de l'implication inverse est directe. \square

Ce résultat enrichit mutuellement les deux approches de l'adaptation :

- L'adaptation par révision généralise l'adaptation différentielle dans le sens où elle permet de considérer des cas sources qui ne sont pas nécessairement spécifiques et, surtout, elle permet de prendre en compte les connaissances du domaine pour l'adaptation différentielle.
- L'adaptation par révision dépend du choix d'un opérateur de révision et les connaissances d'adaptation associées à l'adaptation différentielle peuvent servir à définir un opérateur de révision pertinent pour effectuer l'adaptation.

L'adaptation conservatrice revue et corrigée. Dans [15] et [3], l'adaptation par révision est appelée *conservative adaptation* (adaptation conservatrice). Ce qualificatif semble être mal approprié pour certaines adaptations par révision, en particulier pour celles qui simulent l'adaptation différentielle (qui peuvent difficilement être qualifiées de conservatrices, sauf si $D_{x^s}(dx) = 0$ pour tout $dx \in \mathcal{U}_{pb}$). Quand la \circ -adaptation est-elle conservatrice ? Pour répondre à cette question, on considère la distance suivante sur \mathcal{U} (avec les hypothèses de la proposition 1) :

$$dist(u^s, u^c) = dist_{pb}(x^s, x^c) + dist_{sol}(y^s, y^c) \quad (8)$$

Lorsque la \circ_{dist} -adaptation est appliquée, on peut remarquer que le résultat de la minimisation de $y^c \mapsto dist(u^s, u^c)$ dépend seulement de y^s , elle ne dépend ni de x^s , ni de y^c . La \circ_{dist} -adaptation avec cette distance est conservatrice dans le sens où elle consiste à copier et corriger la solution du cas source. L'adaptation de (3) illustre cela. Contrairement à la distance définie par (8), la distance définie par (7) établit une corrélation entre les instances de problème et les instances de solution : le résultat de la minimisation de la distance dépend de x^s, y^s , et x^c . Cette remarque justifie la proposition suivante pour la définition d'adaptation conservatrice :

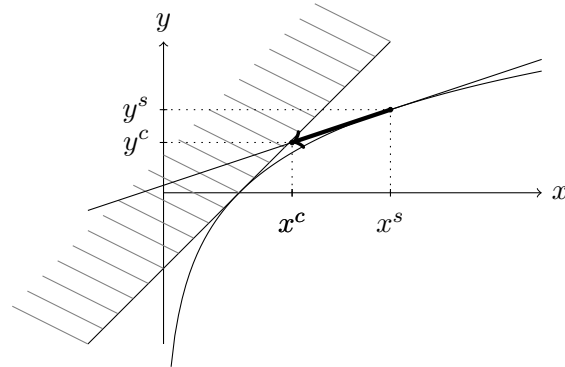
Définition 1 *Une adaptation est dite conservatrice s'il s'agit d'une \circ -adaptation où \circ est un opérateur de révision sur \mathcal{U} tel qu'il existe un opérateur de révision \circ_{pb} sur \mathcal{U}_{pb} et un opérateur de révision \circ_{sol} sur \mathcal{U}_{sol}*

vérifiant la propriété suivante :

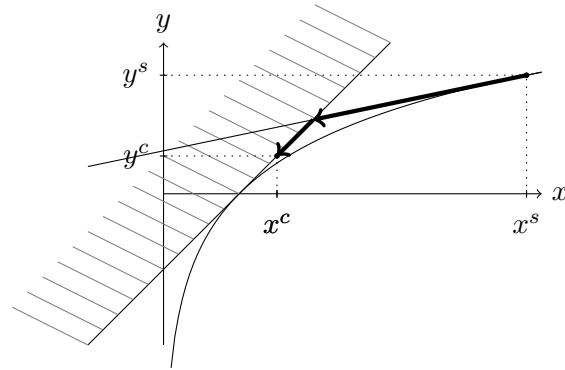
$$\text{si } P_1, P_2 \in 2^{\mathcal{U}_{pb}}, S_1, S_2 \in 2^{\mathcal{U}_{sol}}, \text{ alors } (P_1 \times S_1) \circ (P_2 \times S_2) = (P_1 \circ_{pb} P_2) \times (S_1 \circ_{sol} S_2) \quad (9)$$

En particulier, la \circ_{dist} -adaptation, avec la distance $dist$ définie par (8), est conservatrice puisque la propriété ci-dessus, reliant \circ , \circ_{pb} et \circ_{sol} est vérifié pour $\circ = \circ_{dist}$, $\circ_{pb} = \circ_{dist_{pb}}$, et $\circ_{sol} = \circ_{dist_{sol}}$.

