Reasoning with Co-Variations

Fadi Badra

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

Abstract. Adaptation is what allows a system to maintain consistent behavior across variations in operating environments. 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.

Keywords: co-variations · adaptation · analogical proportions

1 Introduction

Adaptation is what " allows a system to maintain consistent behavior across variations in operating environments "[9]. A variation represents a set of differences between two or more states of affairs. These differences can be "all or nothing", *i.e.*, express the gain, preservation, or loss of a property, (*e.g.*, "false morels are toxic whereas true morels are edible"), or express a change of degree (*e.g.*, "the movie *Mad Max* is more violent than the movie *Cinderella*"). In some previous work [3], we proposed to represent variations as an attribute of an ordered set of objects (at least two), all taken in a same set. The question addressed here is the formalization of the co-occurrences that may exist between two or more variations, and the definition of "rules of inference" that could be used to infer a variation from another. In particular, we would like to define a "modus ponens" inference rule, which may be summarized schematically (for pairs of objects) as follows:

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

Such deductive reasoning enables to determine the value of a variation from another by applying a "IF ... THEN" rule. This rule must express the fact that the variation g coïncides locally with the variation f on a set of pairs of objects that contains the pair (z,t).

In this article, we propose to define co-variations as functional dependencies between variations. This definition enables to define a natural deduction rule on variations, which can be easily extended to define a similarity-based reasoning. A method is proposed to learn co-variations from data. In this method, the co-variations correspond to object implication rules in a pattern structure.

The paper is organized as follows. The next section reviews the literature. Sec. 3 recalls some definitions about variations. In Sec. 4, a co-variation is defined as a functional dependency on variations. A natural deduction scheme is defined in Sec. 5, and extended in Sec. 6 to a similarity-based inference. Sec. 7 presents a method to learn co-variations from data. Sec. 8 concludes and gives future work.

2 Review of Literature

One motivation of this work is the role variations play in the adaptation step in case-based reasoning. In one of its formulations [10], adaptation is presented as the construction of a solution sol(tgt) of a target problem tgt by modifying the solution sol(srce) of a retrieved source problem srce. Adaptation can be decomposed in three steps:

① $(\texttt{srce}, \texttt{tgt}) \mapsto \Delta_{\texttt{pb}}$: the differences between the two problems srce and tgt are represented;
② $(\Delta_{\texttt{pb}}, \texttt{AK}) \mapsto \Delta_{\texttt{sol}}$: some adaptation knowledge AK is used to construct a variation $\Delta_{\texttt{sol}}$ between sol(srce) and the (future) sol(tgt) ;
③ $(\Delta_{\texttt{sol}}, \texttt{sol(srce)}) \mapsto \texttt{sol(tgt)}$: sol(srce) is modified into sol(tgt) by applying $\Delta_{\texttt{sol}}$.

According to this model, performing adaptation requires to be able to infer variations $\Delta_{\texttt{sol}}$ between solutions from variations $\Delta_{\texttt{pb}}$ between problems (step ②).

This paper proposes a symbolic representation of co-variations, and as such can be contrasted with the statistical or graphical approaches often used for correlation detection (see for example [8, 16, 19]). A co-variation expresses a local coïncidence of values of two properties, each of which apply to an ordered set of objects. The notion of co-variation is therefore close to the notion of analogical dissimilarity, which is measured in [7] by counting the number of flips necessary to turn four objects into an analogical proportion, and in [11] by taking the cosine of two vectors in the euclidian space $\mathbb{R}^n$. When variations represent a change in degree of a property, co-variations express monotone correspondences between gradual changes, through rules of the form "the more x is A, the more y is B". Such correspondences are known in language semantics as argumentative *topoï*, and are defined as pairs of gradual predicates, along with the set of monotone correspondences between these gradations [2]. Gradual inference rules [5, 14] may be seen as a numerical modeling of such correspondences using fuzzy logic techniques. These rules have been applied to similarity-based reasoning [13], and even to the modeling of adaptation [6]. However, their semantics is different from the one presented here. A gradual inference rule models uncertainty, whereas the semantics chosen here for co-variations rather models co-occurrence, in the spirit of association rules. Regarding co-variation learning, an algorithm is proposed in [20] to extract gradual inference rules using inductive logic programming techniques. A method is proposed in [17] to learn analogical proportions from formal contexts by reducing iteratively the analogical dissimilarity between pairs of objects.

