A proposition for combining pattern structures and RCA.

ICFCA 2014

Victor Codocedo, Amedeo Napoli

Introduction

Featuring in this presentation:

- Latent Variable models [8].
- Latent Semantic Indexing [2].
- Relational Concept Analysis [9].
- Interval Pattern Structures [7].
- Heterogeneous Pattern Structures.

1

Problem definition

Inspiring example

Latent variables model

Latent variables models (LV-models) (also called “topic models”) use the notion of hidden latent variables in data to “explain” information [2].

Inspiring example

Latent variables model

Latent variables models (LV-models) (also called “topic models”) use the notion of hidden latent variables in data to “explain” information [2].

They work by **unveiling** these variables and using them to process data.

Inspiring example

Latent models available:

- Principal Component Analysis (PCA) (1901) [8].
- Latent Semantic Indexing/Analysis (LSI/LSA) (1988) [2].
- Probabilistic Latent Semantic Indexing (PLSA) (1999) [6].
- Latent Dirichlet Allocation (LDA) (2003) [1].

Inspiring example

Why are latent variables useful? [2, 10, 8]

Several things:

- They can introduce “latent knowledge” not explicit in data.
- They can reduce noise.
- They can reduce the search space.
- They can reduce sparsity.
- They can represent “clusters”.

Inspiring example

Latent Semantic Indexing

LSI is based on the *lower-rank* approximation of a document-term* matrix and the representation of documents in a vectorial space of latent variables. For both of these issues, it uses the *Singular Value Decomposition* process.

	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: Document-term matrix *A*.

Inspiring example

Latent Semantic Indexing

SVD Process:

$$A_{(9\times12)} = U_{(9\times9)} \cdot \Sigma_{(9\times12)} \cdot V^T_{(12\times12)} \tag{1}$$

$$\tilde{A}_{(9\times12)} = U_{(9\times k)} \cdot \Sigma_{(k\times k)} \cdot V^T_{(k\times12)} \quad (\text{with } k \ll min(9,12)) \tag{2}$$

$$A \sim \tilde{A} \tag{3}$$

$$\tilde{A} \cdot \tilde{A}^T = U_{(9\times k)} \cdot \Sigma_{(k\times k)} \cdot V^T_{(k\times12)} \cdot V_{(12\times k)} \cdot \Sigma^T_{(k\times k)} \cdot U^T_{(k\times9)} \tag{4}$$

$$\tilde{A} \cdot \tilde{A}^T = (U_{(9\times k)} \cdot \Sigma_{(k\times k)}) \cdot (U_{(9\times k)} \cdot \Sigma_{(k\times k)})^T \tag{5}$$

Inspiring example

	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: Documents in 2 LVs. (k=2)

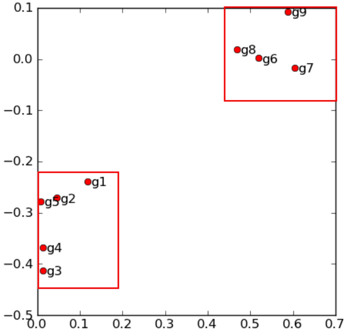
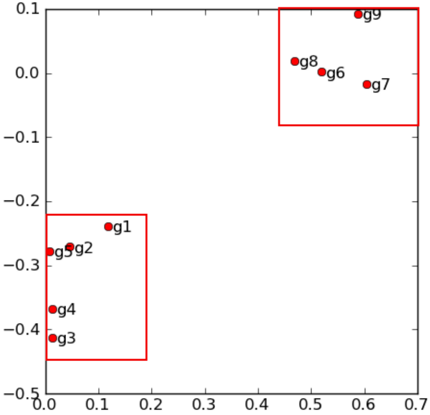


Figure: Graphical representation of documents as points in a 2 dimensional LV space.

Inspiring example

Where are the semantics?



Inspiring example

Where are the semantics?

- Latent variables are abstractions.
- A given LV or a convex region in a LV-space can represent a topic, but they lack a proper characterization.
- Cannot introduce external knowledge sources.

