Une approche pour représenter les variations entre cas — Vers une application à l'extraction de connaissances d'adaptation

Fadi Badra, Jean Lieber

▶ **To cite this version:**

Fadi Badra, Jean Lieber. Une approche pour représenter les variations entre cas — Vers une application à l'extraction de connaissances d'adaptation. 15ème atelier sur le raisonnement à partir de cas - RàPC-07, Amélie Cordier, Jul 2007, Grenoble, France. pp.47–56. inria-00189611

HAL Id: inria-00189611
https://inria.hal.science/inria-00189611

Submitted on 21 Nov 2007

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.

Une approche pour représenter les variations entre cas — Vers une application à l'extraction de connaissances d'adaptation

Fadi Badra et Jean Lieber
LORIA (UMR 7503 CNRS–INPL–INRIA-Nancy 2–UHP),
BP 239, 54 506 Vandœuvre-lès-Nancy, FRANCE
{badra,lieber}@loria.fr

Résumé

Cet article propose une méthode pour représenter les variations entre deux cas dans le cadre d'une représentation des cas par des couples attributs-ensembles de valeurs (ce qui généralise le formalisme, classique dans le cadre du raisonnement à partir de cas, des couples attributs-valeurs). Pour effectuer une telle représentation, on met en évidence tout d'abord les attributs qui ne sont pas communs aux deux cas puis, pour les autres, on compare les ensembles de valeurs par des ensembles de relations binaires. De telles relations sont spécifiées par le concepteur du système de raisonnement à partir de cas sur la base de ce qui est pertinent pour « passer » d'un cas à un autre. Des exemples de telles relations sont proposés. Ce travail doit être appliqué à l'extraction de connaissances d'adaptation par fouille de la base de cas : une telle fouille s'appuie sur la représentation des variations entre cas de la base. L'expressivité des règles d'adaptation ainsi extraites sera fonction de l'expressivité de la représentation des variations entre cas.

1 Introduction

Le raisonnement à partir de cas (RàPC [13]) est un mode de résolution de problème s'appuyant sur une base de cas, un cas étant un problème déjà résolu accompagné de sa solution. On note `cible` le problème à résoudre et (`srce`,`Sol(srce)`) un cas de la base de cas (`srce` : problème *source*, `Sol(srce)` : solution de `srce`). Une session de RàPC consiste en général à sélectionner un cas (`srce`, `Sol(srce)`) de la base jugé similaire à `cible` (étape de remémoration) puis à adapter ce cas dans l'optique de la résolution de `cible` (étape d'adaptation). L'adaptation, selon le principe de l'analogie par transformation [5], peut être décomposée en trois étapes. La première consiste à analyser les ressemblances et dissemblances entre `srce` et `cible`, ce que nous appellerons les *variations* entre ces problèmes et dénoterons par Δ`pb`. La deuxième étape consiste à inférer les variations Δ`sol` entre solutions à partir des Δ`pb` et des connaissances d'adaptation. La troisième étape consiste à appliquer Δ`sol` sur `Sol(srce)` afin de construire une solution `Sol(cible)` de `cible`. Cela motive l'étude des variations entre cas (i.e. des variations entre problèmes et des variations entre solutions), en particulier l'étude de leur formalisation.

Cet article présente une approche pour représenter les variations entre cas. Plus précisément, nous considérerons que les problèmes (resp. les solutions) sont représentés par des ensembles de descripteurs, où un descripteur est un couple (`a`,`V`), `a` étant un attribut et `V` une contrainte sur `a` : étant donné un problème `pb` et (`a`,`V`) $\in$ `pb`, `V` est l'ensemble des valeurs que peut prendre `a` dans le contexte du problème `pb`. Cette recherche est motivée initialement par le développement de l'outil CabamakA d'extraction de connaissances d'adaptation par fouille de la base de cas [6] : à l'heure actuelle, cet outil s'appuie sur une représentation très fruste des variations entre cas et doit être amélioré à terme grâce à une représentation plus sophistiquée de ces variations.

La section suivante introduit les notions, notations et hypothèses relatives au langage choisi pour représenter les cas. Dans la section 3, nous proposons une approche pour représenter les variations

entre deux cas représentés dans ce langage. Quelques exemples de représentations de ces variations sont donnés dans la section 4. L'approche est ensuite appliquée au système CabamakA (section 5). La section 6 discute cette approche et la situe par rapport à des travaux proches.

2 Notions, notations et hypothèses

Langage de description des cas. Les problèmes et les solutions sont décrits par des ensembles finis de descripteurs. Un tel ensemble de descripteurs est appelé *concept* dans cet article. Par exemple :

$$\mathtt{C} = \{\mathtt{d}_1, \mathtt{d}_2, \ldots \mathtt{d}_n\} \tag{1}$$

est un concept. On suppose qu'un descripteur prend la forme d'un couple (a,V), où a désigne un attribut monovalué et V la représentation d'un ensemble de valeurs[1]. L'ensemble de toutes les valeurs possibles associées à a est appelé *co-domaine* de a et est dénoté par $\mathcal{D}_\mathtt{a}$. Par conséquent, si (a,V) est un descripteur, alors $\mathtt{V} \subseteq \mathcal{D}_\mathtt{a}$. L'ensemble des ensembles de valeurs V autorisés pour l'attribut a est noté $\mathcal{E}_\mathtt{a}$ et est un sous-ensemble de l'ensemble $2^{\mathcal{D}_\mathtt{a}}$ des parties de $\mathcal{D}_\mathtt{a}$. On suppose que si $\mathtt{V}_1 \in \mathcal{E}_\mathtt{a}$, $\mathtt{V}_2 \in \mathcal{E}_\mathtt{a}$ et $\mathtt{V}_1 \cap \mathtt{V}_2 \neq \emptyset$, alors $\mathtt{V}_1 \cap \mathtt{V}_2 \in \mathcal{E}_\mathtt{a}$. Par exemple, si on considère les attributs `diamètre` et `luit`, dont on restreint les ensembles de valeur respectifs aux intervalles fermés de $\mathbb{R}^+ = [0; +\infty[$ et aux sous-ensembles de {`oui`, `non`}, on peut définir le concept

$$\mathtt{Planète} = \{(\mathtt{diamètre}, [0,8; 140]), (\mathtt{luit}, \{\mathtt{non}\})\}$$

qui décrit l'ensemble des astres dont le diamètre est situé entre 0,8 et 140 milliers de kilomètres et qui ne brillent pas par eux-mêmes. $\mathcal{E}_\mathtt{diamètre}$ est donc l'ensemble des (représentations d')intervalles fermés de $\mathbb{R}^+$ et $\mathcal{E}_\mathtt{luit} = 2^{\{\mathtt{oui},\mathtt{non}\}}$.

