A Proposition for Combining Pattern Structures and Relational Concept Analysis

Victor Codocedo, Amedeo Napoli

HAL Id: hal-01095870
https://hal.inria.fr/hal-01095870

Submitted on 16 Dec 2014

A proposition for combining pattern structures and relational concept analysis

Víctor Codocedo and Amedeo Napoli

LORIA - CNRS - INRIA - Université de Lorraine, BP 239, 54506 Vandœuvre-les-Nancy.
victor.codocedo@loria.fr, amedeo.napoli@loria.fr,

Abstract. In this paper we propose an adaptation of the RCA process enabling the relational scaling of pattern structures. In a nutshell, this adaptation allows the scenario where RCA needs to be applied in a relational context family composed by pattern structures instead of formal contexts. To achieve this we define the heterogeneous pattern structures as a model to describe objects in a combination of spaces, namely the original object description space and the set of relational attributes derived from the RCA scaling process. We frame our approach in the problem of characterizing latent variables (LV) in a latent variable model of documents and terms. LVs are used as compact and improved dataset representations. We approach the problem of LV characterization missing from the original LV-model, through the application of the adapted RCA process using pattern structures. Finally, we discuss the implications of our proposition.

1 Introduction

Relational Concept Analysis (RCA) [10] is an extension of Formal Concept Analysis (FCA) [4] based on a scaling process. RCA enables the application of FCA algorithms over a relational context family (RCF) which models the situation where different object sets, in different formal contexts ($\mathcal{K}_1$ and $\mathcal{K}_2$) are associated by a binary relation $r \subseteq G_1 \times G_2$ (e.g. people and their professions liking movies with different genres). In this paper we present an adaptation of RCA which enables its application when one of the object sets cannot be described by set of attributes as usual, but rather by complex descriptions (and thus calling for pattern structures [3] for taking into account these complex descriptions). Particularly, we consider the case when the "domain context" (i.e. the context where the object set is the domain of relation r) is a pattern structure of the form $\mathcal{K}_1 = (G_1, (D, \sqcap), \delta)$. To achieve this adaptation, we define the heterogeneous pattern structures as a mean to provide an object with descriptions in different spaces of data, to support both, its original pattern structure description and the relational scaling proposed in RCA.

The inspiration of this problem comes from a model known in information retrieval as "latent variable models" (LV-models), sometimes called "topic models" [11]. LV-models are a long used, cutting-edge and useful manner to index, cluster and retrieve documents [2]. They share the basic notion that the information in a document collection is "generated" by a reduced set of latent variables (LVs) hidden in data, i.e. terms in a given document are a manifestation of topics or LVs (e.g. in an article about "formal

concept analysis", the terms "formal context" and "concept lattice" are expected to be mentioned).

Latent variables, however, are abstractions. While they may represent topics, those topics lack a proper characterization, which makes difficult their interpretation. For example, in the case of latent semantic indexing (LSI) [2] (considered to be seminal work in topic models), LVs are represented by eigenvectors of a document-term matrix. Nevertheless, eigenvectors or convex regions in the eigenvector space (usually called "clusters") can be hardly recognizable as being, for instance, the topic of "formal concept analysis". Usually, we can try to manually recognize the documents and terms in a cluster to give it a "label", however this can be expensive and tedious. Moreover, LV-models do not allow the incorporation of external knowledge sources which could aid the "labelling" task.

Given the capabilities of FCA for classification and the extent/intent representation of concepts, LVs' characterization can be achieved by constructing a RCF containing a context of document descriptions in the latent variable space (a pattern structure), a formal context for terms' annotations from Wordnet[1] (e.g. a "lattice" *is a* "structure"[2]), and a relational context between documents and terms representing the binary relation document *contains* term. Accordingly, a key aspect of this work is to address the issue that relational scaling is not currently supported for pattern structures.

The main contributions of this work are the proposition of a coherent combination of pattern structures and RCA, the resulting description of heterogeneous pattern structures and a characterization technique for latent variables in a LV-model. The remainder of this paper is as follows. Section 2 provides the theoretical background of this work by describing the RCA process and the pattern structure framework. In Section 3 we describe the latent variable characterization problem in the context of the LSI technique and provide the problem statements. Section 4 describes our proposal for a pattern structures-RCA combination and defines the heterogeneous pattern structures framework. Finally, Section 5 answers both questions and discusses their implications, while providing the conclusions for this work.

2 Theoretical framework

In the following, we define the basic notions which support our approach. The examples in this section are illustrative for RCA and pattern structures, respectively, however they do not represent our scenario which is actually introduced in the next section.

2.1 Relational Concept Analysis (RCA)

Hereafter, we briefly introduce the mechanism of RCA as detailed in [9,10]. Different from standard FCA, RCA considers the scenario where an object has not only attributes, but also relations with other objects which have attributes of their own. For example,

---

[1] http://wordnetweb.princeton.edu - Wordnet is an open lexical hierarchy available online.

[2] Hypernym of "lattice".

(a)

	author1	author2	author3	author4
g_1		×	×	
g_2	×	×		
g_3	×	×		
g_4	×	×		×
g_5			×	
g_6				×
g_7			×	×
g_8			×	
g_9	×			×

(b)

	Person	Surgery	Illness	Artefact	Event	Activity
patient	×					
laparoscopy		×				×
scan				×		
user	×					
medicine						×
response					×	
time					×	
MRI				×		
practice						×
complication			×			
arthroscopy		×				×
infection			×			

(c)

	patient	laparoscopy	scan	user	medicine	response	time	MRI	practice	complication	arthroscopy	infection
g_1	×	×	×						×			
g_2		×	×	×	×	×			×			
g_3		×			×	×		×				
g_4	×				×			×				
g_5		×			×	×						
g_6									×		×	
g_7										×	×	
g_8										×	×	×
g_9											×	×

Table 1: Relational context family (RCF) - Table 1a: Formal context $\mathcal{K}_1$ of documents and their authors. Table 1b: Formal context $\mathcal{K}_2$ of terms and their Wordnet annotations. Table 1c: Relational context `aw` representing document "annotated with" term.

