Sur la représentation des co-variations

Fadi Badra

Université Paris 13, Sorbonne Paris Cité, LIMICS, (UMR_S 1142), F-93430, Villetaneuse, France
Sorbonne Universités, UPMC Univ Paris 06, UMR_S 1142, LIMICS, F-75006, Paris, France
INSERM, U1142, LIMICS, F-75006, Paris, France
badra@univ-paris13.fr

Résumé

L'abondance actuelle des données disponibles motive la recherche pour représenter les différences qui existent entre plusieurs traces numériques d'un même processus. Dans des travaux précédents, nous avons proposé une représentation symbolique des variations entre plusieurs éléments d'un même ensemble. Dans cet article, nous allons plus loin et définissons la notion de co-variation comme une dépendance fonctionnelle entre variations. Nous montrons que cette définition permet d'obtenir une règle de déduction naturelle sur les variations, qui peut être facilement étendue pour définir un raisonnement par similarité. Une méthode est ensuite proposée pour apprendre des co-variations à partir de données. Dans cette méthode, les co-variations correspondent aux règles d'implication d'objets dans une structure de motifs.

Abstract

In order to make sense of the ever-growing amount of available data, it is key to be able to capture the differences between two or more digital traces of a same process. In some previous work, a symbolic representation of the variations between two or more elements of a set was proposed. This article goes one step further and defines co-variations as functional dependencies between variations. This gives us a natural deduction rule on variations, which we show can be easily extended to perform similarity-based reasoning. A method is also proposed to learn co-variations from the data. In this method, co-variations correspond to object implication rules in a pattern structure.

1 Introduction

Une variation représente un ensemble de différences entre deux ou plusieurs traces numériques (au sens de [23]) d'un même processus. Ces différences peuvent être de l'ordre du "tout ou rien", c.-à-d., exprimer le gain/la préservation/la perte d'une propriété (par exemple, *les morilles noires sont toxiques alors que les fausses morilles sont comestibles*), ou établir un changement de degré (par exemple, *Le film "Rambo" est plus violent que le film "Cendrillon"*). Être capable de les représenter constitue un véritable enjeu, car elles permettent au sein d'un processus complexe de représenter la façon dont un paramètre évolue avec son environnement. Les applications potentielles sont nombreuses, comme par exemple la comparaison d'objets en reconnaissance de formes [17], l'étude des évolutions cohérentes dans l'expression des gènes [21], la caractérisation de relations sémantiques entre entités linguistiques [14], la modélisation de l'adaptation en raisonnement à partir de cas [12, 18, 24] ou encore l'étude de l'impact des changements dans l'habitat des poissons sur leurs déplacements [5].

Cet article fait suite à de précédents travaux [3], où nous avions proposé une représentation symbolique des variations sous forme d'un attribut d'un ensemble ordonné d'objets, tous pris dans un même ensemble. Nous nous intéressons ici à la formalisation des co-occurrences qui peuvent exister entre une ou plusieurs variations, et à la définition de "règles d'inférences" entre variations qui permettent de les utiliser dans un raisonnement. En particulier, une forme d'inférence qu'il serait bon d'avoir sur les variations est le raisonnement de type modus ponens, dont le schema est le suivant :

$$\frac{\begin{array}{l} f(z,t) \\ \text{SI } f(x,y) \text{ ALORS } g(x,y) \end{array}}{g(z,t)}$$

Ce raisonnement permet de déterminer la valeur d'une variation en un point par application d'une règle de type SI ... ALORS. Cette règle doit traduire le fait que la variation g coïncide localement avec la variation

f sur un ensemble de paires d'objets qui contient le couple (z, t).

Dans cet article, nous proposons de définir la notion de co-variation comme une dépendance fonctionnelle entre variations. Cette définition permet de définir une règle de déduction naturelle sur les variations, qui peut être facilement étendue pour définir un raisonnement par similarité. Une méthode est proposée pour apprendre des co-variations à partir de données. Dans cette méthode, les co-variations correspondent aux règles d'implication d'objets dans une structure de motifs.

L'article est organisé comme suit. La section suivante fait une étude de la littérature autour de la représentation de co-variations. La Sec. 3 rappelle des définitions sur les variations et fait le lien entre variations et proportion analogique. Dans la Sec. 4, une co-variation est définie comme une dépendance fonctionnelle entre variations. Dans la Sec. 5, une règle de déduction naturelle, de type "modus ponens" est présentée. La règle d'inférence est ensuite étendue en Sec. 6 pour définir un raisonnement par similarité. La Sec. 7 présente une méthode d'apprentissage de co-variations à partir de données. La Sec. 8 conclue l'article et donne des perspectives de recherches.

2 Revue de la littérature

Dans cet article, nous proposons une approche symbolique pour représenter les co-variations, là où la détection de corrélations entre variables fait souvent appel à des techniques statistiques ou graphiques (voir par exemple [9, 20, 27]).

Une des motivations de ce travail réside dans l'utilité des variations pour réaliser l'étape d'adaptation en raisonnement à partir de cas. Dans une de ses formulations [12], l'adaptation est présentée comme la construction d'une solution $\texttt{sol}(\texttt{cible})$ pour un problème cible $\texttt{cible}$ par modification de la solution $\texttt{sol}(\texttt{srce})$ d'un problème source $\texttt{srce}$ remémoré. L'adaptation peut être décomposée en trois étapes :

① $(\texttt{srce}, \texttt{cible}) \mapsto \Delta_{\texttt{pb}}$: on représente les différences entre les problèmes $\texttt{srce}$ et $\texttt{cible}$;

② $(\Delta_{\texttt{pb}}, \texttt{CA}) \mapsto \Delta_{\texttt{sol}}$: on utilise des connaissances d'adaptation $\texttt{CA}$ pour construire une variation $\Delta_{\texttt{sol}}$ entre $\texttt{sol}(\texttt{srce})$ et le futur $\texttt{sol}(\texttt{cible})$;

③ $(\Delta_{\texttt{sol}}, \texttt{sol}(\texttt{srce})) \mapsto \texttt{sol}(\texttt{cible})$: $\texttt{sol}(\texttt{srce})$ est modifié en $\texttt{sol}(\texttt{cible})$ en appliquant $\Delta_{\texttt{sol}}$.