Le cadre de représentation proposé est assez général et couvre en particulier le formalisme, classique en RàPC, des descripteurs de la forme (a,v) où v est une valeur (et non un ensemble de valeurs). En effet, il suffit de remplacer (a,v) par (a,{v}) pour se ramener à notre formalisme.

Sémantique. Une interprétation $\mathcal{I}$ est un couple $(\Delta^\mathcal{I}, \cdot^\mathcal{I})$, où $\Delta^\mathcal{I}$ est un ensemble (le *domaine d'interprétation*) et $\cdot^\mathcal{I}$ est une *fonction d'interprétation* qui à un concept C associe un ensemble $\mathtt{C}^\mathcal{I} \subseteq \Delta^\mathcal{I}$ et à un attribut a associe une fonction partielle $\Delta^\mathcal{I} \to \mathcal{D}_\mathtt{a}$.

L'ensemble de descripteurs décrivant un concept s'interprète comme une conjonction, de sorte qu'une définition de concept pourra être notée indifféremment sous la forme (1) ou sous la forme :

$$\mathtt{C} = \mathtt{d}_1 \wedge \mathtt{d}_2 \wedge \ldots \wedge \mathtt{d}_n$$

La conjonction de descripteurs est alors interprétée par une intersection :

$$(\mathtt{d}_k \wedge \mathtt{d}_\ell)^\mathcal{I} = \mathtt{d}_k^\mathcal{I} \cap \mathtt{d}_\ell^\mathcal{I}$$

Ainsi :

$$\mathtt{C}^\mathcal{I} = \bigcap_k \mathtt{d}_k^\mathcal{I}$$

L'interprétation d'un descripteur $\mathtt{d} = (\mathtt{a}, \mathtt{V})$ est[2]

$$(\mathtt{a}, \mathtt{V})^\mathcal{I} = \{x \in \Delta^\mathcal{I} \mid \mathtt{a}^\mathcal{I}(x) \text{ est défini et } \mathtt{a}^\mathcal{I}(x) \in \mathtt{V}^\mathcal{I}\}$$

Dans cette étude, on supposera que tous les concepts sont définis par au moins un attribut. On suppose en outre que dans une même définition de concept un nom d'attribut n'apparaît qu'une seule fois[3].

[1] Notons que dans cet article, nous appellerons souvent « ensemble de valeurs » ce qu'il faudrait appeler plus rigoureusement « représentation d'un ensemble de valeurs », afin de ne pas trop alourdir l'exposé.

[2] Pour les lecteurs familiers avec les logiques de descriptions [2], (a,V) correspond dans ces formalismes à $\exists$a.V, où soit a est un rôle fonctionnel et V est un concept, soit a est un attribut concret et V est un prédicat unaire de domaine concret.

[3] Cela ne réduit pas l'expressivité du langage : cela consiste simplement à remplacer $(\mathtt{a},\mathtt{V}_1) \wedge (\mathtt{a},\mathtt{V}_2)$ par $(\mathtt{a},\mathtt{V}_1 \cap \mathtt{V}_2)$, ces deux expressions ayant la même sémantique.

3 Approche proposée

Le but de cet article est de proposer une méthode pour comparer deux cas. Cela revient à comparer deux problèmes et deux solutions. Pour cela il suffit de savoir comparer deux concepts. Un concept représente un ensemble d'individus.

Dans la suite de la section, nous considérerons deux problèmes `srce` et `cible` comme concepts à comparer, et nous représenterons la variation $\Delta\mathtt{pb}$ lorsqu'on passe du problème `srce` au problème `cible` par un ensemble d'expressions.

Dans notre formalisme, puisque nous nous appuyons sur une représentation utilisant des attributs, nous ramenons la variation $\Delta\mathtt{pb}$ de `srce` à `cible` à la variation $\Delta_\mathtt{a}\mathtt{pb}$ de `srce` à `cible` sous l'angle de l'attribut `a`. Ainsi, $\Delta\mathtt{pb} = \bigcup_\mathtt{a} \Delta_\mathtt{a}\mathtt{pb}$ pour `a` : attribut apparaissant dans les descripteurs de `srce` et/ou de `cible`.

Trois situations peuvent alors se présenter. Si l'attribut `a` est introduit par un descripteur $(\mathtt{a},\mathtt{V}_\mathtt{srce})$ dans la définition de `srce` mais est absent de celle de `cible`, on représentera dans $\Delta_\mathtt{a}\mathtt{pb}$ la variation de `srce` à `cible` par l'expression $(\mathtt{a},\mathtt{V}_\mathtt{srce})^-$, composée du descripteur $(\mathtt{a},\mathtt{V}_\mathtt{srce})$ marqué par un signe « − ». On aura $\Delta_\mathtt{a}\mathtt{pb} = \{(\mathtt{a},\mathtt{V}_\mathtt{srce})^-\}$. Ce signe « − » dénote donc le retrait d'une propriété lors du passage de `srce` à `cible`.

Si l'attribut `a` est introduit par un descripteur $(\mathtt{a},\mathtt{V}_\mathtt{cible})$ dans la définition de `cible` mais est absent de celle de `srce`, on représentera dans $\Delta_\mathtt{a}\mathtt{pb}$ la variation de `srce` à `cible` par l'expression $(\mathtt{a},\mathtt{V}_\mathtt{cible})^+$, composée du descripteur $(\mathtt{a},\mathtt{V}_\mathtt{cible})$ marqué par un signe « + ». On aura $\Delta_\mathtt{a}\mathtt{pb} = \{(\mathtt{a},\mathtt{V}_\mathtt{cible})^+\}$. Ce signe « + » dénote donc l'ajout d'une propriété lors du passage de `srce` à `cible`.

Si enfin l'attribut `a` est introduit à la fois dans la définition de `srce` par un descripteur $(\mathtt{a},\mathtt{V}_\mathtt{srce})$ et dans la définition de `cible` par un descripteur $(\mathtt{a},\mathtt{V}_\mathtt{cible})$, $\Delta_\mathtt{a}\mathtt{pb}$ sera défini sur la base de la variation dans $\mathcal{E}_\mathtt{a}$ entre $\mathtt{V}_\mathtt{srce}$ et $\mathtt{V}_\mathtt{cible}$. Plus précisément si cette variation, dénotée par $\Delta\mathtt{V}$, prend la forme d'un ensemble d'expressions δ, alors $\Delta_\mathtt{a}\mathtt{pb} = \{\mathtt{a}^\delta \mid \delta \in \Delta\mathtt{V}\}$.