3 Variations

This section recalls some definitions about variations.

A *variation* is modelled by a function $f : \mathcal{X}^n \longrightarrow \mathcal{V}$ which associates a value taken in a set $\mathcal{V}$ to the elements of the cartesian product $\mathcal{X}^n$. In the following, we'll assume that $n = 2$, so that variations are attributes of pairs of elements of $\mathcal{X}$. The set of all variations $f : \mathcal{X}^2 \longrightarrow \mathcal{V}$ defined on a set $\mathcal{X}^2$ and with values in $\mathcal{V}$ is denoted by $\mathscr{V}(\mathcal{X}^2, \mathcal{V})$.

Variations and Binary Relations. When $\mathcal{V} = \{0, 1\}$, the set $\mathscr{V}(\mathcal{X}^2, \{0, 1\})$ denotes the indicator functions of binary relations on $\mathcal{X}^2$. For example, if $\mathcal{X} = \mathbb{N}$ is the set of natural numbers, one can define the variation:

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

that returns 1 if x is lower or equal than y and 0 otherwise.

Variations between Sets of Binary Attributes. A special case is when $\mathcal{X} = \mathscr{P}(\mathtt{M})$ denotes the powerset of a set $\mathtt{M}$ of binary attributes. For example, let $\mathtt{M} = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{e}\}$ denote a set of binary attributes, and $x = \{\mathtt{a}, \mathtt{b}, \mathtt{d}\}$ and $y = \{\mathtt{a}, \mathtt{e}\}$ be two sets of attributes of $\mathtt{M}$. A natural way to represent the variations between x and y is to introduce the four sets $x \cap y$, $x \cap \overline{y}$, $\overline{x} \cap y$, and $\overline{x} \cap \overline{y}$, which together form a partition of $\mathtt{M}$. In our example, we get:

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

When x, y, z, and t represent sets of binary attributes, the analogical proportion $x : y :: z : t$ is defined [18] by $x \cap \overline{y} = z \cap \overline{t}$ and $\overline{x} \cap y = \overline{z} \cap t$, which means that the analogical proportion $x : y :: z : t$ holds for two pairs (x,y) and (z,t) if and only if the two pairs take the same value for the variation:

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

If the three sets x, y, and z are known, solving the analogical proportion equation enables to determine the set t. For example, if $x = \{\mathtt{a}, \mathtt{b}, \mathtt{d}\}$, $y = \{\mathtt{a}, \mathtt{e}\}$, and $z = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, then $\upsilon(x, y) = \{\{\mathtt{b}, \mathtt{d}\}, \{\mathtt{e}\}\}$ and all sets of attributes t that verify $\upsilon(z, t) = \upsilon(x, y)$ are in analogical proportion with x, y, and z. Unfortunately, in this example, the equation has no solution (Fig. 1) since z does not contain the attribute 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

Fig. 1: Resolution of the equation $x : y :: z : t = 1$.

4 Co-Variations

In this section, co-variations are defined as functional dependencies between two (sets of) variations. Let us first define the notion of co-variation between two single variations.

Definition 1. *Let $f, g \in \mathscr{V}(\mathcal{X}^2, \mathcal{V})$ be two variations. A variation g co-varies with a variation f on a subset R of $\mathcal{X}^2$, denoted by $f \overset{R}{\rightharpoondown} g$, iff for all (x, y) and (z, t) of R:*

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

This definition expresses the fact that whenever two elements of the subset R of $\mathcal{X}^2$ share a same value for the variation f, they must also share a same value for g. If $R = \mathcal{X}^2$, the set R can be omited, and we will write $f \rightharpoondown g$ for $f \overset{\mathcal{X}^2}{\rightharpoondown} g$.

