HAL open science

Concept analysis-based association mining from linked data: A case in industrial decision making

Mickael Wajnberg, Petko Valtchev, Mario Lezoche, Alexandre Blondin Masse, Hervé Panetto

To cite this version:

Mickael Wajnberg, Petko Valtchev, Mario Lezoche, Alexandre Blondin Masse, Hervé Panetto. Concept analysis-based association mining from linked data: A case in industrial decision making. 2nd International Workshop on Data meets Applied Ontologies in Open Science and Innovation, DAO-SI 2019, Sep 2019, Gratz, Austria. hal-02455243

HAL Id: hal-02455243

https://hal.science/hal-02455243v1

Submitted on 25 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

CEUR-WS.org/Vol-2518/paper-DAOSI3.pdf

Concept Analysis-Based Association Mining From Linked Data: A Case In Industrial Decision Making

Mickael WAJNBERG [a,b], Petko VALTCHEV [a], Mario LEZOCHE [b], Alexandre BLONDIN MASSÉ [a], and Hervé PANETTO [b]

[a] *Université du Québec À Montréal, Canada*

[b] *Université de Lorraine, CNRS, CRAN, Nancy, France*

Abstract. Linked data (LD) is a rich format increasingly exploited in knowledge discovery from data (KDD). To that end, LD is typically structured as graph, but can also fit the multi-relational data mining (MRDM) paradigm, e.g. as multiple types and object properties may be used in the dataset. Formal concept analysis (FCA) has been successfully used as theoretical framework for KDD in a variety of applications, primely in clustering and association rule mining (ARM) tasks. As FCA applicability to LD is limited by its single data table input format, relational concept analysis (RCA) was introduced as a MRDM extension that successfully deals with links in the data, including cyclic ones. While RCA has been mainly adapted for conceptual clustering in the past, we present here an RCA-based ARM method. It exploits the iterative nature of pattern generation to cut cyclic references with a minimal loss of information. The utility of the rules discovered by our method has been validated by an application as a decision support in the aluminum die casting industry.

Keywords. Linked Data, Association Rule Mining, Relational Concept Analysis, Knowledge Management, Industrial Process, Decision Making

1. Introduction

Relational data, i.e. data that comprise both proper features of the represented objects and links between them, is ever more widely used in knowledge discovery from data (KDD). As a special case, linked data (LD) offers arguably the richest semantics, with the linked open data (LOD) cloud and other knowledge graphs among the largest such datasets. As with other massive datasets, decision makers and experts often turn to these data in a search for insights about the underlying phenomena, processes or concepts of the corresponding domain. By covering multiple entity types, LD allows the automated discovery of complex patterns and trends across related types and their links making them particularly valuable, as challenging. While such data naturally form graphs and thus can be targeted with graph mining methods [24,14], a more structured view thereof exists, i.e. as a set of inter-related data tables. The need to properly analyze such data motivated the emergence of multi-relational data mining (MRDM) [18,12].

Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

We are interested in contexts where MRDM supports experts in gaining better understanding of the domain behind the data, e.g. in the design of high-level domain models such as ontologies or causal models, and study a case pertaining to industrial process optimization. A key requirement is intelligibility of mining output to human experts. Moreover, our target data are typically unlabeled, thus they warrant descriptive mining mode [20] as in association discovery or clustering.

Formal concept analysis (FCA) [22] is a mathematically-founded analysis approach that has been successfully applied as a framework for KDD [21]. It reveals the hidden conceptual structure behind an (object x attribute) dataset in the form of a lattice of conceptual abstractions (a conceptual clustering). Alternatively, FCA extracts various families of non redundant association rules [10], based on notions of closure and generator [19]. This makes FCA a particularly versatile KDD framework. Another domain where FCA has been proved to be useful is fault-tolerance [2,15]. Recent work in fault localization crosschecks traces of correct and failing execution traces, it implicitly searches for association rules which indicate that most probably exists a cause for the whole execution to fail [3].

Relational concept analysis (RCA) is a recent extension of FCA that fits the MRDM paradigm, in general, and Linked data (LD), in particular [17]. It organizes multi-type datasets into a family of inter-related lattices, one per object type. To that end, object links are factored into concept definitions in the form of reference-like attributes, dynamically created by propositionalization. So far, RCA is missing effective tools for association discovery. Indeed, while in FCA associations are composed from parts of concept intents, relational intents might hide reference cycles, thus making the relationship between the premise and the conclusion part of a rule hard to interpret. However, if properly formalized, such rules could help uncover non trivial regularities spanning across multiple object types.

As a possible remedy, we designed a mechanism to untangle references in RCA, even in circular concept intents. It resolves concept references by a substitution with a appropriately truncated version of the underlying intent. The result enables the straightforward definition of an association as well as more advanced closure and generator-based ones [19]. The latter offer a valuable trade-off of generalization capacity vs result size that is beyond the reach of existing LD-compatible mining methods such as pure graph pattern miners [19] or association miners for LD relying on logical pattern languages [9] or on flattening techniques [13]. We applied our method to industrial production data, i.e. a dataset describing handles and frames for doors and windows and possible issues with their usage. The overall goal of our study is to assist the expert design of a high-level casual model to support decisions about optimizing the production process. At a first step, our methods was to help the search for causality links between production factors and product anomalies. Thus, it was set up to abstract associations between discovered machined part concepts, on one hand, and observed problem ones, on the other hand, which were then to be submitted to the expert. First experimental results show that our method succeeds in detecting non trivial patterns previously unknown to the expert.