Inspiring example

Formal Concept Analysis [4], on the other hand:

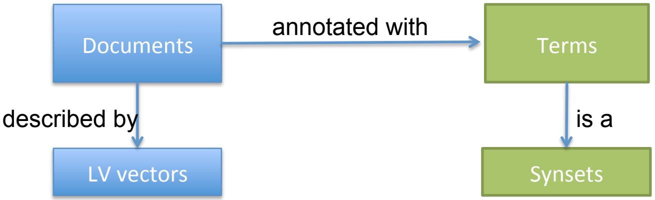
- Provides a formal characterization of concepts through the dual extent/intent descriptions.
- Allows the introduction of external knowledge sources through object relations (RCA).
- Allows the analysis of complex data such as convex regions in a vectorial space (interval pattern structures).

2

Definitions

Scenario

Can we relate convex regions in a LV-space to taxonomical objects?



In fact, this scenario fits with the Relational Concept Analysis process.

Definitions

Relational Concept Analysis (RCA) [9]

RCA describes an iterative scaling process to obtain a family of related concept lattices from a relational context family.

	k1	k2
g_1	×	
g_2		×
g_3		×
g_4	×	×
g_5		×
g_6		×
g_7	×	
g_8	×	×
g_9	×	

Table: Formal Context $\mathcal{K}_1 = (G_1, M_1, I_1)$

	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: Relational Context $aw = (G_1, G_2, I_{aw})$

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

Table: Formal Context $\mathcal{K}_2 = (G_2, M_2, I_2)$

Definitions

Relational Concept Analysis (RCA)

- A *relational context family* (RCF) is composed by:
  - A set of formal contexts $\mathbf{K} = \{\mathcal{K}_1, \mathcal{K}_2\}$.
  - A set of binary relations $\mathbf{R} = \{\mathtt{aw}\}$.
- A *relational context* can be also defined as a function $\mathtt{aw} : \mathtt{G}_1 \to \mathtt{G}_2$, where $\mathtt{dom}(\mathtt{aw}) = \mathtt{G}_1$ and $\mathtt{ran}(\mathtt{aw}) = \mathtt{G}_2$.
- A set of *relational attributes* is built from the concept lattice of the formal context with objects $\mathtt{ran}(\mathtt{aw})$.
- A *relational scaling* process applied in the formal context with objects $\mathtt{dom}(\mathtt{aw})$ assigns a set of relational attributes to an object $\mathtt{g} \in \mathtt{G}_1$ whenever $\mathtt{aw}(\mathtt{g}) \cap \mathtt{extent}(\mathtt{C}) \neq \emptyset$ ($\exists$ quantifier).

Definitions

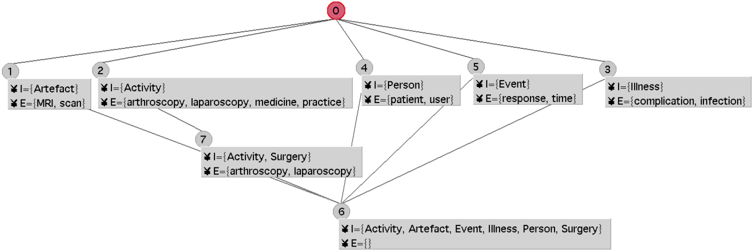


Figure: Concept Lattice for Taxonomic annotations $\mathcal{L}_2$.

RCA - Relational Scaling

$$\mathtt{aw}(\mathtt{g}_1) \cap \mathtt{extent}(\mathtt{C4}) = \{\mathtt{patient}, \mathtt{user}\}$$

$$\Longrightarrow \mathcal{K}_1^{(1)} = (\mathtt{G}_1, \mathtt{M}_1 \cup \{\mathtt{aw}\colon \mathtt{C4}\}, \mathtt{I}_1 \cup \{(\mathtt{g}_1, \mathtt{aw}\colon \mathtt{C4})\})$$

Definitions

Relational Concept Analysis

- Formal concepts in $\mathcal{K}_1^{(1)}$ have intents which relate LV with taxonomical annotations in $\mathcal{K}_2$.
- Nevertheless, $\mathcal{K}_1$ is a many-valued context. Convex regions in a LV-space are better described with interval pattern structures [3, 7].
- **What happens if we apply relational scaling in a many-valued formal context?**

Definitions

	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: Heterogeneous formal context.