This definition can be extended to the co-variation between two sets of variations.

Definition 2. *Let $F \subseteq \mathscr{V}(\mathcal{X}^2, \mathcal{V})$ and $G \subseteq \mathscr{V}(\mathcal{X}^2, \mathcal{V})$ be two sets of variations. The set G co-varies with the set F on a subset R of $\mathcal{X}^2$, denoted by $F \overset{R}{\rightharpoondown} G$, iff for all (x, y) and (z, t) of R :*

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

This definition expresses the fact that each subset of R that shares the same value for all variations of F must also share a same value for all variations of G. Similarly, we will write $F \rightharpoondown G$ for $F \overset{\mathcal{X}^2}{\rightharpoondown} G$.

Co-Variations and Analogical Proportions. Let us assume that $\mathcal{X} = \mathscr{P}(\mathtt{M})$ denotes the powerset of a set $\mathtt{M}$ of binary attributes. Let us consider the variations $\upsilon_{\mathtt{T}}$ defined for $\mathtt{T} \subseteq \mathtt{M}$ by: $\upsilon_{\mathtt{T}}(x, y) = \{x \cap \overline{y} \cap \mathtt{T}\,,\, \overline{x} \cap y \cap \mathtt{T}\}$. The variations $\upsilon_{\mathtt{T}}$ represent the differences that exist between two sets of attributes, but on a subset $\mathtt{T} \subseteq \mathtt{M}$ only. They take the same value for all pairs (x, y) that are in analogical proportion on the subset $\mathtt{T}$. One can write the co-variations $\upsilon_{\mathtt{S}} \overset{R}{\rightharpoondown} \upsilon_{\mathtt{T}}$. These co-variations express the fact that every subset of R that is in analogical proportion on $\mathtt{S}$ (*i.e.*, for which the variation $\upsilon_{\mathtt{S}}$ takes the same value) is also in analogical proportion on $\mathtt{T}$. More formally:

Definition 3. *Let $R \subseteq \mathcal{X}^2$ and $\mathtt{T} \subseteq \mathtt{M}$. The set R is in analogical proportion on $\mathtt{T}$ iff:*

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

A subset R of $\mathcal{X}^2$ is in analogical proportion on a set of attributes $\mathtt{T}$ if the variation $\upsilon_{\mathtt{T}}$ takes the same value on R. Here, $\mathscr{P}(\mathtt{M})$ denotes the powerset of $\mathtt{M}$.

Proposition 1. *The co-variation $\upsilon_{\mathtt{S}} \overset{R}{\rightharpoondown} \upsilon_{\mathtt{T}}$ holds iff every subset of R which is in analogical proportion on $\mathtt{S}$ is also in analogical proportion on $\mathtt{T}$.*

Proof. ⇒: Assume that $\upsilon_{\mathtt{S}} \overset{R}{\rightharpoondown} \upsilon_{\mathtt{T}}$ and that $A \subseteq R$ is in analogical proportion on $\mathtt{S}$. Then, (Def. 3) there exists a v such that $\forall (x, y) \in R$, $\upsilon_{\mathtt{S}}(x, y) = v$. Assume that there exists $(x, y), (z, t) \in R$ such that $\upsilon_{\mathtt{T}}(x, y) \neq \upsilon_{\mathtt{T}}(z, t)$. We have $\upsilon_{\mathtt{S}}(x, y) = \upsilon_{\mathtt{S}}(z, t) = v$ so by Def. 1, $\upsilon_{\mathtt{T}}(x, y) = \upsilon_{\mathtt{T}}(z, t)$. Contradiction.
⇐: Let $(x, y), (z, t) \in R$ be such that $\upsilon_{\mathtt{S}}(x, y) = \upsilon_{\mathtt{S}}(z, t)$. The set $\{(x, y), (z, t)\} \subseteq R$ is in analogical proportion on $\mathtt{S}$ and therefore on $\mathtt{T}$. As a result, (Def. 3) there exists a v such that $\upsilon_{\mathtt{T}}(x, y) = \upsilon_{\mathtt{T}}(z, t) = v$. □

5 A Rule-Based Inference

In this section, co-variations are used in a rule-based inference to predict the value of a variation from the values of one or many other variations.