Selon ce modèle, réaliser l'adaptation nécessite d'être capable d'inférer des variations $\Delta_{\texttt{sol}}$ entre solutions à partir de variations $\Delta_{\texttt{pb}}$ entre problèmes (étape ②).

Une co-variation exprime la coïncidence locale de deux propriétés, qui s'appliquent chacune sur un couple d'objets. En cela, la notion de co-variation est proche de celle de dissimilarité analogique, qui est mesurée de manière numérique dans [8] comme le nombre de flips nécessaires pour que les quatre objets se retrouvent en proportion analogique, ou dans [13] comme le cosinus de deux vecteurs dans un espace euclidien $\mathbb{R}^n$.

Lorsque les variations représentent un changement de degré d'une propriété, les co-variations expriment des correspondances monotones entre deux changements graduels, à travers des règles de la forme "plus x est A, plus y est B". Au niveau lexical, de telles correspondances, appelées schémas topiques (*topoi*), sont définies comme étant la donnée de deux prédicats graduels ainsi que l'ensemble des correspondances monotones entre ces deux gradations [2]. Les topoi ont été identifiées comme jouant un grand rôle dans l'enchaînement dénoncés linguistiques [26].

Les règles d'inférence graduelles [6, 16] sont une modélisation numérique de telles correspondances, en faisant appel à des techniques de la logique floue. Ces règles ont été appliquées au raisonnement par similarité [15], et même utilisées pour modéliser l'adaptation [7]. Néanmoins, leur sémantique est différente de celle présentée ici. Une règle d'inférence graduelle modélise l'incertitude, alors que la sémantique choisie ici pour les co-variations modélise plutôt la co-occurrence, à la manière des règles d'association.

Concernant l'apprentissage de co-variations, un algorithme d'extraction de règles d'inférence graduelles à l'aide de techniques de programmation logique inductive a été proposé dans [28]. Dans [22] est proposée une méthode d'apprentissage de proportions analogiques dans des contextes formels qui met en œuvre une heuristique de diminution itérative de la dissimilarité analogique entre deux paires d'objets.

3 Variations

Cette section rappelle des définitions sur les variations et fait le lien entre variations et proportion analogique.

3.1 Définitions et notations

Une *variation* est modélisée par une fonction $f : \mathcal{X}^n \longrightarrow \mathcal{V}$ qui à un élément du produit cartésien $\mathcal{X}^n$ d'un ensemble $\mathcal{X}$ (avec $n \geq 2$) associe une valeur prise dans un ensemble $\mathcal{V}$. Dans la suite, nous supposerons que $n = 2$, donc nous ne représenterons que des variations entre couples d'éléments d'un ensemble donné $\mathcal{X}$. L'ensemble des fonctions $f : \mathcal{X}^2 \longrightarrow \mathcal{V}$ définies sur un ensemble $\mathcal{X}^2$ et à valeurs dans $\mathcal{V}$ est noté $\mathscr{V}(\mathcal{X}^2, \mathcal{V})$.

3.2 Variations et relations binaires

Lorsque $\mathcal{V} = \{0, 1\}$, l'ensemble $\mathscr{V}(\mathcal{X}^2, \{0, 1\})$ désigne les fonctions indicatrices de relations binaires sur $\mathcal{X}^2$. Par exemple, si $\mathcal{X} = \mathbb{N}$ est l'ensemble des entiers naturels, on peut définir la variation :

$$\mathbf{1}_{\leq}(x, y) = \begin{cases} 1 & \text{si } x \leq y \\ 0 & \text{sinon} \end{cases}$$

qui renvoie 1 si l'entier x est inférieur ou égal à l'entier y et 0 sinon.

3.3 Variations et proportion analogique

La notion de variation est liée à celle de proportion analogique. En effet, l'ensemble $R_v^f = \{(x, y) \in \mathcal{X}^2 \mid f(x, y) = v\}$ des antécédents d'une valeur v prise par une variation f définit une relation binaire sur $\mathcal{X}^2$, et l'égalité

$$\mathbf{1}_{R_v^f}(x, y) = \mathbf{1}_{R_v^f}(z, t)$$

traduit une proportion analogique $x : y :: z : t$ entre les éléments x, y, z et t [11]. Donc chaque valeur v que prend une variation a la propriété de mettre en proportion analogique un ensemble de couples d'éléments (tous les couples d'éléments (x,y) qui prennent la même valeur $f(x, y) = v$ pour cette variation).

3.4 Variations entre ensembles d'attributs binaires

Comment représenter des variations entre deux ensembles d'attributs binaires ? Autrement dit, si $\mathtt{M}$ désigne un ensemble d'attributs binaires, chacune représentant une propriété, et si x et y décrivent chacun un sous-ensemble de $\mathtt{M}$, peut-on trouver une fonction f qui permette d'exprimer les variations entre x et y ? Par exemple, si $\mathtt{M} = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{e}\}$ et que

$$x = \{\mathtt{a}, \mathtt{b}, \mathtt{d}\}$$
$$y = \{\mathtt{a}, \mathtt{e}\}$$

comment représenter les variations entre x et y ?