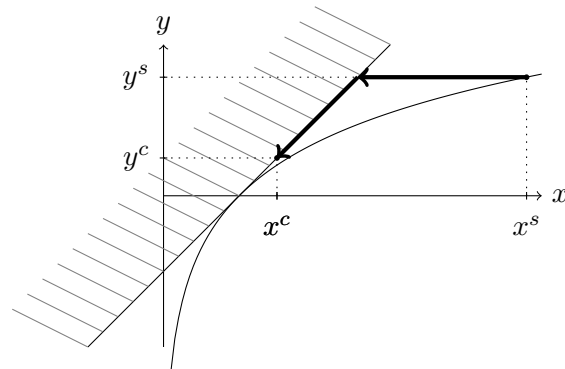
Lorsqu'un opérateur de révision \circ sur \mathcal{U} vérifie (9), on parle d'opérateur de révision *orthogonal* (par rapport à la décomposition problème-solution de \mathcal{U}). Cet adjectif est justifié sur l'exemple ci-dessous.



(a)



(b)



(c)

FIG. 4 – Deux illustrations de l'adaptation par révision suivant un opérateur de révision intégrant des connaissances d'adaptation différentielle. Dans la situation (a), le résultat de l'adaptation différentielle est cohérent avec les connaissances du domaine, il n'y a donc pas besoin de le corriger (cela est équivalent à l'adaptation de la figure 3). Dans la situation (b), le résultat de l'adaptation différentielle n'est pas cohérent avec les connaissances du domaine, il faut donc le corriger. Le résultat de cette adaptation est équivalent à l'adaptation conservatrice représentée en (c).

Application à l'exemple. Avec des distances $\text{dist}_{\mathcal{U}_{\text{pb}}}$ et $\text{dist}_{\mathcal{U}_{\text{so1}}}$ de la forme $|\cdot - \cdot|$ (c.-à-d. les distances L_1 sur les bases canoniques de \mathcal{U}_{pb} et \mathcal{U}_{so1}), l'équation (7) donne

$$\text{dist}(u^s, u^c) = |x^c - x^s| + |y^c - (y^s + \varrho^s \cdot (x^c - x^s))| \quad (10)$$

Donc, sans connaissances du domaine, la \circ_{dist} -adaptation donne le résultat (6) et avec les connaissances du domaine spécifiées dans (1), cette adaptation donne :

$$y^c = \min(y^s + \varrho^s \cdot (x^c - x^s), x^c - 1)$$

La situation $\varrho^s = 0$ correspond à l'adaptation conservatrice (cf. équation (3)). Plus généralement, la distance dist définie par (10) est une distance L_1 dans la base (\vec{e}_1, \vec{e}_2) de \mathbb{R}^2 avec $\vec{e}_1 = \vec{\varepsilon}_1 + \varrho^s \vec{\varepsilon}_2$ et $\vec{e}_2 = \vec{\varepsilon}_2$ où $(\vec{\varepsilon}_1, \vec{\varepsilon}_2)$ est la base canonique ($\vec{\varepsilon}_1 = (1, 0)$ et $\vec{\varepsilon}_2 = (0, 1)$) : la pente de \vec{e}_1 dans $(\vec{\varepsilon}_1, \vec{\varepsilon}_2)$ vaut ϱ^s . (\vec{e}_1, \vec{e}_2) est une base *orthogonale* (par rapport à $(\vec{\varepsilon}_1, \vec{\varepsilon}_2)$) ssi $\varrho^s = 0$ (ce qui correspond à $(\vec{e}_1, \vec{e}_2) = (\vec{\varepsilon}_1, \vec{\varepsilon}_2)$). Ainsi, \circ_{dist} avec la distance dist définie par (10) est orthogonale (avec le sens de (9)) ssi (\vec{e}_1, \vec{e}_2) , la base dans laquelle dist est L_1 , est orthogonale. Cela justifie *a posteriori* l'adjectif « orthogonal » de l'opérateur de révision.

La figure 4 illustre cette adaptation.

Remarque 1 Dans cet exemple, cette adaptation coïncide soit avec l'adaptation différentielle soit avec l'adaptation conservatrice. Ce n'est pas toujours le cas. Par exemple, supposons qu'on cherche à approcher par un système de RÀPC la valeur de la fonction sinus en $\pi/2$ en connaissant sa valeur en 0. $\mathcal{U} = \mathcal{U}_{\text{pb}} \times \mathcal{U}_{\text{so1}}$ avec $\mathcal{U}_{\text{pb}} = \mathcal{U}_{\text{so1}} = \mathbb{R}$, $\text{Source} = \{(0, 0)\}$ et le problème cible à résoudre est $\text{cible} = \pi/2$, on s'appuie sur la valeur de la dérivée du sinus en 0 qui vaut 1 et sur les connaissances du domaine qui établissent que la fonction sinus est à valeurs dans $[-1, 1]$:

$$\text{CD} = \{(x, y) \in \mathcal{U} \mid -1 \leq y \leq 1\}$$

L'adaptation de Source pour résoudre cible donne $\text{CibleComplétée} = \{(x^c, y^c)\}$ avec :

$$x^c = \pi/2 \quad \text{et} \quad \begin{cases} y^c = 0 & \text{avec l'adaptation conservatrice.} \\ y^c = \pi/2 & \text{avec l'adaptation différentielle (sans tenir compte de CD).} \\ y^c = 1 & \text{avec la } \circ_{\text{dist}}\text{-adaptation pour } \text{dist} \text{ définie suivant (10).} \end{cases}$$