Definition 4. *Let $f, g \in \mathscr{V}(\mathcal{X}^2, \mathcal{V})$ be two variations. The "modus ponens" inference rule on variations is as follows:*

$$\frac{f(x, y) = f(z, t) \text{ for } (x, y), (z, t) \in R \qquad f \overset{R}{\rightharpoondown} g}{g(x, y) = g(z, t)} \tag{MP}$$

This rule states that knowing that g co-varies with f on a subset R, if we know that two pairs (x, y) and (z, t) take the same value for f, then we can infer that they also take the same value for g.

This rule can be extended to co-variations $F \overset{R}{\rightharpoondown} G$ between two sets of variations F and G.

Definition 5. *Let $F \subseteq \mathscr{V}(\mathcal{X}^2, \mathcal{V})$ and $G \subseteq \mathscr{V}(\mathcal{X}^2, \mathcal{V})$ be two sets of variations, $f \in F$ and $g \in G$. The "modus ponens" inference rule on sets of variations is as follows:*

$$\frac{\forall f \in F,\ f(x, y) = f(z, t) \text{ for } (x, y), (z, t) \in R \qquad F \overset{R}{\rightharpoondown} G}{\forall g \in G,\ g(x, y) = g(z, t)}$$

Example #1: the A Fortiori *Inference.* When the two variations f and g are indicator functions of partial orders, the modus ponens inference corresponds to an *a fortiori* inference [1]. This type of inference exploits the monotony of two partial orders to estimate the value of an attribute. The authors of [12] give the following example. If we know that whiskey is stronger than beer, and that buying beer is illegal under the age of 18, then we can plausibly derive that buying whiskey is illegal under the age of 18. Let us call $\mathcal{A}$ the set of alcohols, and $\leq_{\text{degree}}$ and $\leq_{\text{legal\_age}}$ the partial orders on $\mathcal{A}$ that order the alcohols respectively on their degree and minimum legal age. The example can be interpreted as the co-variation $\mathbf{1}_{\leq \text{ degree}} \rightharpoondown \mathbf{1}_{\leq \text{ legal\_age}}$ between the two variations $\mathbf{1}_{\leq \text{ degree}}$ and

$\mathbf{1}_{\leq \text{legal\_age}}$ of $\mathscr{V}(\mathcal{A}^2, \{0,1\})$ which represent respectively the indicator functions of the two relations $\leq_{\text{degree}}$ and $\leq_{\text{legal\_age}}$. The inference rule (MP) gives[1]:

$$\frac{\begin{array}{l}\mathbf{1}_{\leq \text{degree}}(\texttt{beer}, \texttt{whiskey}) = 1 \\ \mathbf{1}_{\leq \text{degree}} \rightharpoondown \mathbf{1}_{\leq \text{legal\_age}}\end{array}}{\mathbf{1}_{\leq \text{legal\_age}}(\texttt{beer}, \texttt{whiskey}) = 1}$$

and we conclude that the minimum legal age to drink whiskey is at least equal to the minimum legal age to drink beer.