Il paraît naturel pour représenter ces variations d'introduire les quatre ensembles $x \cap y$, $x \cap \overline{y}$, $\overline{x} \cap y$ et $\overline{x} \cap \overline{y}$, qui forment à eux quatre une partition de $\mathtt{M}$. Dans l'exemple précédent, on obtient :

$$x \cap y = \{\mathtt{a}\}$$
$$x \cap \overline{y} = \{\mathtt{b}, \mathtt{d}\}$$
$$\overline{x} \cap y = \{\mathtt{e}\}$$
$$\overline{x} \cap \overline{y} = \{\mathtt{c}\}$$

Lorsque x, y, z, et t représentent des ensembles d'attributs binaires, la proportion analogique $x : y :: z : t$ est définie [25] par

$$x \cap \overline{y} = z \cap \overline{t} \text{ et } \overline{x} \cap y = \overline{z} \cap t$$

ce qui équivaut à dire que les couples (x,y) et (z,t) qui sont en proportion analogique prennent la même valeur pour la variation :

$$\upsilon(x, y) = \{\, x \cap \overline{y},\, \overline{x} \cap y \,\}$$

L'intérêt de choisir cette variation est que si on connaît x, y et z, la « règle de trois » sur les proportions analogiques permet de déterminer les éléments t. Par exemple, si

$$x = \{\mathtt{a}, \mathtt{b}, \mathtt{d}\}$$
$$y = \{\mathtt{a}, \mathtt{e}\}$$
$$z = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$$

alors $\upsilon(x, y) = \{\{\mathtt{b}, \mathtt{d}\}, \{\mathtt{e}\}\}$ et tous les ensembles d'attributs t qui vérifient $\upsilon(z, t) = \upsilon(x, y)$ sont en proportion analogique avec x, y et z. Malheureusement dans cet exemple, l'équation n'a pas de solution (Fig. 1) car z ne contient pas l'attribut $\mathtt{d}$.

	a	b	c	d	e
x	1	1	0	1	0
y	1	0	0	0	1
z	1	1	1	0	0
t	1	0	1	?	1

FIGURE 1 – Résolution de l'équation $x : y :: z : t = 1$ lorsque $x = \{\mathtt{a}, \mathtt{b}, \mathtt{d}\}$, $y = \{\mathtt{a}, \mathtt{e}\}$ et $z = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$.

4 Co-variations

Dans cette section, une co-variation est définie comme étant une dépendance fonctionnelle entre deux (ensembles de) variations.

4.1 Définitions

Nous commençons par définir la notion de co-variation entre deux variations $f, g \in \mathscr{V}(\mathcal{X}^2, \mathcal{V})$.

Définition 1. *Soient $f, g \in \mathscr{V}(\mathcal{X}^2, \mathcal{V})$. Une variation g co-varie avec une variation f sur une partie R de $\mathcal{X}^2$, noté $f \overset{R}{\rightharpoondown} g$, lorsque pour tout (x, y) et (z, t) de R :*

$$f(x, y) = f(z, t) \Rightarrow g(x, y) = g(z, t)$$

Cette définition traduit le fait que dès que les éléments du sous-ensemble R de $\mathcal{X}^2$ partagent une même valeur pour la variation f, alors ils doivent également partager une même valeur pour g. Si $R = \mathcal{X}^2$, on pourra omettre d'écrire le R : $f \rightharpoondown g \Leftrightarrow f \overset{\mathcal{X}^2}{\rightharpoondown} g$.

Cette définition est étendue à la co-variation entre deux ensembles de variations.

Définition 2. *Soient $F \subseteq \mathscr{V}(\mathcal{X}^2, \mathcal{V})$ et $G \subseteq \mathscr{V}(\mathcal{X}^2, \mathcal{V})$ deux ensembles de variations. L'ensemble G co-varie avec l'ensemble F sur une partie R de $\mathcal{X}^2$, noté $F \overset{R}{\rightharpoondown} G$, lorsque pour tout (x, y) et (z, t) de R :*

$$\forall f \in F, f(x, y) = f(z, t) \Rightarrow \forall g \in G, g(x, y) = g(z, t)$$

Cette définition traduit le fait que tout sous-ensemble de la relation R qui partage une même valeur pour toutes les variations de F doit partager également une même valeur pour toutes les variations de G. De la même façon, si $R = \mathcal{X}^2$, on pourra omettre d'écrire le R et écrire $F \rightharpoondown G$ pour $F \overset{\mathcal{X}^2}{\rightharpoondown} G$.

4.2 Co-variations et proportion analogique

Supposons que les él'ements de $\mathcal{X}$ sont des ensembles d'attributs pris dans un ensemble $\mathtt{M}$ d'attributs binaires et considérons les variations $\upsilon_{\mathtt{T}}$ définies pour tout $\mathtt{T} \subseteq \mathtt{M}$ par :

$$\upsilon_{\mathtt{T}}(x, y) = \{x \cap \overline{y} \cap \mathtt{T}, \overline{x} \cap y \cap \mathtt{T}\}$$