Problems

- Objects are described by heterogeneous patterns mixing values and binary attributes.
- It becomes necessary to define a proper pattern structure which is able to deal with heterogeneous object descriptions.

Proposition

Relational Scaling for pattern structures

Let a RCF contain a relation $\mathtt{r}$ where $\mathtt{dom}(\mathtt{r})$ is an object set in a pattern structure $\mathcal{K}_1 = (\mathtt{G}_1, (\mathtt{D}, \sqcap), \delta)$, then we define:

- The set of all relational attributes extracted for relation $\mathtt{r}$ as $\mathtt{P}_\mathtt{r} = \{\mathtt{r} : \mathtt{C},\ \forall \mathtt{C} \in \mathcal{L}_2\}$.
- An assignation function $\rho^{\exists}_{\mathtt{r}} : \mathtt{G}_1 \to \mathtt{P}_\mathtt{r}$, such as: $\rho^{\exists}_{\mathtt{r}}(\mathtt{g}) = \{\mathtt{r} : \mathtt{C} \in \mathtt{P}_\mathtt{r} \mid \mathtt{r}(\mathtt{g}) \cap \mathtt{extent}(\mathtt{C}) \neq \emptyset\}$.
- An heterogeneous set of descriptions $\mathtt{H} = \mathtt{D} \times \wp(\mathtt{P}_\mathtt{r})$.
- A mapping $\Delta^{\exists} : \mathtt{G}_1 \to \mathtt{H}$, such as $\Delta^{\exists}(\mathtt{g}) = (\delta(\mathtt{g}), \rho^{\exists}_{\mathtt{r}}(\mathtt{g}))$.

Proposition

Relational Scaling for pattern structures

The relational scaling of the pattern structure $(\mathtt{G}_1, (\mathtt{D}, \sqcap), \delta)$ is defined as:

$$\mathtt{sc}^{\exists}_{\mathtt{r}}(\mathcal{K}_1) = (\mathtt{G}_1, (\mathtt{H}, \sqcap_{\mathtt{H}}), \Delta^{\exists})$$

Where $(\mathtt{G}_1, (\mathtt{H}, \sqcap_{\mathtt{H}}), \Delta^{\exists})$ is called a *heterogeneous pattern structure*.

Heterogeneous pattern structures

Properties:

- A heterogeneous object description $\mathtt{h} \in \mathtt{H}$ is a pair $\mathtt{h} = (\mathtt{d}, \mathtt{B})$ where $\mathtt{d} \in \mathtt{D}$ and $\mathtt{B} \subseteq \mathtt{P_r}$.
- For $\mathtt{h_1} = (\mathtt{d_1}, \mathtt{B_1})$ and $\mathtt{h_2} = (\mathtt{d_2}, \mathtt{B_2})$, the similarity operator $\sqcap_{\mathtt{H}}$ is defined as $\mathtt{h_1} \sqcap_{\mathtt{H}} \mathtt{h_2} = (\mathtt{d_1} \sqcap \mathtt{d_2}, \mathtt{B_1} \cap \mathtt{B_2})$.
- With $\sqcap_{\mathtt{H}}$, $(\mathtt{H}, \sqsubseteq)$ is the *direct product* [4] of the partial orders $(\mathtt{D}, \sqsubseteq)$ and $(\mathtt{P_r}, \subseteq)$ and thus, it is an ordered set itself.

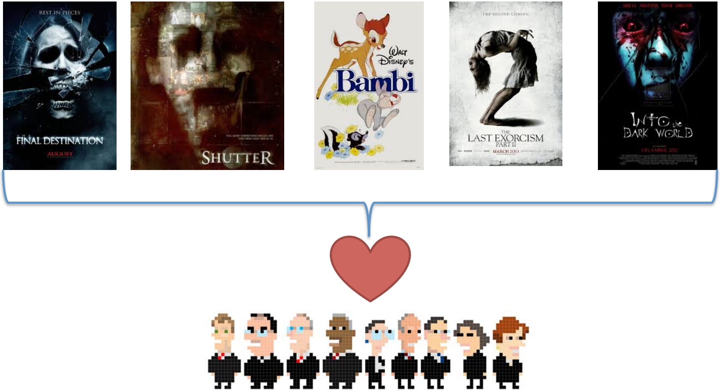
Heterogeneous pattern structures

Properties:

- The derivation operators $(\cdot)^\diamond$ in $(\mathtt{G}_1, (\mathtt{H}, \sqcap_\mathtt{H}), \Delta^\exists)$ for $\mathtt{A} \subseteq \mathtt{G}_1$ and $\mathtt{h} \in \mathtt{H}$ are defined as:
  - $\mathtt{A}^\diamond = \prod_{\mathtt{g}\in\mathtt{A}} {}_\mathtt{H}\, \Delta^\exists(\mathtt{g})$.
  - $\mathtt{h}^\diamond = \{\mathtt{g} \in \mathtt{G}_1 \iff \mathtt{h} \sqsubseteq \Delta^\exists(\mathtt{g})\}$.
- $(\mathtt{A}, \mathtt{h})$ is a heterogeneous pattern concept iff $\mathtt{h}^\diamond = \mathtt{A}$ and $\mathtt{A}^\diamond = \mathtt{h}$.
- $\mathtt{h}^\diamond = \mathtt{d}^\square \cap \mathtt{B}'$, where $()^\square$ is the derivation operator in $(\mathtt{G}_1, (\mathtt{D}, \sqcap), \delta)$.
- $\mathtt{A}^{\diamond\diamond} = \mathtt{A}^{\square\square} \cap \mathtt{A}''$.
- $\mathtt{h}^{\diamond\diamond} = (\mathtt{h}^{\diamond\square}, \mathtt{h}^{\diamond\prime})$.

A simple example

What does a heterogeneous concept represent? Consider the set of movies liked by a group of people.



A closed concept of movies liked by a certain group of people.

A simple example

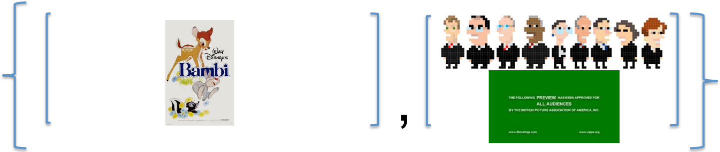
What does a heterogeneous concept represent? Consider the set of movies liked by a group of people.

In a different description space, these movies are separated.

A simple example

What does a heterogeneous concept represent? Consider the set of movies liked by a group of people.

$$A^{\diamond\diamond} = A^{\square\square} \cap A''. \qquad h^{\diamond\diamond} = (h^{\diamond\square}, h^{\diamond\prime}).$$



Heterogeneous concepts catches both spaces together.

Proposition

In our example:

- $(\mathtt{G}_1, (\mathtt{D}, \sqcap), \delta)$ is an interval pattern structure of documents described by convex regions in a LV space.
- $\mathcal{K}_2$ is a formal context of terms and taxonomical annotations (Wordnet synsets).
- $\mathtt{aw} : \mathtt{G}_1 \rightarrow \mathtt{G}_2$ relates documents with a set of annotations (terms).
- An heterogeneous pattern concept (hp-concept) $(\mathtt{A}, \mathtt{h})$ describes in its intent a relation between a convex region in the LV space and a set of taxonomical annotation.
- The set of all hp-concepts generates a set of “labelled clusters” in the LV space.

Proposition

In our example:

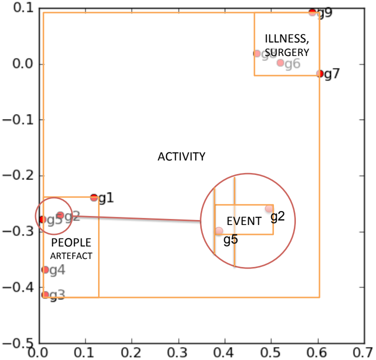


Figure: Labelled document clusters using association rules from the hp-lattice with magnification on documents g_2 and g_5.