Example #2: Analogical Proportions. Let us assume that $\texttt{M} = \{\texttt{a}, \texttt{b}, \texttt{c}, \texttt{d}, \texttt{e}\}$ is a set of binary attributes. For $\texttt{T} \subseteq \texttt{M}$, the variations $\upsilon_{\texttt{T}}$ are the ones defined in Sec. 4. Assume that we know the rule $\upsilon_{\{\texttt{a},\texttt{b}\}} \overset{R}{\rightharpoondown} \upsilon_{\{\texttt{e}\}}$ on the set $R = \{(x, y), (z, t)\}$. This rule says that every pair of R that is in analogical proportion on $\{\texttt{a}, \texttt{b}\}$ is also in analogical proportion on $\{\texttt{e}\}$. If we have, for example, $x = \{\texttt{a}, \texttt{b}, \texttt{d}\}$, $y = \{\texttt{a}, \texttt{e}\}$, and $z = \{\texttt{a}, \texttt{b}, \texttt{c}\}$, and we know that t contains $\texttt{a}$ and not $\texttt{b}$, then the inference rule (MP) gives:

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

So, we can deduce that (z, t) is in analogical proportion on $\{\texttt{e}\}$. By solving the analogical proportion equation, we deduce that t contains $\texttt{e}$.

The "modus ponens" inference is a kind of deductive reasoning, in which a co-variation $f \overset{R}{\rightharpoondown} g$ learned on R can only be applied to a point (z, t) of R. What happens if $(z, t) \notin R$? Can we still apply the rule? Under what conditions?

6 A Similarity-Based Inference

This section defines a hypothetical reasoning scheme, in which co-variations may be applied to some points that are outside of their known domain R of validity.

The idea of the method is to generalize the inference rule (MP) to any pair $(z, t) \in \mathcal{X}^2$. This leads to the following schema:

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

When $(z, t) \in R$, this schema corresponds to previously defined inference rule (MP). But here, the inference rule may be applied to pairs $(z, t) \notin R$, on the basis of the similarity between (z, t) and some points (x, y) of R, which take the same value for f.

[1] Here the rule $\mathbf{1}_{\leq \text{degree}} \rightharpoondown \mathbf{1}_{\leq \text{legal\_age}}$ says that the two relations vary in the same direction but does not explicitely give the direction of the co-variation. So a more rigorous application of the inference rule (MP) would require to compare the pair $(\texttt{beer}, \texttt{whiskey})$ with one or many other pairs of alcohols.

Example #1: the A Fortiori *Inference.* In the example of the alcohols, suppose that the rule $\mathbf{1}_{\leq\text{ degree}} \overset{R}{\rightsquigarrow} \mathbf{1}_{\leq\text{ legal\_age}}$ is known on the set $R = \{\texttt{beer}, \texttt{whiskey}\}^2$ and that we want to estimate the minimum legal age for cider consumption. If we know that $\texttt{cider} \leq_{\text{degree}} \texttt{beer}$, the inference rule gives:

$$\frac{\begin{array}{l}\mathbf{1}_{\leq\text{ degree}}(\texttt{cider}, \texttt{beer}) = \mathbf{1}_{\leq\text{ degree}}(\texttt{beer}, \texttt{whiskey}) \\ \mathbf{1}_{\leq\text{ degree}} \overset{R}{\rightsquigarrow} \mathbf{1}_{\leq\text{ legal\_age}}\end{array}}{\mathbf{1}_{\leq\text{ legal\_age}}(\texttt{cider}, \texttt{beer}) = \mathbf{1}_{\leq\text{ legal\_age}}(\texttt{beer}, \texttt{whiskey})}$$

which makes it possible to formulate the hypothesis that the minimum legal age for cider consumption is lower or equal than the minimum legal age for beer consumption.

Example #2: Analogical Proportions. In the example of analogical proportions, this inference rule allows to apply the co-variation $\upsilon_{\{\texttt{a},\texttt{b}\}} \overset{R}{\rightsquigarrow} \upsilon_{\{\texttt{e}\}}$ to the pair (z, t) even though it is known only on a subset $R = \{(x, y)\}$ that does not contain the pair (z, t). The inference rule gives:

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

which makes it possible to formulate the hypothesis that t contains $\texttt{e}$.