Les variations $\upsilon_{\mathtt{T}}$ expriment les différences entre deux ensembles d'attributs mais sur un sous-ensemble d'attributs $\mathtt{T} \subseteq \mathtt{M}$ seulement. Elles prennent une même valeur pour tous les couples (x, y) qui sont en proportion analogique sur le sous-ensemble d'attributs $\mathtt{T}$. On peut alors écrire les co-variations :

$$\upsilon_{\mathtt{S}} \overset{R}{\rightharpoondown} \upsilon_{\mathtt{T}}$$

Ces co-variations traduisent le fait que tout sous-ensemble de R qui est en proportion analogique sur $\mathtt{S}$ (*i.e.*, pour lequel la variation $\upsilon_{\mathtt{S}}$ prend la même valeur) est également en proportion analogique sur $\mathtt{T}$. Plus formellement :

Définition 3. *Soient $R \subseteq \mathcal{X}^2$ et $\mathtt{T} \subseteq \mathtt{M}$. L'ensemble R est en proportion analogique sur $\mathtt{T}$ ssi*

$$\exists v \in \mathscr{P}(\mathtt{M}) \times \mathscr{P}(\mathtt{M}) \mid \forall (x, y) \in R, \upsilon_{\mathtt{T}}(x, y) = v$$

Une partie R de $\mathcal{X}^2$ est en proportion analogique sur un ensemble d'attributs $\mathtt{T}$ si la variation $\upsilon_{\mathtt{T}}$ prend la même valeur sur R. Ici $\mathscr{P}(\mathtt{M})$ désigne l'ensemble des parties de $\mathtt{M}$.

Proposition 4.1. *$\upsilon_{\mathtt{S}} \overset{R}{\rightharpoondown} \upsilon_{\mathtt{T}}$ ssi tout sous-ensemble de R qui est en proportion analogique sur $\mathtt{S}$ est également en proportion analogique sur $\mathtt{T}$.*

Démonstration. $\Rightarrow$: Supposons que $\upsilon_{\mathtt{S}} \overset{R}{\rightharpoondown} \upsilon_{\mathtt{T}}$ et que $A \subseteq R$ est en proportion analogique sur $\mathtt{S}$. Alors (Déf. 3) il existe un v tel que $\forall (x, y) \in R$, $\upsilon_{\mathtt{S}}(x, y) = v$. Supposons qu'il existe $(x, y), (z, t) \in R$ tels que $\upsilon_{\mathtt{T}}(x, y) \neq \upsilon_{\mathtt{T}}(z, t)$. On a $\upsilon_{\mathtt{S}}(x, y) = \upsilon_{\mathtt{S}}(z, t) = v$ donc par Déf. 1, $\upsilon_{\mathtt{T}}(x, y) = \upsilon_{\mathtt{T}}(z, t)$. Contradiction.
$\Leftarrow$: Soient $(x, y), (z, t) \in R$ tels que $\upsilon_{\mathtt{S}}(x, y) = \upsilon_{\mathtt{S}}(z, t)$. L'ensemble $\{(x, y), (z, t)\} \subseteq R$ est en proportion analogique sur $\mathtt{S}$ donc sur $\mathtt{T}$. Donc (Déf. 3) il existe v tel que $\upsilon_{\mathtt{T}}(x, y) = \upsilon_{\mathtt{T}}(z, t) = v$. □

5 Un raisonnement à base de règles

Dans cette section, les co-variations sont utilisées dans un raisonnement à base de règles pour prédire la valeur d'une variation en un point lorsqu'on connaît déjà les valeurs prises en ce point par une ou plusieurs autres variations.

5.1 Modus ponens sur les variations

Une règle d'inférence de type "modus ponens" est définie sur les variations.

Définition 4. *Soient $f, g \in \mathscr{V}(\mathcal{X}^2, \mathcal{V})$. La règle d'inférence de type "modus ponens" sur les variations est la suivante :*

$$\frac{\begin{array}{l} f(x, y) = f(z, t) \text{ pour } (x, y), (z, t) \in R \\ f \overset{R}{\rightharpoondown} g \end{array}}{g(x, y) = g(z, t)} \tag{MP}$$

Cette règle dit que si g co-varie avec f sur une partie R, alors toute paire de couples (x, y) et (z, t) de R qui prennent la même valeur pour f prennent également la même valeur pour g.

Cette règle peut être étendue à des co-variations $F \overset{R}{\rightharpoondown} G$ entre deux ensembles de variations F et G.

Définition 5. *Pour $F \subseteq \mathscr{V}(\mathcal{X}^2, \mathcal{V})$, $G \subseteq \mathscr{V}(\mathcal{X}^2, \mathcal{V})$, $f \in F$ et $g \in G$, la règle d'inférence de type "modus ponens" sur les variations est la suivante :*

$$\frac{\begin{array}{l} \forall f \in F,\ f(x, y) = f(z, t) \text{ pour } (x, y), (z, t) \in R \\ F \overset{R}{\rightharpoondown} G \end{array}}{\forall g \in G,\ g(x, y) = g(z, t)}$$

Dans la suite de la section, nous donnons deux exemples de ce type de raisonnement.