consider a set of documents with authors as attributes (formal context $\mathcal{K}_1$ in Table 1a) and a set of terms with entities extracted from Wordnet (formal context in Table 1b). Then, we can consider the relation "document *annotated with* term" (denoted as `aw`) which defines a *relational context* as the one shown in Table 1c. RCA defines a *relational context family* (RCF) as a set of contexts $\mathbf{K} = \{\mathcal{K}_1, \mathcal{K}_2\}$ and a set of binary relations $\mathbf{R} = \{\mathbf{r}\}$. A relation $\mathbf{r} \subseteq \mathbf{G}_1 \times \mathbf{G}_2$ connects two object sets, a *domain* $\mathbf{G}_1$, $(\mathrm{dom}(\mathbf{r}) = \mathbf{G}_1)$ and a *range* $\mathbf{G}_2$, $(\mathrm{ran}(\mathbf{r}) = \mathbf{G}_2)$. Moreover, a relation $\mathbf{r}$ can be seen as a set-valued function $\mathbf{r} : \mathbf{G}_1 \rightarrow \wp(\mathbf{G}_2)$ [9].
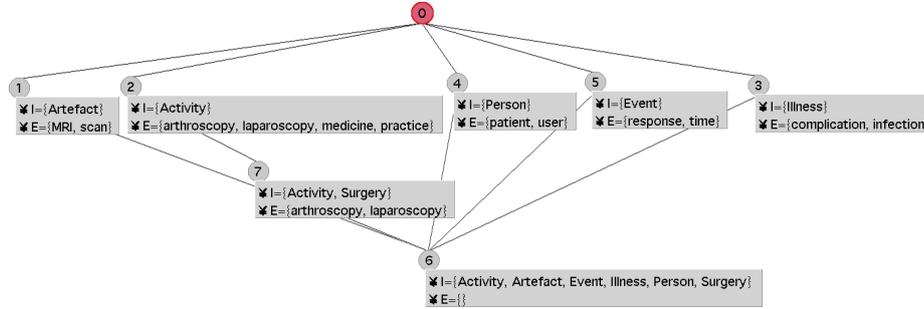
Fig. 1: Concept lattice of formal context $\mathcal{K}_2$ in Table 1b.

For the current example, let $\mathbf{G}_1$ be a set of documents and $\mathbf{G}_2$ be a set of terms. Then the corresponding RCF is composed by contexts $\mathcal{K}_1 = (\mathbf{G}_1, \mathbf{M}_1, \mathbf{I}_1)$ (with $\mathbf{M}_1, \mathbf{I}_1$ as shown in Table 1a), $\mathcal{K}_2 = (\mathbf{G}_2, \mathbf{M}_2, \mathbf{I}_2)$ (with $\mathbf{M}_2, \mathbf{I}_2$ as shown in Table 1b) and the relational context `aw` in Table 1c.

RCA is based on a *relational scaling* mechanism that transforms a relation $\mathbf{r}$ into a set of *relational attributes* that are added to complete the "initial context" describing

the object set $G_1 = \text{dom}(r)$. For each relation r, there is an *initial lattice* for each object set, i.e. $\mathcal{L}_1$ for G_1 and $\mathcal{L}_2$ for G_2.

The RCA mechanism starts from two initial lattices, $\mathcal{L}_1$ and $\mathcal{L}_2$, and builds a series of intermediate lattices by gradually completing the initial context $\mathcal{K}_1$ with new "relational attributes". Relational scaling follows the description logics (DL) semantics of role restrictions.

A relational attribute $\exists r : C$, C being a concept and $\exists$ the existential quantifier, is associated to an object $g \in G_1$ whenever $r(g) \cap \text{extent}(C) \neq \emptyset$ (other quantifiers are available, see [9]). The series of intermediate lattices converges toward a "fixpoint" or "final lattice" and the RCA mechanism is terminated. This is why there is one initial and one final lattice for each context of the considered RCF. For the running example, the lattice (in this case initial and final) in Figure 1 for the formal context $\mathcal{K}_2$ in Table 1b, along with the "relational context" in Table 1c, indicates the "relational attributes" that should be added to the formal context in Table 1a. For instance, using the existential quantifier, the relational attribute $\exists \text{aw} : \text{C1}$ (C1 is the concept with intent "Artefact" in Figure 1) should be added to all documents $g_i \in G_1$ in formal context $\mathcal{K}_1$ in Table 1a if g_i contains terms "MRI" or "scan" in the relational context of Table 1c. Table 2 shows formal context $\mathcal{K}_1$ after the relational scaling process.

	author1	author2	author3	author4	∃aw : C1	∃aw : C2	∃aw : C3	∃aw : C4	∃aw : C5	∃aw : C6	∃aw : C7
g_1		×	×		×	×	×	×			×
g_2	×	×			×	×		×	×		
g_3	×	×			×	×		×			×
g_4	×	×		×	×	×		×			
g_5			×						×	×	
g_6			×		×	×					×
g_7		×	×		×	×					×
g_8			×		×	×					×
g_9	×			×	×	×					×

Table 2: Context $\mathcal{K}_1$ after relational scaling using existential quantifier.

	patient	laparoscopy	scan	user	medicine	response	time	MRI	practice	complication	arthroscopy	infection
g_1	0.25	0.25	0.25	0	0	0	0	0	0	0.25	0	0
g_2	0	0	0.16	0.16	0.16	0.16	0.16	0	0.16	0	0	0
g_3	0	0.25	0	0.25	0.25	0	0	0.25	0	0	0	0
g_4	0.3	0	0	0	0.3	0	0	0.3	0	0	0	0
g_5	0	0	0	0.3	0	0.3	0.3	0	0	0	0	0
g_6	0	0	0	0	0	0	0	0	0.5	0	0.5	0
g_7	0	0	0	0	0	0	0	0	0	0.5	0.5	0
g_8	0	0	0	0	0	0	0	0	0	0.3	0.3	0.3
g_9	0	0	0	0	0	0	0	0	0	0	0.5	0.5

Table 3: Many-valued formal context of term frequencies in each document.

2.2 Pattern structure framework

Pattern structures model a FCA procedure when documents do not have attributes, but rather complex data descriptions such as numerical values, e.g. terms frequency values for each given document as shown in Table 3. In the following, we introduce the pattern structure framework firstly described in [3]. A pattern structure $(G_1, (D, \sqcap), \delta)$ is a generalization of a formal context where G_1 is a set of objects, $(D, \sqcap)$ is a semi-lattice of object descriptions and $\delta : G_1 \rightarrow D$ is a mapping associating a description to an object.

In the "interval pattern structures" setting (deeply discussed in [6]), an object descriptions $g \in G_1$ is a vector of intervals $d \in D$, $d = \langle [l_i, r_i] \rangle_{i \in \{1..|M|\}}$ with $l_i, r_i \in \mathbb{R}$ and $l_i \leq r_i$. For example, from Table 3 we have that the set of objects G_1 is composed by documents $g_1 - g_9$ (we use this notation for all documents between and including g_1 and g_9). The object description $\delta(g_1)$ is defined by the vector of intervals

$\langle[0.25, 0.25], [0.25, 0.25], [0.25, 0.25], [0, 0], ..., [0.25, 0.25], [0, 0], [0, 0]\rangle$. An interval pattern defines a convex region within the given description space.