Considérons l'exemple suivant :

$$\begin{array}{rcccccccc} \mathtt{srce} & = & (\mathtt{b},\mathtt{W}_\mathtt{srce}) & \wedge & (\mathtt{c},\mathtt{X}_\mathtt{srce}) & \wedge & (\mathtt{d},\mathtt{Y}_\mathtt{srce}) & & \\ \mathtt{cible} & = & & & (\mathtt{c},\mathtt{X}_\mathtt{cible}) & \wedge & (\mathtt{d},\mathtt{Y}_\mathtt{cible}) & \wedge & (\mathtt{e},\mathtt{Z}_\mathtt{cible}) \end{array}$$

L'attribut `b` est introduit dans la définition de `srce` par le descripteur $(\mathtt{b},\mathtt{W}_\mathtt{srce})$ mais est absent de la définition de `cible`, donc $\Delta_\mathtt{b}\mathtt{pb} = \{(\mathtt{b},\mathtt{W}_\mathtt{srce})^-\}$. L'attribut `e` est introduit dans la définition de `cible` par le descripteur $(\mathtt{e},\mathtt{Z}_\mathtt{cible})$ mais est absent de la définition de `srce`, donc $\Delta_\mathtt{e}\mathtt{pb} = \{(\mathtt{e},\mathtt{Z}_\mathtt{cible})^+\}$. Les attributs communs à `srce` et à `cible` sont `c` et `d`, donc on représentera $\Delta_\mathtt{c}\mathtt{pb}$ en s'appuyant sur la variation $\Delta\mathtt{X}$ de $\mathtt{X}_\mathtt{srce}$ à $\mathtt{X}_\mathtt{cible}$, et on représentera $\Delta_\mathtt{d}\mathtt{pb}$ en s'appuyant sur la variation $\Delta\mathtt{Y}$ de $\mathtt{Y}_\mathtt{srce}$ à $\mathtt{Y}_\mathtt{cible}$. Si on suppose que $\Delta\mathtt{X} = \{\delta_1,\delta_2\}$ et que $\Delta\mathtt{Y} = \{\delta_3,\delta_4,\delta_5\}$, alors

$$\begin{aligned} \Delta\mathtt{pb} &= \Delta_\mathtt{b}\mathtt{pb} \cup \Delta_\mathtt{c}\mathtt{pb} \cup \Delta_\mathtt{d}\mathtt{pb} \cup \Delta_\mathtt{e}\mathtt{pb} \\ &= \{(\mathtt{b},\mathtt{W}_\mathtt{srce})^-, \mathtt{c}^{\delta_1}, \mathtt{c}^{\delta_2}, \mathtt{d}^{\delta_3}, \mathtt{d}^{\delta_4}, \mathtt{d}^{\delta_5}, (\mathtt{e},\mathtt{Z}_\mathtt{cible})^+\} \end{aligned}$$

Il reste donc à étudier la représentation de la variation $\Delta\mathtt{V}$ entre $\mathtt{V}_\mathtt{srce} \in \mathcal{E}_\mathtt{a}$ et $\mathtt{V}_\mathtt{cible} \in \mathcal{E}_\mathtt{a}$, c'est-à-dire préciser ce qu'est un $\delta \in \Delta\mathtt{V}$. δ représente une relation entre deux éléments de $\mathcal{E}_\mathtt{a}$, c'est donc une relation binaire sur $\mathcal{E}_\mathtt{a}$. N'importe quelle relation binaire sur $\mathcal{E}_\mathtt{a}$ fait-elle l'affaire ? Si on garde à l'esprit le fait qu'elles sont associées à l'attribut `a` dans le cadre d'une application du RàPC, certaines relations binaires font sens et d'autres pas : c'est un problème d'ingénierie des connaissances de mettre en évidence les relations pertinentes, comme l'illustre l'exemple suivant.

Considérons un système de RàPC qui évalue le prix d'un appartement à partir de sa description. Dans ce système, un problème est un concept qui représente un ensemble d'appartements. Supposons qu'un attribut `localisation` intervienne dans la définition de certains concepts pour exprimer des contraintes sur la localisation des appartements et que les ensembles de valeurs qui lui sont associés sont des régions du plan. Plusieurs relations binaires `r` peuvent être établies entre deux régions du plan A et B, en considérant par exemple :

– Leur latitude : A r B si la région A est située au nord de la région B,

– Leur distance au centre ville : A r B si la région A est plus proche du centre ville que la région B ;

Si la position géographique d'un appartement a bien une influence sur son prix, les deux relations binaires ci-dessus n'auront pas, du point de vue de l'expert, le même intérêt pour comparer deux ensembles d'appartements. Pour déterminer le prix d'un appartement dans une ville comme Paris, la distance au centre ville sera certainement jugée plus pertinente que la latitude. La section suivante étudie plusieurs types de variations entre ensembles de valeurs.

4 Représentation de variations entre ensembles de valeurs : exemples

Cette section présente quelques exemples d'ensembles $\mathcal{E}_{\mathtt{a}}$ et de relations binaires δ sur $\mathcal{E}_{\mathtt{a}}$ potentiellement utiles pour un système de RàPC.

4.1 Variations entre singletons d'entiers naturels

Comme nous l'avons remarqué à la section 2, pour représenter un couple attribut-valeur $(\mathtt{a}, \mathtt{v})$, dans notre formalisme, il suffit de le représenter par $(\mathtt{a}, \{\mathtt{v}\})$. Si nous voulons représenter les individus ayant un âge de 30 ans, il suffit d'écrire (âge,{30}). En se limitant à des valeurs précises de l'âge, on prendra $\mathcal{E}_{\mathtt{âge}}$ l'ensemble des singletons sur $\mathbb{N}$: $\mathcal{E}_{\mathtt{âge}} = \{\{x\} \mid x \in \mathbb{N}\}$. Dans la suite de cette section, nous assimilerons un singleton $\{x\}$ à l'élément x, et donc : $\mathcal{E}_{\mathtt{âge}} = \mathbb{N}$.

Les relations binaires δ qui nous intéressent sont donc des relations binaires sur $\mathbb{N}$. Supposons que, pour l'application qui nous intéresse, la variation de l'âge x vers l'âge y s'appuie d'une part sur les relations $<, >, \leqslant, \geqslant, =$ et $\neq$ (par exemple, $<$ permet de dire qu'une personne est strictement moins âgée qu'une autre) et d'autre part sur les différences $y - x \in \mathbb{Z}$ (pour dire qu'une personne a 10 ans de plus qu'une autre). Les relations binaires sont donc $<, >, \leqslant, \geqslant, =$ et $\neq$ ainsi que les relations $\mathtt{ajouter}(\alpha)$ (pour $\alpha \in \mathbb{Z}$) telles que $x \ \mathtt{ajouter}(\alpha) \ y$ si $x + \alpha = y$.