5.2 Exemple : le raisonnement a fortiori

Lorsque les deux variations f et g sont les fonctions indicatrices de relations d'ordre, l'inférence de type modus ponens sur les variations correspond au raisonnement *a fortiori* [1]. Ce mode de raisonnement exploite la co-monotonie de deux relations d'ordre pour évaluer la valeur d'un attribut. Les auteurs de [14] donnent l'exemple suivant. Si l'on considère que plus un alcool est fort, plus l'âge légal pour le boire doit être élevé, et si l'âge légal pour boire de la bière est 18 ans, alors l'âge légal pour boire du whisky doit être au moins 18 ans. Appelons $\mathcal{A}$ l'ensemble des alcools et $\leq_{\text{degré}}$ et $\leq_{\text{âge\_légal}}$ les relations d'ordre sur $\mathcal{A}$ qui ordonnent les alcools respectivement suivant leur degré et l'âge légal autorisé pour les boire. L'exemple peut s'interpréter comme la co-variation $\mathbf{1}_{\leq \text{degré}} \rightharpoonup \mathbf{1}_{\leq \text{âge\_légal}}$ entre deux variations $\mathbf{1}_{\leq \text{degré}}$ et $\mathbf{1}_{\leq \text{âge\_légal}}$ de $\mathscr{V}(\mathcal{A}^2, \{0,1\})$ qui représentent respectivement les fonctions indicatrices des deux relations $\leq_{\text{degré}}$ et $\leq_{\text{âge\_légal}}$. La règle d'inférence (MP) donne [1] :

$$\frac{\begin{array}{l}\mathbf{1}_{\leq \text{degré}}(\texttt{biere}, \texttt{whisky}) = 1 \\ \mathbf{1}_{\leq \text{degré}} \rightharpoonup \mathbf{1}_{\leq \text{âge\_légal}}\end{array}}{\mathbf{1}_{\leq \text{âge\_légal}}(\texttt{biere}, \texttt{whisky}) = 1}$$

et l'on conclut que l'âge légal pour boire du whisky est au moins égal à l'âge légal pour boire de la bière.

5.3 Exemple des proportions analogiques

On suppose que $\mathtt{M} = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{e}\}$ est un ensemble d'attributs binaires. Pour $\mathtt{T} \subseteq \mathtt{M}$, les variations $\upsilon_{\mathtt{T}}$ sont celles définies à la Sec. 4.2. Supposons que l'on connaisse la règle

$$\upsilon_{\{\mathtt{a},\mathtt{b}\}} \overset{R}{\rightharpoonup} \upsilon_{\{\mathtt{e}\}}$$

sur l'ensemble $R = \{(x, y), (z, t)\}$. Cette règle dit que tout couple de R qui est en proportion analogique sur $\{\mathtt{a}, \mathtt{b}\}$ est aussi en proportion analogique sur $\{\mathtt{e}\}$. Si par exemple, on a

$$x = \{\mathtt{a}, \mathtt{b}, \mathtt{d}\}$$
$$y = \{\mathtt{a}, \mathtt{e}\}$$
$$z = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$$

et que l'on sait déjà que t contient $\mathtt{a}$ et non $\mathtt{b}$, alors la règle d'inférence (MP) donne :

$$\frac{\begin{array}{l}\upsilon_{\{\mathtt{a},\mathtt{b}\}}(z, t) = \upsilon_{\{\mathtt{a},\mathtt{b}\}}(x, y) \\ \upsilon_{\{\mathtt{a},\mathtt{b}\}} \overset{R}{\rightharpoonup} \upsilon_{\{\mathtt{e}\}}\end{array}}{\upsilon_{\{\mathtt{e}\}}(z, t) = \upsilon_{\{\mathtt{e}\}}(x, y)}$$

On en déduit que (z, t) est en proportion analogique sur $\{\mathtt{e}\}$. L'application de la "règle de trois" nous permet de déduire que t contient $\mathtt{e}$.

Un raisonnement de type "modus ponens" est un raisonnement déductif, dans lequel une co-variation $f \overset{R}{\rightharpoonup} g$ apprise sur une partie R ne peut être appliquée qu'en un point $(z, t) \in R$. Que se passe-t-il si $(z, t) \notin R$? Peut-on s'autoriser à appliquer la règle quand même ?

1. Ici la règle $\mathbf{1}_{\leq \text{degré}} \rightharpoonup \mathbf{1}_{\leq \text{âge\_légal}}$ dit que les deux relations varient dans le même sens mais ne donne pas explicitement le sens de la co-variation. En toute rigueur, l'application de la règle d'inférence (MP) nécessite de comparer le couple $(\texttt{biere}, \texttt{whisky})$ à un ou plusieurs autres couples d'alcools.

6 Un raisonnement par similarité

Dans cette section, un raisonnement hypothétique est proposé dans lequel on s'autorise à appliquer une co-variation connue sur une partie R à des points qui ne sont pas dans R, mais qui sont d'une certaine façon similaires à ceux de R.

6.1 Principe

L'idée est de généraliser la règle d'inférence (MP) à tout couple $(z, t) \in \mathcal{X}^2$. On obtient :

$$\frac{\begin{array}{l}f(z, t) = f(x, y) \text{ pour } (x, y) \in R \\ f \overset{R}{\rightharpoonup} g\end{array}}{g(z, t) = g(x, y)}$$

Lorsque $(z, t) \in R$, on retrouve la règle (MP) définie précédemment. Mais ici on s'autorise à appliquer la règle d'inférence aux couples $(z, t) \notin R$, sur la base de la similarité entre (z, t) et certains points (x, y) de R, qui prennent la même valeur pour f.

6.2 Exemples

Dans l'exemple des alcools, supposons que l'on connaisse la règle $\mathbf{1}_{\leq \text{degré}} \overset{R}{\rightharpoonup} \mathbf{1}_{\leq \text{âge\_légal}}$ sur l'ensemble $R = \{\texttt{biere}, \texttt{whisky}\}^2$ et que l'on veuille évaluer l'âge légal de consommation du cidre. Si l'on sait que $\texttt{cidre} \leq_{\text{degré}} \texttt{biere}$, la règle d'inférence nous donne :

$$\frac{\begin{array}{l}\mathbf{1}_{\leq \text{degré}}(\texttt{cidre}, \texttt{biere}) = \mathbf{1}_{\leq \text{degré}}(\texttt{biere}, \texttt{whisky}) \\ \mathbf{1}_{\leq \text{degré}} \overset{R}{\rightharpoonup} \mathbf{1}_{\leq \text{âge\_légal}}\end{array}}{\mathbf{1}_{\leq \text{âge\_légal}}(\texttt{cidre}, \texttt{biere}) = \mathbf{1}_{\leq \text{âge\_légal}}(\texttt{biere}, \texttt{whisky})}$$

ce qui permet de faire l'hypothèse que l'âge légal de consommation du cidre est inférieur ou égal à l'âge légal de consommation de la bière.