In $(D, \sqcap)$ the *similarity* operator $\sqcap$ applied to two object descriptions $d_1 = \langle[l_i^1, r_i^1]\rangle$ and $d_2 = \langle[l_i^2, r_i^2]\rangle$ with $i \in \{1..|M|\}$ and $d_1, d_2 \in D$, returns the convex hull described in Equation 1 while the subsumption order $\sqsubseteq$ between them is given by Equation 2.

$$d_1 \sqcap d_2 = \langle[\min(l_i^1, l_i^2), \max(r_i^1, r_i^2)]\rangle \tag{1}$$

$$d_1 \sqsubseteq d_2 \iff d_1 \sqcap d_2 = d_1 \tag{2}$$

A Galois connection between $\wp(G_1)$ (powerset of G_1) and $(D, \sqcap)$ for $A \subseteq G_1$ and $d \in D$ is defined as follows:

$$A^{\square} = \bigsqcap_{g \in A} \delta(g) \qquad\qquad d^{\square} = \{g \in G | d \sqsubseteq \delta(g)\} \tag{3, 4}$$

Where $A^{\square}$ represents the common description to all objects in A while $d^{\square}$ represents the set of objects respecting the description d. A pair (A, d) such as $A^{\square} = d$ and $d^{\square} = A$ is called an *interval pattern concept (ip-concept)* with extent A and pattern intent d. Interval pattern concepts can be ordered in an interval pattern concept lattice (ip-concept lattice).

3 Inspiring problem - Latent Semantic Indexing

3.1 Latent variables characterization problem

As previously discussed, LV-models lack a proper characterization for the LVs found through its application. For instance, Latent Semantic Indexing (LSI) [2], a technique commonly used in information retrieval (IR) for indexation, clustering and dimension reduction purposes, is based on the idea that within a document-term matrix (as the one shown in Table 3) there is a set of hidden "latent variables" (LVs) that explain the data which constitutes the matrix. Consequently, LSI describes a technique to uncover these LVs through a "lower-rank approximation" of the original document-term matrix using linear algebra methods (specifically, singular value decomposition (SVD) [12]). Documents can later be described not as vectors of term frequencies, but as vectors of LV values in a reduced vectorial space. Latent variables are supposed to capture the "semantics" in the set of documents, nevertheless it is difficult to grasp this notion while documents are still described by vectors of numeric values. In the following, we provide a further description of the LSI process as described in [2].

3.2 Latent Semantic Indexing

Let us consider the values in the formal context in Table 3 as a matrix A of dimensions 9×12. LSI works through the SVD of matrix A and the consequent calculation of the reduced space of LVs as follows:

$$A_{(9\times12)} = U_{(9\times9)} \cdot \Sigma_{(9\times12)} \cdot V_{(12\times12)}^T \tag{5}$$

$$\tilde{A}_{(9\times12)} = U_{(9\times k)} \cdot \Sigma_{(k\times k)} \cdot V_{(k\times12)}^T \quad \text{(with } k \ll min(9,12)) \tag{6}$$

$$A \sim \tilde{A} \tag{7}$$

$$\tilde{A} \cdot \tilde{A}^T = U_{(9\times k)} \cdot \Sigma_{(k\times k)} \cdot V_{(k\times12)}^T \cdot V_{(12\times k)} \cdot \Sigma_{(k\times k)}^T \cdot U_{(k\times9)}^T \tag{8}$$

$$\tilde{A} \cdot \tilde{A}^T = (U_{(9\times k)} \cdot \Sigma_{(k\times k)}) \cdot (U_{(9\times k)} \cdot \Sigma_{(k\times k)})^T \tag{9}$$

Where $(A)^T$ denotes the "transpose" of matrix A; U, V are orthonormal matrices and Σ is a diagonal matrix of "singular values". We have on one side the lower-rank approximation (Equation 7) to a matrix of rank k which is ensured to be the best $k-rank$ matrix approximation by the Frobenius norm difference [12]. On the other hand, we have the dimensional reduction (Equation 9) using matrix $U_{(9\times k)} \cdot \Sigma_{(k\times k)}$ as the space of documents in k LVs. Table 4 shows this space for matrix A with $k = 2$. Furthermore, Figure 2 presents a graphical representation of documents as points in a plane where we can appreciate the presence of 2 document groups, usually called "clusters". In this paper we use the notion of "cluster" as a convex region in the LV space. In fact, one of the main uses of LSI is to provide a more compact representation of documents so that clusters are easier to find in the space of LVs. Incidentally, an interval pattern in this space represents a cluster (rectangles in Figure 2).

3.3 Problem statement

In Figure 2, while the clusters are easily distinguishable, it is not possible to say why they exist or what are their features. In order to characterize them we need to rely on

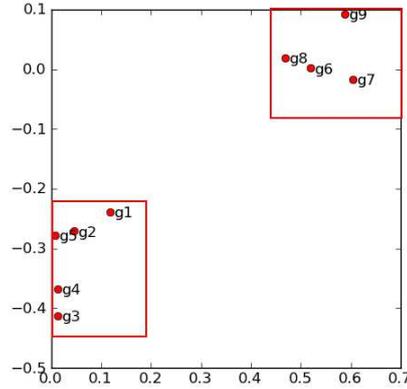|	k1	k2
g_1	0.118	-0.238
g_2	0.046	-0.271
g_3	0.014	-0.413
g_4	0.014	-0.368
g_5	0.008	-0.277
g_6	0.519	0.002
g_7	0.603	-0.017
g_8	0.469	0.02
g_9	0.588	0.092

Table 4: Documents in 2 LVs.

Fig. 2: Graphical representation of documents as points in a 2 dimensional LV space.

their relations with terms. For example, we know that documents $g_6 - g_9$ share the term "arthroscopy". While this is not totally clear with documents $g_1 - g_5$ which do not share a common term, we can see that documents $g_1 - g_4$ share the term "patient" and g_2, g_3 and g_5 share the term "user". Both terms are related through the annotation "People" extracted from Wordnet (see Table 1b) which lead us to think that LVs can represent differences in this concern.

One way to automatically make these characterizations is through the use of the RCA framework where we can model documents and terms as objects, LV values as document descriptions, and Wordnet annotations as term attributes, while the document-term relation is given by `aw` in Table 1c. Nevertheless, as explained in the previous section, LSI generates document descriptions in the form of vectors of LV values, while clusters in the LV-space are better represented by interval pattern structures.

The main problem tackled in this work is how to enable the application of RCA in these kinds of scenarios. We provide an adaptation of RCA which allows the relational scaling in pattern structures. We achieve this by the introduction of heterogeneous pattern structures described in the following section. A sub-goal of this work is to find out if domain knowledge can explain the existence and the "semantics" in LVs. We met this sub-goal by the characterization of LVs through the proposed combination of RCA and pattern structures. Given that LVs define a k-dimensional space (k being the number of LVs) where documents are organized, we formulate the following questions: Is it possible for us to find sub-regions in the space of LV values related to domain knowledge elements such as Wordnet annotations? And if so, how can we characterize these sub-regions?.

4 Adapting RCA for pattern structures

In this section we firstly describe the formal model description in which pattern structures are considered into a RCF. We show that the adaptation of the relational scaling operators induces a new space of heterogeneous object descriptions which we support in the framework of heterogenous pattern structures. Following, we provide a full description of this novel pattern structures instance.