Ainsi si (âge,{40}) ∈ src e et (âge,{30}) ∈ cible alors $\Delta_{\mathtt{âge}}\mathtt{pb} = \{\mathtt{âge}^{\mathtt{ajouter}(-10)}, \mathtt{âge}^{>}, \mathtt{âge}^{\geqslant}, \mathtt{âge}^{\neq}\}$.

4.2 Variations entre intervalles fermés de $\mathbb{R}^+$

Les relations binaires qui nous intéressent sont des relations binaires entre intervalles fermés non vides de $\mathbb{R}^+$. On notera, pour $a \in \mathbb{R}^+$, $b \in \mathbb{R}^+ \cup \{+\infty\}$ et $a \leqslant b$, $[a; b] = \{x \in \mathbb{R}^+ \mid a \leqslant x \leqslant b\}$. Supposons que dans notre cadre applicatif les relations binaires jugées les plus pertinentes pour comparer deux de ces intervalles soient les 13 relations binaires de Allen [1]. Parmi ces relations, la relation de Allen o (*overlaps*) est définie pour deux intervalles $i_1 = [a_1; b_1]$ et $i_2 = [a_2; b_2]$ par $i_1 \ o \ i_2$ si $a_1 < a_2$, $a_2 < b_1$ et $b_1 < b_2$. Par ailleurs, la relation b (*before*) est définie par $i_1 \ b \ i_2$ si $b_1 < a_2$.

Cette relation permet de comparer sous l'angle de l'attribut diamètre les deux concepts

$$\begin{aligned} \mathtt{Astéroïde} &= \{(\mathtt{diamètre}, [0; 2]), (\mathtt{luit}, \{\mathtt{non}\})\} \\ \mathtt{Planète} &= \{(\mathtt{diamètre}, [0,8; 140]), (\mathtt{luit}, \{\mathtt{non}\})\} \end{aligned}$$

car on a la relation $[0; 2]\ o\ [0,8; 140]$ entre les intervalles de valeurs associés à l'attribut diamètre. Donc $\mathtt{diamètre}^o \in \Delta\mathtt{pb}$.

L'*algèbre* de Allen est l'ensemble des 2^{13} relations binaires obtenues par disjonction des 13 relations de base. Par exemple, $m \vee o$ est la relation définie par $i_1\ (m \vee o)\ i_2$ si $i_1\ m\ i_2$ ou $i_1\ o\ i_2$ (avec m tel que $[a_1; b_1]\ m\ [a_2; b_2]$ si $b_1 = a_2$: m est l'initiale de *meets*). Parmi ces 2^{13} relations qualitatives entre intervalles fermés de $\mathbb{R}^+$, seules certaines ont un intérêt pour une application particulière : leur choix est affaire de conception d'un système de RàPC. Si $m \vee o$ fait partie des relations choisies, l'exemple suivant donnera donc $\mathtt{diamètre}^{m \vee o} \in \Delta\mathtt{pb}$.

Les relations ci-dessus, ainsi que les relations $<, >, \leqslant, \geqslant, =$ et $\neq$ définies à la section 4.1 sont souvent dites *qualitatives*, alors que les relations *ajouter*(α) $(\alpha \in \mathbb{Z})$ de 4.1 sont dites *quantitatives*. On peut

définir entre intervalles fermés de $\mathbb{R}^+$ les relations quantitatives $\texttt{ajouter-aux-bornes}(\alpha;\beta)$ ($\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$) définies par $[a_1;b_1]$ $\texttt{ajouter-aux-bornes}(\alpha;\beta)$ $[a_2;b_2]$ si $a_1 + \alpha = a_2$ et $b_1 + \beta = b_2$.

Pour reprendre l'exemple ci-dessus, on aura donc : $\texttt{diamètre}^{\texttt{ajouter-aux-bornes}(0,8;138)} \in \Delta\texttt{pb}$.

4.3 Relations ensemblistes

Les exemples ci-dessus étaient de la forme suivante : $\mathcal{E}_\mathtt{a}$ est une partie de $2^{\mathcal{D}_\mathtt{a}}$ où $\mathcal{D}_\mathtt{a}$ est un ensemble muni d'une structure. Dans 4.1 cette structure était $(\mathbb{N}, \leqslant, +)$, dans 4.2, c'était $(\mathbb{R}^+, \leqslant, +)$. Les relations δ sur $\mathcal{E}_\mathtt{a}$ s'appuyaient sur cette structure : la relation de Allen o sur $\mathbb{R}^+$ s'appuie sur la relation $\leqslant$ sur $\mathbb{R}^+$. Que faire quand on ne dispose pas *a priori* de structure sur $\mathcal{D}_\mathtt{a}$? Une possibilité est d'utiliser les relations ensemblistes $\subseteq$ et $\supseteq$ sur $\mathcal{E}_\mathtt{a}$.

Si $\mathtt{V}_\mathtt{srce}, \mathtt{V}_\mathtt{cible} \in \mathcal{E}_\mathtt{a}$ et $\mathtt{V}_\mathtt{srce} \subseteq \mathtt{V}_\mathtt{cible}$ alors $(\mathtt{a}, \mathtt{V}_\mathtt{srce})^\mathcal{I} \subseteq (\mathtt{a}, \mathtt{V}_\mathtt{cible})^\mathcal{I}$ quelle que soit l'interprétation $\mathcal{I}$. Cette relation correspond donc à une généralisation de la source vers la cible. De façon duale, si $\mathtt{V}_\mathtt{srce} \supseteq \mathtt{V}_\mathtt{cible}$ alors $(\mathtt{a}, \mathtt{V}_\mathtt{srce})^\mathcal{I} \supseteq (\mathtt{a}, \mathtt{V}_\mathtt{cible})^\mathcal{I}$: $\supseteq$ correspond à une spécialisation de la source vers la cible.