Dans l'exemple des proportions, cette règle d'inférence nous permet d'appliquer la co-variation $\upsilon_{\{\mathtt{a},\mathtt{b}\}} \overset{R}{\rightharpoonup} \upsilon_{\{\mathtt{e}\}}$ au couple (z, t) même si cette règle n'est connue que sur une partie $R = \{(x, y)\}$ ne contenant pas (z, t). La règle d'inférence donne :

$$\frac{\begin{array}{l}\delta_{\{\mathtt{a},\mathtt{b}\}}(z, t) = \delta_{\{\mathtt{a},\mathtt{b}\}}(x, y) \\ \delta_{\{\mathtt{a},\mathtt{b}\}} \overset{R}{\rightharpoonup} \delta_{\{\mathtt{e}\}}\end{array}}{\delta_{\{\mathtt{e}\}}(z, t) = \delta_{\{\mathtt{e}\}}(x, y)}$$

ce qui permet de formuler l'hypothèse, lorsqu'on sait que t contient $\mathtt{a}$ mais non $\mathtt{b}$, que t contient $\mathtt{e}$.

6.3 Lien avec l'adaptation

La modélisation de l'adaptation présentée en Sec. 2 peut être abordée comme un raisonnement par similarité sur des variations. Dans ce cadre, la partie $R \subseteq \mathcal{X}^2$ représente un ensemble de couples de cas sources. L'élément z représente un cas source $(\texttt{srce}, \texttt{sol}(\texttt{srce}))$ remémoré et l'élément t représente le cas cible $(\texttt{cible}, \texttt{sol}(\texttt{cible}))$ dont on cherche à construire la partie solution $\texttt{sol}(\texttt{cible})$. Une règle d'adaptation $\Delta_{\texttt{pb}} \overset{R}{\rightsquigarrow} \Delta_{\texttt{sol}}$ apprise sur R relie un ensemble de variations $\Delta_{\texttt{pb}}$ entre problèmes à un ensemble de variations $\Delta_{\texttt{sol}}$ entre solutions. Le raisonnement par similarité que nous avons présenté peut permettre de construire $\texttt{sol}(\texttt{cible})$ à partir de règles d'adaptation valides sur un ensemble de cas sources qui ne contient pas $\texttt{cible}$ et de les utiliser néanmoins pour construire $\texttt{sol}(\texttt{cible})$.

Une telle approche suppose de connaître un ensemble de co-variations valides sur l'ensemble R.

7 Apprentissage de co-variations

Cette section décrit une méthode d'apprentissage de co-variations à partir de données. L'idée de la méthode est d'extraire des co-variations dans une structure de partitions, à la manière de ce qui est fait dans [4].

7.1 Structures de motifs

Dans cette section, nous rappelons certaines définitions relatives aux structures de motifs (en anglais, *pattern structure*).

Soit $(G, (D, \sqcap), \delta)$ un triplet dans lequel G est un ensemble, $(D, \sqcap)$ est un semi-treillis-infimum [2] et δ est une fonction $\delta : G \longrightarrow D$ qui à tout élément de G associe sa "description" dans D. Alors $(G, (D, \sqcap), \delta)$ est une *structure de motifs* [19]. Les éléments de D sont appelés motifs et sont ordonnés par une relation de subsomption $\sqsubseteq$: $c \sqsubseteq d$ ssi $c \sqcap d = c$. Les opérateurs de dérivation $(.)^\square$ définis par

$$A^\square = \bigsqcap_{g \in A} \delta(g) \text{ pour } A \subseteq G$$

et

$$d^\square = \{g \in G \mid d \sqsubseteq \delta(g)\} \text{ pour } d \in D$$

forment un connexion de Galois entre $(\mathscr{P}(G), \subseteq)$ et $(D, \sqsubseteq)$. Pour $A, B \subseteq G$, une *implication d'objets* $A \rightarrow B$ est vérifiée si $A^\square \sqsubseteq B^\square$.

7.2 Structures de partition

Une structure de partition [4] est une structure de motifs $(G, (D, \sqcap), \delta)$ dans laquelle l'ensemble des descriptions D est l'ensemble des partitions d'un ensemble $\mathcal{U}$ et la relation $\sqcap$ donne l'infimum de deux partitions. Cette section rappelle en quoi consiste un tel ensemble de descriptions et pourquoi on peut l'utiliser.

2. Un semi-treillis infimum est un ensemble partiellement ordonné dans lequel tout sous-ensemble fini non vide admet un infimum.