5 De l'adaptation conservatrice à l'adaptation différentielle optimale

Cette section traite d'adaptations qui sont « entre » l'adaptation conservatrice (AC) et l'adaptation différentielle optimale (ADO), étant donné des connaissances du domaine. L'ADO est une adaptation différentielle avec des connaissances complètes sur les $\left[\frac{\partial y_j}{\partial x_i}(x^s)\right]_{i,j}$, c.-à-d., un choix de D_{x^s} donnant les résultats optimaux pour l'adaptation différentielle suivant des critères prédéfinis (pour l'exemple, ce choix optimal est expliqué dans la note 3 en bas de page 5). L'adaptation conservatrice correspond, elle, au choix arbitraire $\frac{\partial y_j}{\partial x_i}(x^s) = 0$ (pour chaque x^s, i et j).

D'un côté, l'ADO peut être préférée à l'AC, puisqu'elle donne la meilleure adaptation parmi les adaptations différentielles. D'un autre côté, l'AC ne demande par d'effort d'acquisition de connaissances d'adaptation ni d'apprentissage alors que l'ADO en demande beaucoup. L'idée d'adaptations intermédiaires entre AC et ADO a donc émergé, elle consiste à définir une procédure d'acquisition de connaissances d'adaptation, pour affiner graduellement l'opérateur d'adaptation en partant de l'AC comme état initial, et en convergeant vers l'ADO. Cette procédure pourrait être implantée suivant les principes de l'acquisition opportuniste de connaissances d'adaptation (voir p. ex. [9]) ou de l'extraction opportuniste de connaissances d'adaptation (voir p. ex. [2]).

6 Conclusion et perspectives

Cet article compare l'adaptation par révision (APR) et l'adaptation différentielle (AD), deux approches complémentaires de l'adaptation qui permettent d'aborder le problème de conception d'un module d'adaptation sous des angles différents. Ces approches d'adaptation sont toutes deux définies au niveau des inférences

(c.-à-d., indépendamment des algorithmes réalisant ces inférences). À ce niveau, il est montré que l'AD est plus spécifique que l'APR : un opérateur de révision prenant en compte les connaissances d'adaptation de l'adaptation différentielle peut être défini de sorte, qu'en l'absence de connaissances du domaine, l'APR utilisant cet opérateur de révision coïncide avec l'AD. Ce résultat implique que ces deux approches de l'adaptation peuvent bénéficier l'une de l'autre :

- L'AD peut être étendue grâce à l'APR pour prendre en compte des connaissances du domaine, elle peut aussi bénéficier des travaux sur la révision des connaissances.
- L'approche de l'APR ne précise pas comment choisir un opérateur de révision pertinent pour effectuer l'adaptation, les connaissances d'AD peuvent guider le choix d'un « bon » opérateur de révision pour l'APR (défini suivant la topologie induite par la connaissance des $\left[\frac{\partial y_j}{\partial x_i}(x^s) \right]_{ij}$).

Une autre contribution de cet article est le retour sur la notion d'adaptation conservatrice. Dans les travaux précédents, les termes adaptation conservatrice et APR étaient utilisés comme synonymes. Dans cet article, l'adaptation conservatrice est redéfinie comme une APR pour laquelle l'opérateur de révision a la propriété de pouvoir être décomposé orthogonalement suivant \mathcal{U}_{pb} et \mathcal{U}_{so1} , ce qui correspond à une adaptation par copie et correction. Inversement, une adaptation non conservatrice correspond à un opérateur de révision « non orthogonal » : une variation dans l'espace des problèmes implique une variation dans l'espace des solutions.

La dernière contribution de cet article s'appuie sur une comparaison de deux approches de l'adaptation extrêmes : l'adaptation conservatrice, qui ne requiert aucune acquisition de connaissances d'adaptation et constitue une adaptation novice, et l'adaptation différentielle optimale, très exigeante en termes d'acquisition de connaissances d'adaptation. Une procédure d'acquisition de connaissances d'adaptation est envisagée, qui part de l'adaptation conservatrice et qui devrait converger vers l'adaptation différentielle optimale. Cette procédure pourrait, par exemple, s'appuyer sur les principes de l'acquisition opportuniste de connaissances. Une première perspective serait de concevoir une procédure d'acquisition suivant ce modèle et d'étudier expérimentalement l'évolution des performances d'un système de RÀPC suivant cette acquisition.