4.1 Formal Model

Consider the simple case when we have a single relation between two sets of objects $r \subseteq G_1 \times G_2$, the domain of which is an object set in a pattern structure such as $\mathcal{K}_1 = (G_1, (D, \sqcap), \delta)$. The range of the relation is an object set inside a binary formal context $\mathcal{K}_2 = (G_2, M, I)$. Let us also define the relation as the set-valued function $r : G_1 \rightarrow \wp(G_2)$ and let $\mathcal{L}_1 = \underline{\mathfrak{B}}(\mathcal{K}_1)$ and $\mathcal{L}_2 = \underline{\mathfrak{B}}(\mathcal{K}_2)$ be the pattern concept lattice and the concept lattice of $\mathcal{K}_1, \mathcal{K}_2$ respectively. Thus, we define the relational context family $(\mathbf{K}, \mathbf{R})$ where $\mathbf{K} = \{\mathcal{K}_1, \mathcal{K}_2\}$ and $\mathbf{R} = \{r\}$. The usual RCA procedure induces iterations of formal context $\mathcal{K}_1$ through a "relational scaling" task using $\mathcal{L}_2$ ("target lattice" of r), until the derived concept lattice $\mathcal{L}_1$ converges. For this reason, the scaling operators (universal, existential, etc.) are defined over a space of formal contexts into a space of formal contexts. This is the first complication in our model. Since in our setting

$\mathcal{K}_1$ is a pattern structure and not a formal context, we cannot directly apply the scaling operators as defined in [9]. Thus, we move forward to redefine relational scaling operators which support pattern structures. To achieve this, let us define, for a relation $\mathtt{r}$, a function that assigns a set of relational attributes to a given object in the pattern structure depending on the type of relational scaling applied (universal, existential, etc.).

Definition 1. Let $\mathtt{r} \subseteq \mathtt{G}_1 \times \mathtt{G}_2$ be a relation between two object sets where $\mathcal{L}_2$ is its target lattice composed by formal concepts $\mathtt{C}$. We define the potential set of all possible relational attributes $\mathtt{P}_\mathtt{r}$ scaled from relation $\mathtt{r}$ as follows [3]:

$$\mathtt{P}_\mathtt{r} = \{\mathtt{r} : \mathtt{C}, \ \forall \mathtt{C} \in \mathcal{L}_\mathtt{j}\} \tag{10}$$

For $\mathtt{g} \in \mathtt{G}_1$, we also define two functions $\rho_\mathtt{r}^\exists, \rho_\mathtt{r}^{\forall\exists} : \mathtt{G}_1 \rightarrow \wp(\mathtt{P}_\mathtt{r})$ which assign a set of relational attributes to a given object using the 'existential quantifier operator ($\exists$)" and the "universal-existential quantifier operator ($\forall\exists$)" respectively.

$$\rho_\mathtt{r}^\exists(\mathtt{g}) = \{\mathtt{r} : \mathtt{C} \in \mathtt{P}_\mathtt{r} \mid \mathtt{r}(\mathtt{g}) \cap \mathtt{extent}(\mathtt{C}) \neq \emptyset\} \tag{11}$$

$$\rho_\mathtt{r}^{\forall\exists}(\mathtt{g}) = \{\mathtt{r} : \mathtt{C} \in \mathtt{P}_\mathtt{r} \mid \mathtt{r}(\mathtt{g}) \neq \emptyset, \mathtt{r}(\mathtt{g}) \subseteq \mathtt{extent}(\mathtt{C})\} \tag{12}$$

Hereafter we refer to $\rho_\mathtt{r}^\exists(g)$ or $\rho_\mathtt{r}^{\forall\exists}(g)$ as the "relations of g".

Example 1. Let the following scenario be the running example for the remainder of this article. Consider a relational context family of two contexts $\mathbf{K} = \{\mathcal{K}_1, \mathcal{K}_2\}$ where $\mathcal{K}_1 = (\mathtt{G}_1, (\mathtt{D}, \sqcap), \delta)$ is the interval pattern structure of documents and their LV values shown in Table 4 and $\mathcal{K}_2$ is the formal context of terms and their Wordnet annotations shown in Table 1b. Consider as well the relation "document *annotated with* term" as shown in Table 1c such as $\mathbf{R} = \{\mathtt{r}\}$. From the initial lattice shown in Figure 1 we can construct the set of relational attributes $\mathtt{P}_\mathtt{r} = \{\mathtt{aw} : \mathtt{C}_\mathtt{i}\}$ where $\mathtt{i} \in [0, 7]$ (i.e. each $\mathtt{C}_\mathtt{i}$ corresponds to one formal concept shown in the lattice[4]). Then, we have:

$$\mathtt{r}(\mathtt{g}_1) = \{\mathtt{patient}, \mathtt{laparoscopy}, \mathtt{scan}, \mathtt{complication}\}$$
$$\mathtt{extent}(\mathtt{C1}) = \{\mathtt{MRI}, \mathtt{scan}\}$$
$$\mathtt{r}(\mathtt{g}_1) \cap \mathtt{extent}(\mathtt{C1}) = \{\mathtt{scan}\} \neq \emptyset \Longrightarrow \mathtt{aw} : \mathtt{C1} \in \rho_\mathtt{r}^\exists(\mathtt{g}_1)$$
$$\rho_\mathtt{r}^\exists(\mathtt{g}_1) = \{\mathtt{aw} : \mathtt{C1}, \mathtt{aw} : \mathtt{C2}, \mathtt{aw} : \mathtt{C3}, \mathtt{aw} : \mathtt{C4}, \mathtt{aw} : \mathtt{C7}\}$$

Definition 2. Let $(\mathtt{G}_1, (\mathtt{D}, \sqcap), \delta)$ be a pattern structure for a set of objects $\mathtt{G}_1$ which are also associated with relational attributes in a set $\mathtt{P}_\mathtt{r}$ through $\rho_\mathtt{r}^\exists$ or $\rho_\mathtt{r}^{\forall\exists}$. We define the scaled pattern structure $(\mathtt{G}_1, (\mathtt{H}, \sqcap), \Delta)$ with mappings $\Delta^\exists, \Delta^{\forall\exists} : \mathtt{G}_1 \rightarrow \mathtt{H}$ as follows:

$$\mathtt{H} = \mathtt{D} \times \wp(\mathtt{P}_\mathtt{r}) \tag{13}$$

$$\Delta^\exists(\mathtt{g}) = (\delta(\mathtt{g}), \rho_\mathtt{r}^\exists(\mathtt{g})) \tag{14}$$

$$\Delta^{\forall\exists}(\mathtt{g}) = (\delta(\mathtt{g}), \rho_\mathtt{r}^{\forall\exists}(\mathtt{g})) \tag{15}$$

---

[3] Normally, the relational attributes $\mathtt{r} : \mathtt{C}$ have the operator $\exists$ or $\forall\exists$ attached as a prefix indicating the scaling operation applied. In this work, we omit the prefixes in favour of generality. Nevertheless, the scaling function will remain indicated at each step.