Soit $\mathscr{E}$ l'ensemble des relations d'équivalences sur un ensemble $\mathcal{U}^2$. Lorsque $\cap$ et $\cup$ dénotent respectivement les opérations d'intersection et d'union sur $\mathcal{U}^2$, on peut montrer que $(\mathscr{E}, \cap, \cup)$ forme un treillis, c'est-à-dire que tout couple de relations de $\mathscr{E}$ admet un infimum et un supremum. La notion de partition est liée à celle de relation d'équivalence. Une partition de $\mathcal{U}$ est un ensemble $P \subseteq \mathscr{P}(\mathcal{U})$ tel que $\bigcup_{p_i \in P} p_i = \mathcal{U}$ et $p_i \cap p_j = \varnothing$ pour tout i, j, $i \neq j$. L'ensemble des partitions sur $\mathcal{U}$ est en bijection avec l'ensemble des relations d'équivalences sur cet ensemble, ce qui veut dire que toute partition représente une relation d'équivalence sur $\mathcal{U}^2$. Par exemple, si $\mathcal{U} = \{1, 2, 3\}$, la partition $\{\{1, 2\}, \{3\}\}$ représente la relation $\{(1, 2), (2, 1), (1, 1), (2, 2), (3, 3)\}$.

Si D désigne l'ensemble des partitions sur $\mathcal{U}$, on peut donc définir un opérateur d'intersection $\sqcap$ et un opérateur d'union $\sqcup$ qui sont les équivalents de $\cap$ et $\cup$ pour les relations d'équivalence. Par exemple, si $\mathcal{U} = \{1, 2, 3, 4\}$,

$$\{\{1, 3\}, \{2, 4\}\} \sqcap \{\{1, 2, 3\}, \{4\}\} = \{\{1, 3\}, \{2\}, \{4\}\}$$

car

$$\{(1, 3), (2, 4)\} \cap \{(1, 2), (1, 3), (2, 3)\} = \{(1, 3)\}$$

(la réflexivité est omise ici pour des questions de lisibilité). Tout comme son pendant pour les relations, l'ensemble $(D, \sqcap, \sqcup)$ forme un treillis, donc en particulier $(D, \sqcap)$ est un semi-treillis-infimum (la relation $\sqcap$ est idempotente, associative et commutative) et peut être utilisé comme un ensemble de description d'une structure de motifs.

7.3 Structures de variation

Une *structure de variation* est une structure de partition qui associe à chaque variation une partition de $\mathcal{X}^2$ dans laquelle chaque classe regroupe les couples (x, y) qui prennent la même valeur pour cette variation.

Définition 6. *Une structure de variation sur une partie $R \subseteq \mathcal{X}^2$ est une structure de partition $(G, (D, \sqcap), \delta)$ dans laquelle :*
- *$G \subseteq \mathscr{V}(\mathcal{X}^2, \mathcal{V})$;*
- *D est l'ensemble des partitions de R ;*
- *$\sqcap$ donne l'infimum de deux partitions ;*
- *$\delta(\upsilon_i)$ est définie par la relation $\equiv_{\upsilon_i}$ suivante :*

$$(x, y) \equiv_{\upsilon_i} (z, t) \text{ ssi } \upsilon_i(x, y) = \upsilon_i(z, t)$$

Dans cette structure, les objets sont des variations $\upsilon_i \in \mathscr{V}(\mathcal{X}^2, \mathcal{V})$, donc sont des attributs de couples d'éléments de $\mathcal{X}$. À chaque variation υ_i est associée une partition $\delta(\upsilon_i)$ de R telle que deux couples (x, y) et

(z, t) sont dans la même classe s'ils prennent la même valeur pour la variation υ_i.

Dans une structure de variation, les co-variations correspondent aux implications d'objets.

Proposition 7.1. *Étant donné une structure de variation $(G, (D, \sqcap), \delta)$ sur R et $A, B \subseteq G$:*

$$A \overset{R}{\rightsquigarrow} B \text{ ssi } A^\square \sqsubseteq B^\square$$

Démonstration. Découle directement de la proposition 3 de [4], en remarquant qu'une co-variation est une relation de dépendance fonctionnelle entre attributs. $\square$

Ce résultat est intéressant car il montre qu'on peut facilement adapter des algorithmes de fouille existants pour extraire des co-variations à partir de données. En effet, tout algorithme qui extrait dans cette structure des règles de la forme $A \to AB$ avec $\{A\}^\square = \{AB\}^\square$ permet d'extraire des co-variations valides sur R.

8 Conclusion et perspectives

Dans cet article, nous avons défini une co-variation comme une dépendance fonctionnelle entre variations, ce qui nous a permis d'obtenir une règle de déduction naturelle sur les variations. Nous avons montré que cette règle peut être facilement étendue pour raisonner par similarité sur les variations, et nous avons proposé une méthode pour apprendre des co-variations à partir de données.

Les perspectives pour donner suite à ce travail sont nombreuses. La première chose à faire serait d'implémenter un algorithme d'extraction de co-variations à partir de données et de tester sur plusieurs jeux de données. En particulier, nous voudrions appliquer cet algorithme pour fouiller des données de trajectoires patient à l'hôpital, dans le but de détecter des évolutions dans les pratiques d'un service à l'autre, d'un hôpital à l'autre, ou d'une année sur l'autre. Une autre idée serait, en s'appuyant sur les travaux de [10], de travailler sur des données biologiques et de mettre au point un processus d'extraction de connaissances pour extraire des évolutions cohérentes d'expression des gènes. Par ailleurs, une idée qui ressort de ce travail est que l'adaptation s'apparenterait à un raisonnement par similarité sur des variations. Il serait intéressant de tester cette idée en mettant au point une modélisation de l'adaptation sur ce principe et en la comparant aux approches existantes.

Remerciements

L'auteur tient à remercier les relecteurs de cet article pour leurs remarques constructives.

Références

[1] M. Abraham, Dov M. Gabbay, and U. Schild. Analysis of the Talmudic Argumentum a Fortiori inference rule (kal vachomer) using Matrix Abduction. *Stud. Log.*, 92(3) :281–364, 2009.