Si $\mathcal{D}_\mathtt{a}$ est un ensemble fini et de petite taille (correspondant à un « type énuméré ») et $\mathcal{E}_\mathtt{a} = 2^{\mathcal{D}_\mathtt{a}}$, l'implantation de $\subseteq$ et $\supseteq$ revient à un simple problème d'inclusion entre ensembles décrits en extension. En s'appuyant sur ces deux relations, si $(\texttt{luit}, \{\texttt{oui}\}) \in \texttt{Étoile} = \texttt{srce}$ et $(\texttt{luit}, \{\texttt{oui}, \texttt{non}\}) \in \texttt{ObjetCéleste} = \texttt{cible}$ alors $\texttt{luit}^\subseteq \in \Delta\texttt{pb}$.

Si $\mathcal{D}_\mathtt{a}$ est de grande taille voire infini, une façon de faire consiste à utiliser une expression — atomique ou définie — pour représenter un $\mathtt{V} \in \mathcal{E}_\mathtt{a}$; par exemple une classe dans les représentations par objets [11] ou un concept dans les logiques de descriptions [2]. La relation $\subseteq$ est alors implantée par la relation de subsomption $\sqsubseteq$ (et $\supseteq$ par $\sqsupseteq$).

À première vue, la composition des deux relations $\subseteq$ et $\supseteq$ n'est pas pertinente car elle correspond à la relation universelle : pour tous $\mathtt{V}_\mathtt{srce}, \mathtt{V}_\mathtt{cible} \in \mathcal{E}_\mathtt{a}$ il existe $\mathtt{V} \in 2^{\mathcal{D}_\mathtt{a}}$ tel que $\mathtt{V}_\mathtt{srce} \subseteq \mathtt{V}$ et $\mathtt{V} \supseteq \mathtt{V}_\mathtt{cible}$. Il suffit en effet de prendre $\mathtt{V} = \mathcal{D}_\mathtt{a}$ ou, si on veut le $\mathtt{V}$ le plus spécifique possible, $\mathtt{V} = \mathtt{V}_\mathtt{srce} \cup \mathtt{V}_\mathtt{cible}$ (ou encore, en logique de descriptions de prendre un plus petit subsumant commun de $\mathtt{V}_\mathtt{srce}$ et $\mathtt{V}_\mathtt{cible}$, ce qui coïncide avec la disjonction des deux concepts si ce constructeur appartient à la logique de descriptions choisie, ce qui ramène à l'union). Pourtant, c'est un mode de comparaison fréquemment utilisé en RàPC : si $\mathtt{V}_\mathtt{srce}$ et $\mathtt{V}_\mathtt{cible}$ partagent le fait d'être sous-ensembles de $\mathtt{V}$, alors ils sont similaires. Pour résoudre ce paradoxe apparent, il suffit de rajouter des conditions sur $\mathtt{V}$: seuls certains super-ensembles $\mathtt{V}$ communs à $\mathtt{V}_\mathtt{srce}$ et $\mathtt{V}_\mathtt{cible}$ conviennent et ce choix est, encore une fois, affaire de conception du système de RàPC considéré. Par exemple, si $\mathtt{V}_\mathtt{srce}$ est l'ensemble des canards et $\mathtt{V}_\mathtt{cible}$ l'ensemble des moustiques, pour une application du RàPC sur le domaine du mode de locomotion animal, on pourra considérer l'ensemble $\mathtt{V}$ des animaux volants, lequel ne sera pas pertinent pour une application culinaire si on considère que la capacité de voler n'a pas de rôle à jouer en cuisine (et, de fait, le moustique ne se cuisine pas comme le canard). On peut ainsi définir les relations $(\subseteq\mathtt{V}\supseteq)$ définies par $\mathtt{V}_\mathtt{srce}$ $(\subseteq\mathtt{V}\supseteq)$ $\mathtt{V}_\mathtt{cible}$ si $\mathtt{V}_\mathtt{srce} \subseteq \mathtt{V}$ et $\mathtt{V} \supseteq \mathtt{V}_\mathtt{cible}$ pour les $\mathtt{V} \in \mathcal{E}_\mathtt{a}$ qui sont jugés pertinents (une extension de cette relation est proposée dans [10]).

4.4 Autres représentations de variations

Autres singletons, autres intervalles. Ce que nous avons présenté à la section 4.1 sur les singletons d'entiers se transpose aux singletons de réels ou à d'autres singletons : les relations binaires sur un ensemble E se transposent en relations binaires sur $\mathcal{E} = \{\{x\} \mid x \in E\}$. De la même façon, ce que nous avons présenté à la section 4.2 sur les intervalles de réels se transpose aux intervalles d'entiers et, de façon plus générale, aux intervalles fermés d'un ensemble E muni d'un ordre total. L'extension à des intervalles quelconques de $\mathbb{R}$ (avec p. ex. des intervalles ouverts) devrait s'appuyer sur des idées similaires.

Relations temporelles et spatiales. Les relations de Allen ont été conçues à l'origine pour une représentation temporelle qualitative et nous les avons réutilisées à la section 4.2 pour comparer des

intervalles. De la même façon on peut réutiliser les relations RCC-8 pour comparer deux régions de l'espace [12].[4] De façon plus générale, on peut s'intéresser aux travaux sur les représentations temporelles et spatiales [4]. Ainsi, l'exemple donné en fin de section 3 (sur les appartements) pourrait être traité à l'aide de relations spatiales telles que la distance qualitative, la position relative, l'orientation, etc.

Les travaux sur les représentations temporelles et spatiales apportent beaucoup à l'étude des comparaisons entre unités de temps (instants, intervalles de temps), d'espace (point, région) voire d'unités spatio-temporelles. Il est donc naturel de s'inspirer de ces travaux et nous ne nous en privons pas, y compris pour représenter la variation entre deux cas de deux intervalles ne représentant pas des intervalles temporels mais, par exemple, des intervalles de tailles de tumeurs, même s'il convient d'être prudent et d'étudier soigneusement quelles relations de Allen font sens dans ce contexte.

5 Vers une application à l'extraction de connaissances d'adaptation

Selon le principe de l'analogie par transformation, l'adaptation d'un cas source ($\texttt{srce}, \texttt{Sol}(\texttt{srce})$) pour résoudre un problème $\texttt{cible}$ se fait en inférant la variation $\Delta\texttt{sol}$ entre solutions à partir de la variation $\Delta\texttt{pb}$ entre problèmes. Pour mener à bien l'étape d'adaptation, un système de raisonnement à partir de cas suivant ce principe doit être pourvu des connaissances d'adaptation nécessaires à la construction d'une variation entre solutions $\Delta\texttt{sol}$ pour chaque variation entre problèmes $\Delta\texttt{pb}$ rencontrée.