Co-Variations and Adaptation. The modelling of adaptation presented in Sec. 2 can be formulated as a similarity-based reasoning on variations. The set $R \subseteq \mathcal{X}^2$ represents a set of pairs of source cases. The element z represents a retrieved source case $(\texttt{srce}, \texttt{sol}(\texttt{srce}))$ and the element t represents the target case $(\texttt{tgt}, \texttt{sol}(\texttt{tgt}))$ for which we want to construct the solution part $\texttt{sol}(\texttt{tgt})$. An adaptation rule $\Delta_{\texttt{pb}} \overset{R}{\rightsquigarrow} \Delta_{\texttt{sol}}$ learned on R relates a set $\Delta_{\texttt{pb}}$ of variations between problems to a set $\Delta_{\texttt{sol}}$ of variations between solutions. The similarity-based reasoning that we have presented can be used to construct $\texttt{sol}(\texttt{tgt})$ from adaptation rules that are known on a set of source cases that do not contain $\texttt{tgt}$, and still use them to construct $\texttt{sol}(\texttt{tgt})$.

Such an approach requires to have acquired some co-variations, along with their domains R of validity. The next section proposes a method to learn co-variations from data.

7 Learning Co-Variations

This section describes a method to learn co-variations from data. The idea of the method is to extract co-variations from a partition pattern structure, in the spirit of what is done in [4].

Pattern Structures. Let G be a set of objects, $(D, \sqcap)$ a meet-semilattice[2], and δ a mapping $\delta : G \longrightarrow D$ that associates to each element of G its "description" in D. Then, $(G, (D, \sqcap), \delta)$ is a *pattern structure* [15]. The elements of D are called patterns and are ordered by a subsumption relation $\sqsubseteq$: $c \sqsubseteq d$ iff $c \sqcap d = c$. The derivation operators $(.)^\square$ defined by:

$$A^\square = \bigsqcap_{g \in A} \delta(g) \text{ for } A \subseteq G \quad \text{and} \quad d^\square = \{g \in G \mid d \sqsubseteq \delta(g)\} \text{ for } d \in D$$

form a Galois connection between $(\mathscr{P}(G), \subseteq)$ and $(D, \sqsubseteq)$. For $A, B \subseteq G$, an *object implication* $A \rightarrow B$ holds if $A^\square \sqsubseteq B^\square$.

Partition Structures. A *partition structure* [4] is a pattern structure $(G, (D, \sqcap), \delta)$ in which the set of descriptions D is the set of partitions of a set $\mathcal{U}$, and the relation $\sqcap$ gives the meet of two partitions. Let $\mathscr{E}$ denote the set of equivalence relations on a set $\mathcal{U}^2$, and $\cap$, $\cup$ denote respectively the intersection and union operation on $\mathcal{U}^2$. It can be shown that $(\mathscr{E}, \cap, \cup)$ forms a lattice, *i.e.*, every pair of equivalence relations of $\mathscr{E}$ has an infimum and a supremum. There is a one-to-one correspondence between the set $\mathscr{E}$ of equivalence relations on $\mathcal{U}^2$ and the set of partitions of $\mathcal{U}$. A partition of $\mathcal{U}$ is a set $P \subseteq \mathscr{P}(\mathcal{U})$ such that $\bigcup_{p_i \in P} p_i = \mathcal{U}$ and $p_i \cap p_j = \varnothing$ for all i, j, $i \neq j$. For example, if $\mathcal{U} = \{1, 2, 3\}$, the partition $\{\{1, 2\}, \{3\}\}$ represents the relation $\{(1, 2), (2, 1), (1, 1), (2, 2), (3, 3)\}$. Let D be the set of partitions of $\mathcal{U}$. An intersection operator $\sqcap$ and an union operator $\sqcup$ can be defined that correspond to the $\cap$ and $\cup$ operators on equivalence relations. For example, if $\mathcal{U} = \{1, 2, 3, 4\}$, $\{\{1, 3\}, \{2, 4\}\} \sqcap \{\{1, 2, 3\}, \{4\}\} = \{\{1, 3\}, \{2\}, \{4\}\}$ since $\{(1, 3), (2, 4)\} \cap \{(1, 2), (1, 3), (2, 3)\} = \{(1, 3)\}$ (reflexivity is omitted here for the sake of readability). As its relational counterpart, the set $(D, \sqcap, \sqcup)$ forms a lattice, and $(D, \sqcap)$ is a meet-semilattice ($\sqcap$ is idempotent, associative, and commutative) so it can be used as a set of descriptions in a pattern structure.