[4] aw stands for "annotated with". In the remainder of this article we will always work with the existential quantifier.

Where H contains heterogeneous descriptions of objects in G_1 combining both, a pattern $\delta(g) \in D$ and a set of relational attributes in P_r.

Definition 3. Let r be a relation between two objects sets, then the existential scaling operator $(\text{sc}_r^{\exists})$ and the universal scaling operator $\text{sc}_r^{\forall\exists}$ for a pattern structure $\mathcal{K}_1$ are defined as:

$$\text{sc}_r^{\exists}(\mathcal{K}_1) = (G_1, (H, \sqcap), \Delta^{\exists}) \qquad \text{sc}_r^{\forall\exists}(\mathcal{K}_1) = (G_1, (H, \sqcap), \Delta^{\forall\exists}) \qquad (16, 17)$$

As shown in Definitions 2 and 3, in order to apply the relational scaling operation to a pattern structure, it is necessary to define a new different pattern structure in which we can consider the original object description $\delta(g)$ and its relational attributes $\rho_r^{\exists}(g)$ or $\rho_r^{\forall\exists}(g)$. This combination of descriptions or "heterogeneous descriptions" H is a Cartesian product between the set of object descriptions and the powerset of P_r to which objects are mapped through $\Delta^{\exists} : G_1 \rightarrow H$. We denominate this new pattern structure instance "heterogeneous pattern structures". In the following, we provide a complete description of its characteristics and capabilities.

Example 2. Table 5 shows a representation of the heterogeneous pattern structure of documents with LVs and relational attributes, where we can find an object description such as:

$$\Delta^{\exists}(g_1) = (\delta(g_1), \rho_r^{\exists}(g_1))$$
$$\delta(g_1) = \langle [0.118, 0.118], [-0.238, -0.238] \rangle$$
$$\rho_r^{\exists}(g_1) = \{\texttt{aw} : \texttt{C1}, \texttt{aw} : \texttt{C2}, \texttt{aw} : \texttt{C3}, \texttt{aw} : \texttt{C4}, \texttt{aw} : \texttt{C7}\}$$

	D		P_r						
	k1	k2	aw : C1	aw : C2	aw : C3	aw : C4	aw : C5	aw : C6	aw : C7
g_1	0.118	-0.238	×	×	×	×			×
g_2	0.046	-0.271	×	×		×	×		
g_3	0.014	-0.413	×	×		×			×
g_4	0.014	-0.368	×	×		×			
g_5	0.008	-0.277				×	×		
g_6	0.519	0.002		×	×				×
g_7	0.603	-0.017		×	×				×
g_8	0.469	0.02		×	×				×
g_9	0.588	0.092		×	×				×

Table 5: Result of relational scaling in the example pattern structure represented in a hybrid formal context. We have removed the relational attribute aw : C0 usually assigned to every object.

4.2 Heterogeneous pattern structures

Definition 4. Let $H = D \times \wp(P_r)$ be a set of heterogeneous object descriptions, where $h_1 = (d_1, B_1)$ and $h_2 = (d_2, B_2)$ are two heterogeneous object descriptions with d_1, $d_2 \in D$, $B_1, B_2 \subseteq P_r$ and $h_1, h_2 \in H$ (the elements d and B are referred to as the "components" of h). We define the "similarity operator" $\sqcap$ between h_1 and h_2 as:

$$h_1 \sqcap h_2 = (d_1 \sqcap d_2, B_1 \cap B_2) \tag{18}$$

Example 3. The similarity operator applied to the object descriptions of g_1 and g_2 is:

$$\Delta^{\exists}(g_1) \sqcap \Delta^{\exists}(g_2) = (\delta(g_1) \sqcap \delta(g_2), \rho_r^{\exists}(g_1) \cap \rho_r^{\exists}(g_2))$$
$$\delta(g_1) \sqcap \delta(g_2) = \langle [0.046, 0.118], [-0.271, -0.238] \rangle$$
$$\rho_r^{\exists}(g_1) \cap \rho_r^{\exists}(g_2) = \{\texttt{aw} : \texttt{C1}, \texttt{aw} : \texttt{C2}, \texttt{aw} : \texttt{C4}\}$$

Proposition 1. $(H, \sqsubseteq)$ *with $\sqcap$ as described in Definition 4 is the direct product of the ordered sets $(D, \sqsubseteq)$ and $(\wp(P_r), \subseteq)$ and thus is an ordered set itself.*

Proof. In order to prove that $(H, \sqsubseteq)$ is the direct product of $(D, \sqsubseteq)$ and $(\wp(P_r), \subseteq)$, we show that $h_1 \sqsubseteq h_2 :\iff d_1 \sqsubseteq d_2$ and $B_1 \subseteq B_2$ (as described in [4]).

$$
\begin{align}
h_1 \sqsubseteq h_2 &\iff h_1 \sqcap h_2 = h_1 & \text{Equation 2} \tag{19}\\
&\iff (d_1 \sqcap d_2, B_1 \cap B_2) = (d_1, B_1) & \text{Definition 4} \tag{20}\\
&\iff d_1 \sqcap d_2 = d_1 \text{ and } B_1 \cap B_2 = B_1 & \tag{21}\\
&\iff d_1 \sqsubseteq d_2 \text{ and } B_1 \subseteq B_2 & \tag{22}
\end{align}
$$

$\square$

Because of Proposition 1, we would like to know how the heterogeneous pattern concept lattice is related to the concept lattices of its components, namely the pattern concept lattice $(G_1, (D, \sqcap), \delta)$ and the concept lattice of the formal context of objects and their respective relational attributes (G_1, P_r, I) where the incidence relation I is defined in Equation 23. Regarding this, for the following definitions we introduce an alternative description for the standard FCA derivation operator $(\cdot)'$ in Equation 24 for a subset of objects $A \in G_1$ using the function $\rho_r^{\exists}$.

$$I = \bigcup_{g \in G_1} \{(g, m), \forall m \in \rho_r^{\exists}(g)\} \qquad\qquad A' = \bigcap_{g \in A} \rho_r^{\exists}(g) \tag{23, 24}$$

Definition 5. The derivation operators $(\cdot)^{\diamond}$ in $(G_1, (H, \sqcap), \Delta^{\exists})$ for an object set $A \in G_1$ and a heterogeneous element $h \in H$ are defined as:

$$A^{\diamond} = \bigsqcap_{g \in A} \Delta^{\exists}(g) \qquad\qquad h^{\diamond} = \{g \in G_1 \iff h \sqsubseteq \Delta^{\exists}(g)\} \tag{25, 26}$$

A heterogeneous pattern concept (hp-concept) is then defined as the pair (A, h) where $h^{\diamond} = A$ and $A^{\diamond} = h$.