CabamakA est un système d'extraction de connaissances qui permet d'acquérir ces connaissances d'adaptation par une analyse systématique des variations entre cas au sein de la base de cas [3, 6]. L'idée est de faire jouer à un cas source de la base de cas le rôle du cas remémoré et à un autre cas source le rôle du problème cible à résoudre, accompagné de sa solution. Le système représente pour chacun de ces couples de cas sources de la base de cas la variation entre cas $\Delta\texttt{cas} = \Delta\texttt{pb} \cup \Delta\texttt{sol}$, puis cherche dans l'ensemble des couples ainsi formés des associations fréquentes entre variations entre problèmes et variations entre solutions.

Cette recherche est effectuée en représentant sous la forme d'un objet — un objet étant un ensemble de propriétés booléennes — chacune des variations $\Delta\texttt{cas}$ entre cas sources de la base de cas, puis en appliquant l'algorithme Charm d'extraction de motifs fermés fréquents [16] sur l'ensemble d'objets ainsi obtenu. Dans cet article, la variation $\Delta\texttt{cas}$ entre deux cas est représentée sous la forme d'un ensemble d'expressions. Pour obtenir une représentation équivalente de $\Delta\texttt{cas}$ sous la forme d'un objet, il suffit de faire correspondre à chacune de ces expressions une propriété booléenne.

Supposons par exemple que parmi les variations entre cas de la base de cas pour une application du RàPC à l'aide à la décision thérapeutique dans le cadre du traitement du cancer du sein [9] figurent les deux variations entre cas $\Delta\texttt{cas}$ et $\Delta\texttt{cas'}$ suivantes :

$$\Delta\texttt{cas} = \{\texttt{taille-tumeur}^{b}, \texttt{âge}^{\texttt{ajouter-aux-bornes}(15;10)}, \texttt{âge}^{b}, \texttt{DOSE-CHIMIO}^{\texttt{ajouter}(-10)}, \texttt{DOSE-CHIMIO}^{\geqslant}\}$$

$$\Delta\texttt{cas'} = \{\texttt{âge}^{\texttt{ajouter-aux-bornes}(15;10)}, \texttt{âge}^{b}, \texttt{DOSE-CHIMIO}^{\texttt{ajouter}(-10)}, \texttt{DOSE-CHIMIO}^{\geqslant}, (\texttt{DOSE-RADIO}, [0;50])^{+}\}$$

(on utilise les minuscules pour les attributs de problèmes et les majuscules pour les attributs de solutions). De l'ensemble des variations entre cas sources de la base de cas, CabamakA pourra extraire le motif

$$m = \{\texttt{âge}^{b}, \texttt{âge}^{\texttt{ajouter-aux-bornes}(15;10)}, \texttt{DOSE-CHIMIO}^{\geqslant}, \texttt{DOSE-CHIMIO}^{\texttt{ajouter}(-10)}\}$$

Ce motif est une généralisation des variations entre cas $\Delta\texttt{cas}$ et $\Delta\texttt{cas'}$ et constitue une règle d'adaptation concrète au sens où elle permet de construire une solution pour un problème $\texttt{cible}$ à partir d'un cas source ($\texttt{srce}, \texttt{Sol}(\texttt{srce})$). En effet, si les deux problèmes $\texttt{srce}$ et $\texttt{cible}$ varient uniquement par les ensembles de valeurs qui leur sont associés pour l'attribut $\texttt{âge}$, et que ces ensembles de valeurs sont des intervalles $[a_1; b_1]$ pour $\texttt{srce}$ et $[a_2; b_2]$ pour $\texttt{cible}$ tels que $a_2 = a_1 + 15$ et $b_2 = b_1 + 10$, alors

[4] RCC-8 forme une algèbre de relations topologiques entre régions qui permet, par exemple, d'indiquer qu'une région est tangentiellement incluse dans une autre région.

une solution Sol(cible) peut être construite pour cible à partir de la solution Sol(srce) de srce en retranchant 10 à la dose de chimiothérapie.

Si alors parmi les motifs extraits par CabamakA se trouve également le motif

$$m' = \left\{\hat{\texttt{age}}^{\texttt{ajouter-aux-bornes(20;10)}}, \hat{\texttt{age}}^{b}, \texttt{DOSE-CHIMIO}^{\texttt{ajouter}(-20)}, \texttt{DOSE-CHIMIO}^{>}\right\}$$

on trouvera aussi (toujours extrait par Charm) le motif

$$M = m \cap m' = \left\{\hat{\texttt{age}}^{b}, \texttt{DOSE-CHIMIO}^{>}\right\}$$

qui généralise m et m' et traduit le fait que quand l'âge augmente, la dose prescrite pour la chimiothérapie diminue.

Ce motif constitue une règle abstraite (i.e., non concrète) car elle n'est pas directement opérationnalisable pour construire une solution à un problème cible à partir d'un cas source. Néanmoins cette règle a l'avantage d'être intelligible pour l'analyste, ce qui facilite sa validation, et peut être utilisée pour regrouper les règles que M généralise, telles que m et m'.

Le formalisme utilisé ici pour représenter les variations entre cas est plus expressif que celui qui était utilisé jusqu'à présent dans CabamakA, et qui permettait uniquement de représenter les variations de présence/absence de propriétés booléennes lorsqu'on passe d'un cas à un autre. L'utilisation d'un formalisme plus expressif pour représenter les variations entre cas permet d'obtenir des règles d'adaptation plus expressives.

6 Discussion et travaux proches

Vers une application à l'adaptation différentielle. Dans le rapport de recherche [7], une approche générale de l'adaptation est présentée que nous qualifions d'adaptation différentielle puisqu'elle s'appuie métaphoriquement sur la formule suivante (en la généralisant) :

$$dy_j = \sum_i \frac{\partial y_j}{\partial x_i} dx_i$$

pour y : fonction différentiable de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Ce qui donne l'approximation :

$$\Delta y_j \simeq \sum_i \frac{\partial y_j}{\partial x_i} \Delta x_i$$

L'idée de l'adaptation différentielle est de s'inspirer de cette formule pour inférer une variation $\Delta_{\texttt{A}}\texttt{sol} = \Delta y_j$ de descripteur solution (l'indice j correspond à l'attribut A dans cette métaphore) à partir des variations $\Delta_{\texttt{a}}\texttt{pb} = \Delta x_i$ de descripteurs problème (i correspond à a) et de la dépendance $\frac{\partial y_j}{\partial x_i}$ (qui indique comment varie l'ensemble de valeurs associé à A au voisinage de $\texttt{V}_{\texttt{Sol(srce)}}$ quand varie l'ensemble de valeurs associé à a au voisinage de srce). Dans le rapport de recherche cité ci-dessus, l'adaptation différentielle est définie plus précisément. Ce que le travail présenté dans le présent article peut apporter à cette approche de l'adaptation, c'est une étude sur ce que peuvent être les variations Δx_i et Δy_j pour une application particulière utilisant l'adaptation différentielle.