Variation Structures. A *variation structure* is a partition pattern structure that associates to each variation a partition of $\mathcal{X}^2$ in which each class groups the pairs (x, y) that take the same value for this variation.

Definition 6. *A variation structure on a subset $R \subseteq \mathcal{X}^2$ is a partition structure $(G, (D, \sqcap), \delta)$ such that:*

- *$G \subseteq \mathscr{V}(\mathcal{X}^2, \mathcal{V})$;*
- *D is the set of partitions of R;*
- *$\sqcap$ gives the meet of two partitions;*
- *$\delta(\upsilon_i)$ is defined by the following $\equiv_{\upsilon_i}$ equivalence relation:*

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

[2] A meet-semilattice is a partially ordered set which has a greatest lower bound for any non-empty finite subset.

In this structure, the objects are variations $\upsilon_i \in \mathscr{V}(\mathcal{X}^2, \mathcal{V})$ that can be seen as attributes of pairs of elements of $\mathcal{X}$. To each variation υ_i is associated a partition $\delta(\upsilon_i)$ of R such that two pairs (x, y) and (z, t) are in the same class if they take the same value for the variation υ_i.

Co-variations on R correspond to object implications in this structure.

Proposition 2. *Let $(G, (D, \sqcap), \delta)$ be a variation structure on R and $A, B \subseteq G$:*

$$A \overset{R}{\rightsquigarrow} B \text{ iff } A^\Box \sqsubseteq B^\Box$$

Proof. $\Rightarrow$: Assume two pairs $(x, y), (z, t) \in \mathcal{X}^2$ are in the same equivalence class of the partition $A^\Box$. Then, $\forall f \in A$, $f(x, y) = f(z, t)$ (Def. 6). By definition of the co-variation $A \rightsquigarrow B$ (Def. 2), we have that $\forall g \in B$, $g(x, y) = g(z, t)$, which means that $(x, y) \equiv_g (z, t)$ (Def. 6) for all $g \in B$, so the two pairs $(x, y), (z, t)$ are in the same equivalence class for the partition $B^\Box$.
$\Leftarrow$: Let two pairs $(x, y), (z, t) \in \mathcal{X}^2$ be such that $\forall f \in A$, $f(x, y) = f(z, t)$. Then, $(x, y) \equiv_f (z, t)$ (Def. 6) for all $f \in A$, and as $A^\Box \sqsubseteq B^\Box$, two pairs $(x, y), (z, t)$ that are in the same equivalence class for the partition $A^\Box$ are also in the same class for the partition $B^\Box$. Thus, $\forall g \in B$, $g(x, y) = g(z, t)$ (Def. 6). □

This result is interesting because it shows that co-variations can be extracted from data by adapting existing pattern mining algorithms.

8 Conclusion and Future Work

In this article, co-variations are defined as functional dependencies between variations, which enables to come up with a natural deduction rule on variations. We showed that the natural "modus ponens" inference on variations can be easily extended to perform similarity-based reasoning on variations, and a method was proposed to learn co-variations from data.

Future work include implementing an algorithm to extract co-variations from data and testing on various data sets. In particular, we would like to apply such an algorithm to mine patient trajectories of patients in french hospitals, in order to explain how health care practices evolve over time or vary between two care units of a same hospital, or between two hospitals. Besides, an idea that this work suggests is that adaptation in case-based reasoning can be viewed as a similarity-based reasoning on variations. It would be interesting to further develop this idea and to compare a modeling of adaptation based on this principle to existing approaches.