[2] Jean-Claude Anscombre. La Théorie Des Topoï : Sémantique Ou Rhétorique ? *Hermes (Wiesb).*, 1(15) :185–198, 1995.

[3] Fadi Badra. Representing and Learning Variations. In *Int. Conf. Tools Artif. Intell.*, 2015.

[4] Jaume Baixeries, Mehdi Kaytoue, and Amedeo Napoli. Characterizing functional dependencies in formal concept analysis with pattern structures. *Ann. Math. Artif. Intell.*, 72(1-2) :129–149, 2014.

[5] Susanne Bleisch, Matt Duckham, Antony Galton, Patrick Laube, and Jarod Lyon. Mining candidate causal relationships in movement patterns. *Int. J. Geogr. Inf. Sci.*, 28(March 2015) :363–382, 2013.

[6] Bernadette Bouchon-Meunier, Anne Laurent, Marie Jeanne Lesot, and Maria Rifqi. Strengthening fuzzy gradual rules through "all the more" clauses. *2010 IEEE World Congr. Comput. Intell. WCCI 2010*, (July), 2010.

[7] Bernadette Bouchon-Meunier, Christophe Marsala, and Maria Rifqi. Fuzzy analogical model of adaptation for case-based reasoning. In *IFSA-EUSFLAT*, pages 1625–1630, 2009.

[8] Myriam Bounhas, Henri Prade, and Gilles Richard. Analogical classification : A new way to deal with examples. In *Front. Artif. Intell. Appl.*, volume 263, pages 135–140, 2014.

[9] Toon Calders, Bart Goethals, and Szymon Jaroszewicz. Mining rank-correlated sets of numerical attributes. In *Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD '06*, page 96, 2006.

[10] Victor Codocedo and Amedeo Napoli. Lattice-based biclustering using partition pattern structures. *Front. Artif. Intell. Appl.*, 263 :213–218, 2014.

[11] William Correa Beltran, Henri Prade, and Gilles Richard. Constructive Solving of Raven's IQ Tests with Analogical Proportions. *Int. J. Intell. Syst.*, 2014.

[12] M. D'Aquin, F. Badra, S. Lafrogne, J. Lieber, A. Napoli, and L. Szathmary. Case base mining for adaptation knowledge acquisition. In *IJCAI Int. Jt. Conf. Artif. Intell.*, pages 750–755, Hyderabad, India, 2007.

[13] Joaquin Derrac and Steven Schockaert. Characterising semantic relatedness using interpretable

directions in conceptual spaces. In *Front. Artif. Intell. Appl.*, volume 263, pages 243–248, 2014.

[14] Joaquín Derrac and Steven Schockaert. Inducing semantic relations from conceptual spaces : A data-driven approach to plausible reasoning. *Artif. Intell.*, 228 :66–94, 2015.

[15] Didier Dubois, Eyke Hüllermeier, and Henri Prade. Fuzzy set-based methods in instance-based reasoning. *IEEE Trans. Fuzzy Syst.*, 10(3) :322–332, 2002.

[16] Didier Dubois and Henri Prade. Gradual inference rules in approximate reasoning. *Inf. Sci. (Ny).*, 61(1-2) :103–122, 1992.

[17] Robert P W Duin and Elbieta Pekalska. The dissimilarity space : Bridging structural and statistical pattern recognition. *Pattern Recognit. Lett.*, 33(7) :826–832, 2012.

[18] Béatrice Fuchs, Jean Lieber, Alain Mille, and Amedeo Napoli. Differential adaptation : An operational approach to adaptation for solving numerical problems with CBR. *Knowledge-Based Syst.*, 68 :103–114, apr 2014.

[19] Bernhard Ganter and Sergei O. Kuznetsov. Pattern Structures and Their Projections. In *9th Int. Conf. Concept. Struct. ICCS 2001*, volume 2120, pages 129–142. Springer Berlin Heidelberg, 2001.

[20] Yichong Li, Mei Zhang, Yong Jiang, and Fan Wu. Co-variations and clustering of Chronic Disease behavioral risk factors in China : China Chronic Disease and Risk Factor Surveillance, 2007. *PLoS One*, 7(3), 2012.

[21] Sara C. Madeira and Arlindo L. Oliveira. Biclustering algorithms for biological data analysis : A survey. *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, 1(1) :24–45, 2004.

[22] Laurent Miclet, Henri Prade, and David Guennec. Looking for analogical proportions in a formal concept analysis setting. In *Concept Lattices Appl.*, volume 959, pages 295–307. CEUR-WS, 2011.

[23] Alain Mille. Des traces à l'ère du Web. *Intellectica*, 59(59) :7–28, 2013.

[24] Sanja Petrovic, Gulmira Khussainova, and Rupa Jagannathan. Knowledge-light adaptation approaches in case-based reasoning for radiotherapy treatment planning. *Artif. Intell. Med.*, pages 1–12, 2016.

[25] Henri Prade and Gilles Richard. Reasoning with Logical Proportions. In *Int. Conf. Princ. Knowl. Represent. Reason.*, pages 545–555, 2010.

[26] P. Y. Raccah. *Topoi et gestion de connaissances.* Masson, 1996.

[27] Karima Sedki and Louis Bonneau De Beaufort. Cognitive Maps and Bayesian Networks for Knowledge Representation and Reasoning. 2012.

[28] Mathieu Serrurier, Didier Dubois, Henri Prade, and Thomas Sudkamp. Learning fuzzy rules with their implication operators. *Data Knowl. Eng.*, 60(1) :71–89, 2007.