Proposition 2. *The derivation operator applied to a heterogeneous element* $\mathtt{h} = (\mathtt{d}, \mathtt{B})$ *is equal to the intersection of the derivation operator on its components:*

$$(\mathtt{d}, \mathtt{B})^{\diamond} = \mathtt{d}^{\square} \cap \mathtt{B}' \qquad (27)$$

Proof. Let $\mathtt{g} \in \mathtt{h}^{\diamond}$, with $\mathtt{h} = (\mathtt{d}, \mathtt{B})$, by Equation 26 we have:

$$\mathtt{g} \in \mathtt{h}^{\diamond} \iff \mathtt{h} \sqsubseteq \Delta^{\exists}(\mathtt{g}) \iff \mathtt{d} \sqsubseteq \delta(\mathtt{g}) \text{ and } \mathtt{B} \subseteq \rho_{\mathtt{r}}^{\exists}(\mathtt{g}) \qquad \text{Proposition 1}$$

The right side of last formula shows two conditions. Using Equation 4, we have that the first condition yields $\mathtt{d} \sqsubseteq \delta(\mathtt{g}) \iff \mathtt{g} \in \mathtt{d}^{\square}$. As for the second condition, in Equation 23, we have that $(\mathtt{g}, \mathtt{m}) \in \mathtt{I}, \forall \mathtt{m} \in \rho_{\mathtt{r}}^{\exists}(\mathtt{g})$. Then, $\forall \mathtt{m} \in (\mathtt{B} \subseteq \rho_{\mathtt{r}}^{\exists}(\mathtt{g}))$ we have that $(\mathtt{g}, \mathtt{m}) \in \mathtt{I}$ and thus $\mathtt{g} \in \mathtt{B}'$. With this we have that:

$$\mathtt{g} \in (\mathtt{d}, \mathtt{B})^{\diamond} \iff \mathtt{g} \in \mathtt{d}^{\square} \text{ and } \mathtt{g} \in \mathtt{B}'$$
$$(\mathtt{d}, \mathtt{B})^{\diamond} = \mathtt{d}^{\square} \cap \mathtt{B}' \qquad \square$$

Proposition 3. *The closure of a set of objects* $\mathtt{A} \in \mathtt{G_1}$ *(an extent) is equal to the intersection of its closures in each component.*

$$\mathtt{A}^{\diamond\diamond} = \mathtt{A}^{\square\square} \cap \mathtt{A}'' \qquad (28)$$

Proof.

$$\mathtt{A}^{\diamond\diamond} = (\prod_{\mathtt{g} \in \mathtt{A}} \Delta^{\exists}(\mathtt{g}))^{\diamond} = (\prod_{\mathtt{g} \in \mathtt{A}} \delta(\mathtt{g}), \bigcap_{\mathtt{g} \in \mathtt{A}} \rho_{\mathtt{r}}^{\exists}(\mathtt{g}))^{\diamond} = (\mathtt{A}^{\square}, \mathtt{A}')^{\diamond} = \mathtt{A}^{\square\square} \cap \mathtt{A}'' \qquad \square$$

From Proposition 3, we can see three different conditions for a heterogeneous extent $\mathtt{A}$, namely it can be closed in both of its components ($\mathtt{A}^{\diamond\diamond} = \mathtt{A}^{\square\square} = \mathtt{A}''$), in only one (either $\mathtt{A}^{\square\square} \subseteq \mathtt{A}''$ or $\mathtt{A}'' \subseteq \mathtt{A}^{\square\square}$), or in none ($\mathtt{A}^{\square\square} \not\subseteq \mathtt{A}''$ or $\mathtt{A}'' \not\subseteq \mathtt{A}^{\square\square}$). Further in is this section, we provide a full description for these kinds of extents. Nevertheless, Proposition 3 provides us with two ways to calculate the set of heterogeneous pattern concepts. Firstly, Equation 28 is a canonical test which can be used in standard FCA algorithms such as AddIntent [13]. Secondly, we can calculate the complete set of extents from both, the formal context and the pattern structure separately and intersect them to calculate each possible heterogeneous extent.

Example 4. Consider the object set $\mathtt{A_1}$ in Table 6. The closure in the fifth column shows that $\mathtt{A_1} = \mathtt{A_1}^{\diamond\diamond}$ and thus it is a proper extent of $(\mathtt{G_1}, (\mathtt{H}, \sqcap), \Delta^{\exists})$. This is not the case for $\mathtt{A_2}$.

Proposition 4. *The closure of a heterogeneous description* $h \in \mathtt{H}$ *is given by:*

$$h^{\diamond\diamond} = (h^{\diamond\square}, h^{\diamond\prime}) \qquad (29)$$

Proposition 4 can be demonstrated analogously to Proposition 3. We are interested in Proposition 4 because it allows us to easily calculate the heterogeneous intents as we show next.

Extent A_i	$(A_i)^{\square\square}$	$(A_i)''$	$A^{\Diamond\Diamond} = A^{\square\square} \cap A''$	$(A_i)^{\Diamond}$
$A_1 = \{g_1, g_3\}$	$\{g_1 - g_4\}$	$\{g_1, g_3\}$	$\{g_1, g_3\}$	-
$A_2 = \{g_5, g_9\}$	$\{g_1, g_2, g_5, g_6, g_8, g_9\}$	G_1	$\{g_1, g_2, g_5, g_6, g_8, g_9\}$	-
$A_3 = \{g_1, g_6 - g_9\}$	A_3	A_3	A_3	$(A_3^{\square}, A_3')$
$A_4 = \{g_6, g_7\}$	A_4	A_3	A_4	$(A_4^{\square}, A_4')$
$A_5 = \{g_1, g_3, g_7\}$	$\{g_1 - g_4, g_7\}$	$\{g_1, g_3, g_6 - g_9\}$	A_5	$(A_5^{\square}, A_5')$

Table 6: Table showing different object sets under different closures. A_1 is a proper extent of $(G_1, (H, \sqcap), \Delta^{\exists})$ because its closed under $(\cdot)^{\Diamond}$ while A_2 is not. A_3 and A_4 are examples of "pure hp-concepts". A_5 is an example of a "mixed hp-concept".