Une perspective plus générale serait de comparer l'adaptation par révision et d'autres approches de l'adaptation. Une première comparaison avec les approches de l'adaptation par règles a été effectuée dans [6]. D'autres comparaisons peuvent être faites, par exemple, avec l'adaptation à partir de cas (voir, p. ex., [10]).

Remerciements. Les auteurs remercient les relecteurs dont les remarques ont permis d'améliorer cet article.

Références

- [1] C. E. Alchourrón, P. Gärdenfors, et D. Makinson. On the Logic of Theory Change : partial meet functions for contraction and revision. *Journal of Symbolic Logic*, 50 :510–530, 1985.
- [2] F. Badra, A. Cordier, et J. Lieber. Opportunistic Adaptation Knowledge Discovery. In *Case-Based Reasoning Research and Development (ICCBR 2009)*, 2009. 60–74.
- [3] J. Cojan et J. Lieber. Conservative Adaptation in Metric Spaces. In *Proceedings of the 9th Conference on Case-Based Reasoning (ECCBR-08)*, Lecture Notes in Artificial Intelligence 5239, pages 135–149. Springer, 2008.
- [4] J. Cojan and J. Lieber. Formalisation de l'adaptation conservatrice dans les espaces métriques. In *16^{ème} atelier RàPC*, pages 92–107, 2008.
- [5] J. Cojan et J. Lieber. Belief Merging-based Case Combination. In *Case-Based Reasoning Research and Development (ICCBR 2009)*, pages 105–119, 2009.
- [6] J. Cojan et J. Lieber. Adapter des cas en utilisant un opérateur de révision ou des règles. In Sébastien Konieczny Laurence Cholvy, editor, *Journée Intelligence Artificielle Fondamentale*, Strasbourg France, 06 2010. Laurence Cholvy, Sébastien Konieczny.
- [7] J. Cojan et J. Lieber. An Algorithm for Adapting Cases Represented in an Expressive Description Logic. In Springer, editor, *Proceedings of the 18th International Conference on Case-Based Reasoning (ICCBR-2010)*, volume 6176 of *Lecture Notes in Artificial Intelligence*, pages 51–65, 2010.

- [8] J. Cojan and J. Lieber. Un algorithme d'adaptation avec des cas exprimés dans la logique de descriptions *ALC*. In *18^{ème} atelier RàPC*, 2010.
- [9] A. Cordier, B. Fuchs, L. Lana de Carvalho, J. Lieber, et A. Mille. Opportunistic Acquisition of Adaptation Knowledge and Cases – The IakA Approach. In *Advances in Case-Based Reasoning, 9th European Conference, ECCBR-2008, Trier, Germany. Proceedings*, Lecture Notes in Artificial Intelligence 5239, pages 150–164. Springer, 2008.
- [10] S. Craw, J. Jarmulak, et R. Rowe. Learning and Applying Case-Based Adaptation Knowledge. In D. W. Aha et I. Watson, editors, *Case-Based Reasoning Research and Development — Fourth International Conference on Case-Based Reasoning (ICCBR-01)*, Lecture Notes in Artificial Intelligence 2080, 2001.
- [11] M. Dalal. Investigations into a theory of knowledge base revision : Preliminary report. In *AAAI*, pages 475–479, 1988.
- [12] B. Fuchs, J. Lieber, A. Mille, et A. Napoli. A general strategy for adaptation in Case-Based Reasoning. Technical Report RR-LIRIS-2006-016, LIRIS UMR 5205 CNRS/INSA de Lyon/Université Claude Bernard Lyon 1/Université Lumière Lyon 2/École Centrale de Lyon, 2006.
- [13] H. Katsuno et A. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52(3) :263–294, 1991.
- [14] J. Kolodner. *Case-Based Reasoning*. Morgan Kaufmann, Inc., 1993.
- [15] J. Lieber. Application of the Revision Theory to Adaptation in Case-Based Reasoning : the Conservative Adaptation. In *Proceedings of the 7th International Conference on Case-Based Reasoning (ICCBR-07)*, Lecture Notes in Artificial Intelligence 4626, pages 239–253. Springer, Belfast, 2007.
- [16] K. Maximini, R. Maximini, et R. Bergmann. An investigation of generalized cases. In K. D. Ashley et D. Bridge, editors, *Proceedings of the 5th International Conference on Case Base Reasoning (ICCBR'03)*, volume 2689 of *LNAI*, pages 261–275, Trondheim, Norway, June 2003. Springer.
- [17] C. K. Riesbeck et R. C. Schank. *Inside Case-Based Reasoning*. Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey, 1989.