Aknowledgements. The author wishes to thank the reviewers for their constructive remarks.

References

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

2. Anscombre, J.C.: La Théorie Des Topoï: Sémantique Ou Rhétorique ? Hermes (Wiesb). 1(15), 185–198 (1995)
3. Badra, F.: Representing and Learning Variations. In: Int. Conf. Tools Artif. Intell. pp. 950–957. IEEE, Vietri sul Mare (2015)
4. Baixeries, J., Kaytoue, M., Napoli, A.: Characterizing functional dependencies in formal concept analysis with pattern structures. Ann. Math. Artif. Intell. 72(1-2), 129–149 (2014)
5. Bouchon-Meunier, B., Laurent, A., Lesot, M.J., Rifqi, M.: Strengthening fuzzy gradual rules through "all the more" clauses. In: FUZZ-IEEE 2010, IEEE Int. Conf. Fuzzy Syst. pp. 1–7. Barcelona, Spain (2010)
6. Bouchon-Meunier, B., Marsala, C., Rifqi, M.: Fuzzy analogical model of adaptation for case-based reasoning. In: IFSA-EUSFLAT. pp. 1625–1630 (2009)
7. Bounhas, M., Prade, H., Richard, G.: Analogical classification: A new way to deal with examples. In: Front. Artif. Intell. Appl. vol. 263, pp. 135–140 (2014)
8. Calders, T., Goethals, B., Jaroszewicz, S.: Mining rank-correlated sets of numerical attributes. In: Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD '06. p. 96 (2006)
9. Coutaz, J., Crowley, J.L., Dobson, S., Garlan, D.: Context is Key. Commun. ACM 48(3), 49–53 (2005)
10. D'Aquin, M., Badra, F., Lafrogne, S., Lieber, J., Napoli, A., Szathmary, L.: Case base mining for adaptation knowledge acquisition. In: IJCAI Int. Jt. Conf. Artif. Intell. pp. 750–755. Hyderabad, India (2007)
11. Derrac, J., Schockaert, S.: Characterising semantic relatedness using interpretable directions in conceptual spaces. In: Front. Artif. Intell. Appl. vol. 263, pp. 243–248 (2014)
12. Derrac, J., Schockaert, S.: Inducing semantic relations from conceptual spaces: A data-driven approach to plausible reasoning. Artif. Intell. 228, 66–94 (2015)
13. Dubois, D., Hüllermeier, E., Prade, H.: Fuzzy set-based methods in instance-based reasoning. IEEE Trans. Fuzzy Syst. 10(3), 322–332 (2002)
14. Dubois, D., Prade, H.: Gradual inference rules in approximate reasoning. Inf. Sci. (Ny). 61(1-2), 103–122 (1992)
15. Ganter, B., Kuznetsov, S.O.: Pattern Structures and Their Projections. In: 9th Int. Conf. Concept. Struct. ICCS 2001, Lect. Notes Comput. Sci. vol. 2120, pp. 129–142. Springer (2001)
16. Li, Y., Zhang, M., Jiang, Y., Wu, F.: 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)
17. Miclet, L., Prade, H., Guennec, D.: Looking for analogical proportions in a formal concept analysis setting. In: Concept Lattices Appl. vol. 959, pp. 295–307. CEUR-WS (2011)
18. Prade, H., Richard, G.: Reasoning with Logical Proportions. In: Princ. Knowl. Represent. Reason. Proc. Twelfth Int. Conf. pp. 545–555. Toronto, Canada (2010)
19. Sedki, K., Beaufort, L.B.D.: Cognitive Maps and Bayesian Networks for Knowledge Representation and Reasoning. In: IEEE 24th Int. Conf. Tools with Artif. Intell. ICTAI 2012. pp. 1035–1040. Athens, Greece (2012)
20. Serrurier, M., Dubois, D., Prade, H., Sudkamp, T.: Learning fuzzy rules with their implication operators. Data Knowl. Eng. 60(1), 71–89 (2007)