Proposition 5. *Let A_1 be an extent in $(G_1, (D, \sqcap), \delta)$ and A_2 be an extent in (G_1, P_r, I) where $A_1 \subseteq A_2$ and for any other extent A in (G_1, P_r, I) we have $A_1 \subseteq A \subseteq A_2 \iff A_2 = A$, i.e. A_2 is the cover of A_1. Then for $h = (A_1^{\square}, A_2')$, h is a heterogeneous intent and (A_1, h) is a hp-concept.*

Proof. We show that $(A_1^{\square}, A_2')^{\Diamond\Diamond} = (A_1^{\square}, A_2')$

$$(A_1^{\square}, A_2')^{\Diamond\Diamond} = (A_1^{\square\square} \cap A_2'')^{\Diamond} = (A_1 \cap A_2)^{\Diamond} = (A_1)^{\Diamond}$$
$$= \prod_{G \in A_1} \Delta^{\exists}(g) = (\prod_{G \in A_1} \delta(g), \bigcap_{G \in A_1} \rho_r^{\exists}(G))$$
$$= (A_1^{\square}, A_1') = (A_1^{\square}, A_2')$$

The last step can be shown by the restrictions imposed to A_1 and A_2 as follows:
$A_1 \subseteq A_2 \implies A_1 \subseteq A_1'' \subseteq A_2 \implies A_1'' = A_2 \implies A_1' = A_2'$ $\qquad\square$

Similarly, it can be shown that when $A_2 \subseteq A_1$, the hp-concept $(A_2, (A_2^{\square}, A_2'))$ exists. Proposition 5 shows that the extents in the pattern structure $(G_1, (D, \sqcap), \delta)$ and in (G_1, P_r, I) will be present in the lattice of hp-concepts. Nevertheless, these do not cover the whole set of hp-concepts in $(G_1, (H, \sqcap), \Delta^{\exists})$.

As previously discussed, the set of hp-concepts (denoted as $\mathfrak{B}((G_1, (H, \sqcap), \Delta^{\exists}))$) can be characterized as containing three types of extents, those that are closed under both components, those that are closed under one of its components and those that are an intersection of two different closed extents. We call these types "pure hp-concepts", "semi-pure hp-concepts" and "mixed hp-concepts" respectively.

Definition 6. Given a hp-concept $(A, h) \in (G_1, (H, \sqcap), \Delta^{\exists})$ we say that:

$$(A, h) \text{ is "pure" iff } A^{\square\square} = A'' \tag{30}$$

$$(A, h) \text{ is "semi-pure" iff } A^{\square\square} \subseteq A'' \text{ or } A'' \subseteq A^{\square\square} \tag{31}$$

$$(A, h) \text{ is "mixed" iff } A^{\square\square} \cap A'' \neq \emptyset \text{ and } A^{\square\square} \nsubseteq A'' \text{ and } A'' \nsubseteq A^{\square\square} \tag{32}$$

Example 5. In Table 6, A_3 is a pure hp-concept extent since it is closed in both components. A_4 is a semi-pure hp-concept extent since it is closed in the pattern structure

component but not in the relational attribute component. A_5 is a mixed hp-concept as it is closed in the hp-lattice but not in either of its components.

In order to obtain the whole set of hp-concepts, it is not sufficient to calculate the sets of pattern concepts and formal concepts from its respective components and match them using Proposition 5. Doing so only provides us with the set of pure and semi-pure hp-concepts, while the set of mixed hp-concepts will be missing. In the following, we describe our method to compute the whole set of hp-concepts.

4.3 Calculating the hp-lattice

The heterogeneous pattern structure $(G_1, (H, \sqcap), \Delta^{\exists})$ has been defined as a standard pattern structure and thus a standard algorithm to calculate pattern concept lattices can be used to obtain the hp-lattice. Some of these algorithms have been described and discussed in [6]. However, a much simpler manner to calculate the hp-lattice is through the use of a "scaled representation context".

A "representation context", as explained in [3], is a mechanism of complex data binarization. The pattern concepts of a pattern structure and the formal concepts of its derived representation context are in 1-1 correspondence and furthermore, their extents are the same [3,7]. In the particular case of a heterogeneous pattern structure as described in this work, we use the representation context of the pattern structure component which is later "relationally scaled" in terms of traditional RCA (see Section 2.1).

Definition 7. Let $(G_1, (H, \sqcap), \Delta^{\exists})$ be a heterogeneous pattern structure with components $(G_1, (D, \sqcap), \delta)$ and (G_1, P_r, I). The "scaled representation context" is defined as $(G_1, D \cup P_r, J)$ where the incidence relation is:

$$(g, x) \in J \iff x \sqsubseteq \delta(g) \text{ or } x \in \rho_r^{\exists}(g); \forall g \in G_1 \text{ and } x \in (D \cup P_r)$$

In other words, $(G_1, D \cup P_r, J)$ is the representation context of the pattern structure $(G_1, (D, \sqcap), \delta)$ plus the relational scaling of (G_1, P_r, I). It can be shown that, in fact, this "scaled representation context" is isomorphic to the representation context of $(G_1, (H, \sqcap), \Delta^{\exists})$. For the running example, we constructed the scaled representation context as depicted in Table 7. In this context, patterns and relational attributes are treated equally, hence the attribute set $D \cup P_r$. Patterns in D were filtered using a similarity threshold as described in [5], since the complete non-restricted pattern lattice contain a little more than 100 concepts. Incidentally, the filter by similarity applied to the calculation of D caused the hp-lattice derived from the context in Table 7 to contain only pure and semi-pure, i.e. their extents are either closed under $(\cdot)^{\square}$ or $(\cdot)'$ or both.

While there are some drawbacks w.r.t. the computational costs associated with the calculation of the formal concepts of the representation context, in this work we disregard them favouring the simplicity of the combined model.

5 Discussion and conclusions

In Section 3.3 we proposed two questions that we discuss in the following.

										D																									P_r						
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	aw:C0	aw:C1	aw:C2	aw:C3	aw:C4	aw:C5	aw:C7
g1	×	×	×	×	×	×	×																										×	×	×	×	×		×
g2		×	×	×	×	×	×	×	×	×	×	×																					×	×	×		×	×	
g3			×	×				×	×			×	×	×																			×	×	×		×		×
g4		×	×	×	×				×	×	×	×		×	×	×	×																×	×	×		×		
g5			×	×	×				×	×	×		×		×		×	×															×				×	×	
g6																			×	×	×	×	×	×	×	×							×		×	×			×
g7																				×	×	×	×				×	×					×		×	×			×
g8																					×	×		×	×			×	×				×		×	×			×
g9																					×	×		×	×		×		×	×	×		×		×	×			×

Table 7: Scaled representation context for the running example. Patterns in D are represented by cardinals from 2 to 33 (number 1 was eliminated as it references the pattern concept $\top$).

Is it possible for us to find sub-regions in the space of LV values related to domain knowledge elements? Indeed, we can. A hp-concept describes exactly this in its intent as a relation of an interval pattern and a set of annotations in the Wordnet taxonomy. Moreover, these relations can be better described in the form of association rules [4]. Particularly, we are searching for those association rules with a premise in the space of latent variables and a consequence in the space of relational attributes, For example, we have the rule $6 \leftrightarrow$ aw : C4 which means that the latent variable region in the interval pattern numbered 6 implies the Wordnet concept "People" as shown in Figure 3. While all kinds of association rules exist in the lattice of the scaled representation context, we are only interested in those related to our specific problem. Figure 3 presents a graphical representation for the association rules extracted on the running example. The map represents what can be called a "labelled hierarchical document clustering" [8] over the space of latent variables. In the map, the region marked as "Activity" is actually a union of two contiguous regions.