3
Discussion

Discussion

Formal Context Constructions

- Direct sum:
  - $\mathcal{K}_1 + \mathcal{K}_2 = (\mathtt{G}_1 \cup \mathtt{G}_2, \mathtt{M}_1 \cup \mathtt{M}_2, \mathtt{I}_1 \cup \mathtt{I}_2 \cup (\mathtt{G}_1 \times \mathtt{M}_2) \cup (\mathtt{G}_2 \times \mathtt{M}_1))$.
- Semi-product:
  - $\mathcal{K}_1 \boxtimes \mathcal{K}_2 = (\mathtt{G}_1 \times \mathtt{G}_2, \mathtt{M}_1 \cup \mathtt{M}_2, \nabla)$.
  - $(\mathtt{g}_1, \mathtt{g}_2)\nabla(\mathtt{j}, \mathtt{m}) :\iff \mathtt{g}_\mathtt{j}\mathtt{I}_\mathtt{j}\mathtt{m}$, for $\mathtt{j} \in \{1, 2\}$
- Direct product:
  - $\mathcal{K}_1 \times \mathcal{K}_2 = (\mathtt{G}_1 \times \mathtt{G}_2, \mathtt{M}_1 \times \mathtt{M}_2, \nabla)$.
  - $(\mathtt{g}_1, \mathtt{g}_2)\nabla(\mathtt{m}_1, \mathtt{m}_2) :\iff \mathtt{g}_1\mathtt{I}_1\mathtt{m}_1$ or $\mathtt{g}_2\mathtt{I}_2\mathtt{m}_2$.
- Heterogeneous composition (if $\mathtt{G} = \mathtt{G}_1 = \mathtt{G}_2$):
  - $\mathcal{K}_1 \bowtie \mathcal{K}_2 = (\mathtt{G}, \mathtt{M}_1 \times \mathtt{M}_2, \nabla)$.
  - $\mathtt{g}\nabla(\mathtt{m}_1, \mathtt{m}_2) :\iff \mathtt{g}\mathtt{I}_1\mathtt{m}_1$ and $\mathtt{g}\mathtt{I}_2\mathtt{m}_2$.

4

Conclusions

Conclusions

Final remarks:

- Latent variable models can benefit from the capabilities of FCA by allowing an enriched description of otherwise, abstract characterizations.
- The flexibility of pattern structures allows for objects to be described by mixed representations, i.e. heterogeneous data.
- Relational Concept Analysis can be also applied on complex data by the use of heterogeneous pattern structures.
- The use of heterogeneous pattern structures may allow the implementation of FCA algorithms on further applications of data mining, such as high-order heterogeneous data co-clustering [5].

THE END

References I

[1] Blei, D. M., Ng, A. Y., and Jordan, M. I.
Latent dirichlet allocation.
Journal of Machine Learning Research 3 (2003), 993–1022.

[2] Deerwester, S., Dumais, S., and Furnas, G.
Indexing by latent semantic analysis.
Journal of the (1990).

[3] Ganter, B., and Kuznetsov, S. O.
Pattern Structures and their projections.
Conceptual Structures: Broadening the Base (2001).

[4] Ganter, B., and Wille, R.
Formal Concept Analysis: Mathematical Foundations.
Springer, Dec. 1999.

References II

[5] Gao, B., Liu, T.-Y., and Ma, W.-Y.

Star-structured high-order heterogeneous data co-clustering based on consistent information theory.

In *Data Mining, 2006. ICDM'06. Sixth International Conference on* (2006), IEEE, pp. 880–884.

[6] Hofmann, T.

Probabilistic latent semantic analysis.

In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence* (1999), Morgan Kaufmann Publishers Inc., pp. 289–296.

[7] Kaytoue, M., Kuznetsov, S. O., and Napoli, A.

Revisiting numerical pattern mining with formal concept analysis.

Proceedings of the Twenty-Second international joint conference on Artificial Intelligence - Volume Volume Two (Nov. 2011), 1342–1347.

References III

[8] Manning, C. D., Raghavan, P., and Schtze, H.
Introduction to Information Retrieval.
July 2008.

[9] Rouane-Hacene, M., Huchard, M., Napoli, A., and Valtchev, P.
Relational concept analysis: mining concept lattices from multi-relational data.
Annals of Mathematics and Artificial Intelligence 67, 1 (Mar. 2013), 81–108.

[10] Srivastava, A., and Sahami, M.
Text Mining: Classification, Clustering, and Applications, 1st ed.
Chapman & Hall/CRC, 2009.