Relations fonctionnelles ou non fonctionnelles. Dans les sections 4.1 et 4.2, la distinction entre relations quantitatives et qualitatives coïncide avec la distinction entre relations fonctionnelles et non fonctionnelles. On rappelle que δ est une relation fonctionnelle si on a l'implication ($x\ \delta\ y_1$ et $x\ \delta\ y_2$) $\Rightarrow y_1 = y_2$. La distinction entre relations fonctionnelles et non fonctionnelles est pertinente en particulier pour l'adaptation de la solution Sol(srce) en une solution Sol(cible) s'appuyant sur Δsol. Si $\texttt{A}^{\delta} \in \Delta\texttt{sol}$ et $(\texttt{A}, \texttt{V}_{\texttt{Sol(srce)}}) \in \texttt{Sol(srce)}$ alors :

– Si δ est fonctionnelle alors il n'existe (au plus) qu'un $\mathtt{V}_{\mathtt{Sol(cible)}} \in \mathcal{E}_{\mathtt{A}}$ tel que $\mathtt{V}_{\mathtt{Sol(srce)}}\ \delta\ \mathtt{V}_{\mathtt{Sol(cible)}}$, ce qui résout `cible` pour l'attribut solution `A` ;
– Si δ est non fonctionnelle alors $\mathtt{V}_{\mathtt{Sol(cible)}} \in \mathcal{E}_{\mathtt{A}}$ est un des $\mathtt{V} \in \mathcal{E}_{\mathtt{A}}$ qui satisfait la contrainte $\mathtt{V}_{\mathtt{Sol(srce)}}\ \delta\ \mathtt{V}$.

Choix des relations pertinentes. Dans cet article, nous n'avons pas abordé la question du choix des relations δ pertinentes : nous nous sommes contentés de dire que c'est le travail d'ingénierie des connaissances du concepteur du système de RàPC. Insistons cependant sur l'importance de ce choix et rappelons le lien qu'il a avec la problématique de l'adaptation et celle de la remémoration guidée par l'adaptabilité. L'adaptation consiste à « passer » d'une solution à une autre et ce passage doit être « minimisé » car il correspond à l'« effort d'adaptation » : les relations δ entre ensembles de valeurs de descripteurs solutions constituant de tels « passages » à « minimiser », leur choix n'est pas indifférent[5]. La remémoration guidée par l'adaptabilité [15] est le principe selon lequel le cas remémoré doit idéalement être adaptable pour résoudre le problème cible avec un « effort d'adaptation » minimal. Ainsi, le choix des relations δ entre problèmes (sur lequel la *similarité* — ou connaissance de remémoration — doit être définie) doit refléter le choix des relations δ entre solutions (qui sont elles liées aux connaissances d'adaptation).

Autres travaux sur la représentation des variations entre cas. Dans [14] est présentée une approche pour comparer deux cas représentés par des concepts `Cas1` et `Cas2` de la logique de descriptions C-Classic. Plus précisément, cette comparaison produit trois concepts `PPSC`, $\mathtt{DIFF}_{12}$ et $\mathtt{DIFF}_{21}$: `PPSC` est le plus petit subsumant commun de `Cas1` et `Cas2` (unique dans ce formalisme) et $\mathtt{DIFF}_{12}$ (resp. $\mathtt{DIFF}_{21}$) est un concept approchant dans ce formalisme la différence ensembliste de `Cas1` et `Cas2` (resp., de `Cas2` et `Cas1`). Cette approche rejoint celle présentée à la section 4.3 au sens où elle ne s'appuie pas sur des relations sur les ensembles de valeurs, en dehors de relations ensemblistes ($\subseteq$, $\supseteq$, etc.). Néanmoins, une étude plus approfondie de ce travail devrait être utile pour approfondir les idées présentées dans 4.3.

Une autre approche pour réifier la similarité (et donc, les variations) d'un problème `srce` à un problème `cible` est d'utiliser un *chemin de similarité*, à savoir un chemin dans l'espace des problèmes structuré par un ensemble $\mathcal{R}$ de relations binaires : un tel chemin a donc la forme $\mathtt{pb}_0\ \mathtt{r}_1\ \mathtt{pb}_0\ \mathtt{r}_1\ \mathtt{pb}_2\ \ldots\ \mathtt{pb}_{q-1}\ \mathtt{r}_1\ \mathtt{pb}_q$ tel que $\mathtt{pb}_0 = \mathtt{srce}$, $\mathtt{pb}_q = \mathtt{cible}$, $\mathtt{r}_i \in \mathcal{R}$ $(1 \leqslant i \leqslant q)$ et $\mathtt{pb}_i$ est un problème dit intermédiaire $(1 \leqslant i \leqslant q-1)$. De même, on définit un *chemin d'adaptation*, qui consiste en l'adaptation de $\mathtt{Sol}(\mathtt{srce}) = \mathtt{Sol}(\mathtt{pb}_0)$ en $\mathtt{Sol}(\mathtt{pb}_q) = \mathtt{Sol}(\mathtt{cible})$ et contient les q étapes d'adaptation de $\mathtt{Sol}(\mathtt{pb}_{i-1})$ en $\mathtt{Sol}(\mathtt{pb}_i)$ $(1 \leqslant i \leqslant q)$, chaque étape s'appuyant sur une *reformulation* $(\mathtt{r}_i, \mathcal{A}_{\mathtt{r}_i})$ où $\mathcal{A}_{\mathtt{r}_i}$ est une fonction d'adaptation expliquant comment $\mathtt{Sol}(\mathtt{pb}_{i-1})$ peut être adaptée en $\mathtt{Sol}(\mathtt{pb}_i)$ sous la condition $\mathtt{pb}_{i-1}\ \mathtt{r}_i\ \mathtt{pb}_i$. Cette approche, décrite plus en détail par exemple dans [8], permet donc la construction d'une similarité complexe sur la base de similarités (variations) simples. L'approche présentée dans le présent article s'intéresse à l'inverse aux variations simples entre parties de cas : ces deux approches sont donc complémentaires.