How can we characterize the relations among sub-regions in the space of LV values and domain knowledge elements? We have already described three types of hp-concepts, namely pure, semi-pure and mixed. In the following, we provide them with a characterization. Let us first introduce the Jaccard index [8] in terms of the hp-concept's extents and the extents of its components as follows ($|\cdot|$ represents set cardinality):

$$J(\mathtt{A}^{\square\square}, \mathtt{A}'') = \frac{|\mathtt{A}^{\square\square} \cap \mathtt{A}''|}{|\mathtt{A}^{\square\square} \cup \mathtt{A}''|} = \frac{|\mathtt{A}^{\diamond\diamond}|}{|\mathtt{A}^{\square\square} \cup \mathtt{A}''|}$$

Pure hp-concepts are interesting since they represent strong coherent relations between clusters in different spaces. Moreover, for any given pure hp-concept $(\mathtt{A}, \mathtt{h})$, the Jaccard index $J(\mathtt{A}^{\square\square}, \mathtt{A}'') = 1$. Consider for example, the pure hp-concept with extent $g_1 - g_5$ (region 6) which represents a "closed" region in the latent variable space related to the topic "People", i.e. outside this region, there are no documents related to "People". We can also relate "pure hp-concepts" as describing *necessary and sufficient conditions* of a defined concept in the description logics framework (DL) [1]. In this case, documents in region 6 have the necessary and sufficient condition of being labelled with the annotation "People".

A semi-pure hp-concept represents a directional coherence, i.e. either $\mathtt{A}^{\square\square} \subseteq \mathtt{A}''$ or $\mathtt{A}'' \subseteq \mathtt{A}^{\square\square}$. The Jaccard index is determined by $\mathtt{A}^{\square\square}/\mathtt{A}''$ in the first case or $\mathtt{A}''/\mathtt{A}^{\square\square}$ in the later. For example, the hp-concept with extent $g_6 - g_7$ (region 22) contains documents related to "Illness and "Surgery", but it does not contain all of them (i.e. g_1 is an exception). Thus, we can call a semi-pure hp-concept an "open" region in the latent variable space. In DL terms, semi-pure hp-concepts represent necessary conditions, i.e. region 22 have the necessary but not sufficient condition of being labelled with the annotation "Surgery". Mixed hp-concepts represent a weak coherence of clusters. In general, their Jaccard index will be lower than the index of semi-pure hp-concepts.

$\mathtt{A}_i$	objects	$(\mathtt{A}_i)^{\square\square}$	$(\mathtt{A}_i)''$	$\mathtt{A}^{\diamond\diamond}$
$\mathtt{A}_a$	$\{g_1, g_3, g_4\}$	$\{g_1, g_2, g_3, g_4\}$	$\{g_1, g_3, g_4, g_5\}$	$\{g_1, g_3, g_4\}$
$\mathtt{A}_a^\diamond = (\langle [0.014, 0.118][-0.413, -0.238]\rangle, \{\mathtt{aw} : \mathtt{C4}\})$				

Table 8: Table showing an imaginary mixed hp-concept.

Finally, we can conclude that the technique presented in this paper is able to find useful relations among convex latent variable regions and domain knowledge which allows giving a proper characterization to the latent variable space, and hence, the latent variables themselves. This is possible due to the simultaneous representation of documents in the latent variable vectorial space and the set of relational attributes as hp-concepts.

The implications of this work are multiple. In this work we have superficially described some connections with descriptions logics. Furthermore, the notion of mixed hp-concepts, left unexplored in this work, lead us to think that they may be useful for annotation and data correction purposes. Other application domains seem also to fit as heterogeneous pattern structures. For example, in image annotation, images are characterized as vectors of features which are then aligned with annotations in the Wordnet taxonomy.

References

1. Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications.* Cambridge University Press, New York, NY, USA, 2003.
2. Scott Deerwester, Susan T. Dumais, Thomas K. Landauer, George W. Furnas, and Richard Harshman. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6):1097–4571, 1990.
3. Bernhard Ganter and Sergei O. Kuznetsov. Pattern structures and their projections. In *ICCS*, volume 2120 of *Lecture Notes in Computer Science*, pages 129–142. Springer, 2001.
4. Bernhard Ganter and Rudolf Wille. *Formal Concept Analysis: Mathematical Foundations.* Springer, December 1999.
5. Mehdi Kaytoue, Zainab Assaghir, Amedeo Napoli, and Sergei O. Kuznetsov. Embedding tolerance relations in formal concept analysis. In *Proceedings of the 19th ACM international conference on Information and knowledge management - CIKM '10*, page 1689, New York, New York, USA, October 2010. ACM Press.
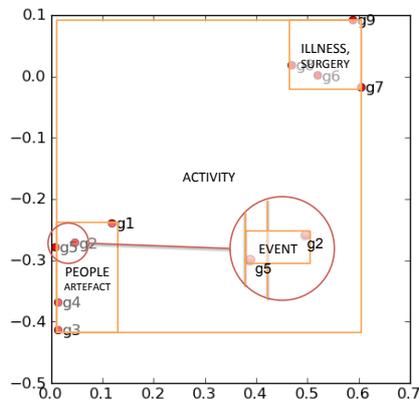
Fig. 3: Labelled document clusters using association rules from the hp-lattice with magnification on documents g_2 and g_5.

6. Mehdi Kaytoue, Sergei O. Kuznetsov, and Amedeo Napoli. Revisiting numerical pattern mining with formal concept analysis. *Proceedings of the Twenty-Second international joint conference on Artificial Intelligence - Volume Volume Two*, pages 1342–1347, November 2011.
7. Sergei O. Kuznetsov. Pattern Structures for Analyzing Complex Data. In *Proceedings of the 12th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing*, volume 5908 of *Lecture Notes in Computer Science*, pages 33–44. Springer Berlin Heidelberg, December 2009.
8. Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schtze. *Introduction to Information Retrieval*. July 2008.
9. Mohamed Rouane-Hacene, Marianne Huchard, Amedeo Napoli, and Petko Valtchev. A proposal for combining Formal Concept Analysis and description Logics for mining relational data. In *Proceedings of ICFCA 2007*, pages 51–65. LNAI 4390, Springer, 2007.
10. Mohamed Rouane-Hacene, Marianne Huchard, Amedeo Napoli, and Petko Valtchev. Relational concept analysis: mining concept lattices from multi-relational data. *Annals of Mathematics and Artificial Intelligence*, 67(1):81–108, March 2013.
11. Ashok Srivastava and Mehran Sahami. *Text Mining: Classification, Clustering, and Applications*. Chapman & Hall/CRC, 1st edition, 2009.
12. Lloyd Trefethen and David Bau. *Numerical Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, 1997.
13. Dean van der Merwe, Sergei Obiedkov, and Derrick Kourie. AddIntent: A New Incremental Algorithm for Constructing Concept Lattices. In *Concept Lattices*, volume 2961 of *Lecture Notes in Computer Science*, pages 205–206. Springer, Berlin/Heidelberg, 2004.