7 Conclusion et perspectives

Cet article présente une approche pour représenter la variation entre cas dans le cadre d'un formalisme attribut-ensemble de valeurs (ou attribut-contrainte). L'idée est de mettre en évidence (−) les descripteurs apparaissant dans le premier cas et pas dans le deuxième, (+) les descripteurs apparaissant dans le deuxième cas et pas dans le premier et (δ) pour un attribut a apparaissant dans les deux cas de se ramener à la représentation de la variation $\Delta\mathtt{V}$ entre deux ensembles de valeurs $\mathtt{V}_{\mathtt{srce}}$ et $\mathtt{V}_{\mathtt{cible}}$. $\Delta\mathtt{V}$ est constitué d'un ensemble de relations binaires δ telles que $\mathtt{V}_{\mathtt{srce}}\ \delta\ \mathtt{V}_{\mathtt{cible}}$, parmi

[5] Les nombreux guillemets de cette phrase signifient qu'elle s'appuie sur des notions intuitives ou, pour botter en touche, dépendantes de l'application du RàPC considérée.

celles que le concepteur du système de RàPC a choisi de représenter comme étant pertinentes pour passer d'un ensemble de valeurs de l'attribut a à un autre. Certaines de ces relations δ sont atomiques (telles que $<$, o ou $\supseteq$), d'autres sont définies (telles que $\texttt{ajouter}(\alpha)$, $m \vee o$ ou $(\subseteq \mathbb{V} \supseteq)$). Des exemples de telles relations ont été présentées, qui nous semblent utiles pour diverses applications.

La motivation initiale de ce travail est le développement du système CabamakA d'extraction de connaissances d'adaptation par fouille de la base de cas : une représentation de variations plus sophistiquée que celle actuellement implantée dans ce système devrait permettre de faciliter le travail de l'analyste en regroupant les règles d'adaptation actuellement extraites par l'introduction de règles d'adaptation abstraites. Cette étude devrait également s'appliquer à la théorie de l'adaptation différentielle : celle-ci s'appuie sur des variations entre descripteurs de cas et les exemples présentés ici tout comme la démarche générale proposée devrait être utiles pour instancier à d'autres descripteurs que les seuls descripteurs numériques cette approche de l'adaptation.

D'un point de vue applicatif, la perspective principale de ce travail est son application effective à CabamakA.

D'un point de vue théorique, outre l'étude plus approfondie de l'adaptation différentielle à travers ce travail, il est envisagé de pousser plus avant l'étude de la représentation des variations. L'idéal serait d'arriver à la proposition d'une syntaxe ayant une sémantique formelle associée, à l'image de ce qui se fait pour les logiques de descriptions. Ainsi, on aimerait pouvoir faire des inférences déductives sur les ensembles de relations δ (ensembles interprétés conjonctivement), telles que : $\{\texttt{ajouter}(-10)\} \models \{>, \geqslant, \neq\}$ ou $\{\leqslant, \geqslant\} \equiv \{=\}$.

Une dernière perspective serait d'étudier comment systématiser le passage d'une structure pertinente (du point de vue du système de RàPC à concevoir) sur $\mathcal{D}_a$ à une structure pertinente sur le sous-ensemble $\mathcal{E}_a$ de $2^{\mathcal{D}_a}$.

Remerciements

Les auteurs tiennent à remercier les relecteurs pour leurs remarques encourageantes et intéressantes.

Références

[1] J. F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11) :832–843, 1983.

[2] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, et P. Patel-Schneider, editors. *The Description Logic Handbook*. Cambridge University Press, cambridge, UK, 2003.

[3] F. Badra et J. Lieber. Extraction de connaissances d'adaptation par l'analyse de la base de cas. In *Extraction et gestion des connaissances (EGC'2007), Actes des septièmes journées Extraction et Gestion des Connaissances, Namur, Belgique, 23-26 janvier 2007, 2 Volumes*, Revue des Nouvelles Technologies de l'Information, pages 751–760. Cépaduès-Éditions, 2007.

[4] Florence Le Ber et Gérard Ligozat. Actes de l'atelier « Représentation et raisonnement sur le temps et l'espace », plateforme AFIA, Nice, 2005.

[5] J. Carbonell. Learning by analogy : Formulating and generalizing plans from past experience. In R. Michalski, J. Carbonell, et T. Mitchell, editors, *Machine Learning : An Artificial Intelligence Approach*, pages 137–162. Tioga, Cambridge, MA, 1983.

[6] M. d'Aquin, F. Badra, S. Lafrogne, J. Lieber, A. Napoli, et L. Szathmary. Case base mining for adaptation knowledge acquisition. In *Proceedings of the International Conference on Artificial Intelligence, IJCAI'07*, pages 750–756, 2007.

[7] 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/Ecole Centrale de Lyon, 2006.

[8] J. Lieber. Reformulations and Adaptation Decomposition. In J. Lieber, E. Melis, A. Mille, et A. Napoli, editors, *Formalisation of Adaptation in Case-Based Reasoning*. Third International Conference on Case-Based Reasoning Workshop, ICCBR-99 Workshop number 3, S. Schmitt and I. Vollrath (volume editor), LSA, University of Kaiserslautern, 1999.

[9] J. Lieber, M. d'Aquin, F. Badra, et A. Napoli. Case-Based Treatment Recommendations for Breast Cancer. *Applied Intelligence (an International Journal)*, devrait paraître en 2007.

[10] J. Lieber et P. Marquis. Domain-Independent Similarity Relations for Case-Based Reasoning in a Logical Framework. In *Proceedings of the Poster Session of the Ninth International Symposium on Methodologies for Intelligent Systems (ISMIS'96), Zakopane (Poland), 9-13 June*, pages 230–241. Oak Ridge National Laboratory, 1996.

[11] A. Napoli, C. Laurenço, et R. Ducournau. An object-based representation system for organic synthesis planning. *International Journal of Human-Computer Studies*, 41(1/2) :5–32, 1994.

[12] D. A. Randell, Z. Cui, et A. G. Cohn. A spatial logic based on regions and connection. In *Knowledge Representation*, pages 165–176, 1992.

[13] C. K. Riesbeck et R. C. Schank. *Inside Case-Based Reasoning*. Lawrence Erbaum Associates, Inc., Hillsdale, New Jersey, 1989.

[14] S. Salotti et V. Ventos. Approche formelle du raisonnement à partir de cas dans une logique de descriptions. *Revue d'Intelligence Artificielle*, 13 :37–7, 1999.

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

[16] M. J. Zaki et C.-J. Hsiao. CHARM : An efficient algorithm for closed itemset mining. In Robert L. Grossman, Jiawei Han, Vipin Kumar, Heikki Mannila, et Rajeev Motwani, editors, *Proceedings of the Second SIAM International Conference on Data Mining, Arlington, VA, USA, April 11-13, 2002*. SIAM, 2002.