The remainder of the paper is as follows: Section 2 motivates our study while section 3 provides background on FCA, RCA and association rules. Sections 4 and 5 describe our association rule approach and the experiment with production data, respectively. Section 6 summarizes then related work and section 7 concludes.

2. Motivation

A new industrial paradigm is emerging, the first of which is the gradual disappearance of boundaries between industry (in the traditional sense) and services, which are brought to cohabit. The data becomes raw materials of strategic importance. The industry of the future places great importance on non-price competitiveness as a success factor for companies on the markets, which requires a change in the offer, its design. The new industrial world is made up of intertwining knowledge, skills and competences. Today, thanks to the progress made in the areas of digitalization, information and communication technologies (ICT), the industrial sector is developing quickly towards ”smarter manufacturing” and ”connected factories”. This fourth industrial revolution also ensures complementarity and collaboration between several scientific disciplines. They need decision support for maximum reactivity. The great challenge is to have good quality products while minimizing working time and reducing manufacturing costs. To meet these constraints, it is important to detect risk phenomena through continuous monitoring of the machining process. This requires an instrumentation of the machines to ensure the acquisition and analysis of the large amount of heterogeneous data and knowledge, representative of these processes. To ensure its management, it is essential to structure the data collected through models to classify the different elements of a machining process, coupled with the global context of production. These models must also include a classification of major failure factors and their causality links to process elements.

We study here a case of metal processing: handles and frames for doors and windows are manufactured by aluminum die casting. A four steps process transforms the aluminum, placed in the machine starting compartment, into machined parts: The machine heats to melt aluminum (1), then the liquid aluminum is sent to a pump (2), where a piston injects the aluminum into a mould (3) and finally the piston compresses the aluminum until solidification (4). Right now, given its size, the manufacturer considers the production machine control a too tedious and costly task to be done directly and preemptively. Instead, the control is achieved monitoring the products themselves : occasionally, non compliant parts will provoke failures, halting the production process to check and fix the machine. When such a problem occurs, the machined part that triggered the stop is highly likely to be recast (melted again and completely reprocessed). When the products are discarded for not compliance to the quality standards, the operators fixes the problem and restarts the production.

Understanding causality between machined parts variations and observed anomaly features can help anticipate potential failures, hence schedule a targeted machine control and avoid recasting a piece, which in the overall, reduce the costs. Therefore, applying forensic analysis to the collected data to discover associations between those is a particularly promising approach. Since available data is unlabeled and comprise many-to-many relations between two types of objects, a MRDM approach for association mining seems a natural choice.

3. Background on FCA and RCA

We recall results from FCA [22] and association rules [1], and then present RCA [17].

In FCA, data is introduced as a (formal) context, a triple $\mathcal{K} = (\mathcal{O}, \mathcal{A}, \mathcal{I})$, where $\mathcal{O}$ is a set of objects, $\mathcal{A}$ a set of attributes and $\mathcal{I} \subseteq \mathcal{O} \times \mathcal{A}$ is a binary relation ($(o,a) \in \mathcal{I}$ means

$\mathcal{K}_1$	A_0	A_1	A_2	A_3
O_0			$\times$	$\times$
O_1	$\times$	$\times$		
O_2	$\times$	$\times$	$\times$	
O_3		$\times$	$\times$	

Figure 1. Formal Context $\mathcal{K}_1$ Example

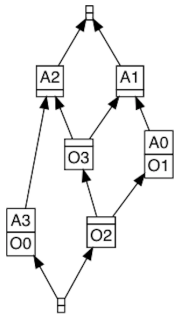

Figure 2. $\mathcal{K}_1$ Lattice

object o has attribute a). A context is drawn as (object x attribute) data table as in Fig. 1. Two derivation operators, both denoted $'$, connect $\mathcal{O}$ and $\mathcal{A}$ both ways: $O' \subseteq \mathcal{A}$ is the set of all attributes shared by every $o \in O$ (ex : $\{O_1, O_2\}' = \{A_0, A_1\}$) and $A' \subseteq \mathcal{O}$ works dually (e.g. $\{A_2, A_3\}' = \{O_0\}$). A (formal) *concept* is a pair $(O, A) \subseteq \mathcal{O} \times \mathcal{A}$ s.t. $O = A'$ and $A = O'$. Thus, a concept is a maximal pair as both its components are maximally extended w.r.t. each-other. In a concept (O, A), O is called its *extent* and A its *intent*. Moreover, both compound derivation operators $'\circ'$, denoted $''$, are closure operators on $\wp(\mathcal{O})$ and $\wp(\mathcal{A})$, respectively. For instance, $\{A_0, A_2\}'' = \{A_0, A_1, A_2\}$.

FCA extracts all the concepts of a context and organizes them into a lattice w.r.t. extent inclusion (see Fig. 2). The lattice can be used to generate the association rules of the context. These have the form $M \rightarrow N$ where $M, N \subseteq \mathcal{A}$. Typical quality measures for association rule include support and confidence, i.e. the percentage of objects incident to attributes in $M \cup N$ in $\mathcal{O}$, and the proportion of the previous objects in the larger set of objects incident to attributes in M, respectively [1]. For instance, we have from figure 1 the following association rule $A_1, A_2 \rightarrow A_0$. The support of the rule is 25%: only one object out of the set of all objects in the base (four) hold these three attributes; while the confidence is 50%: the number of objects holding the three attributes (O_2) out of the numbers of objects holding the A_1, A_2 of the rule (O_2, O_3).

We are interested in exact associations, a.k.a. *implication rules*, which have confidence of 100%. These rules admit no exceptions as opposed to general associations whose confidence may be lower. While exact rules are only a small part of all possible rules (of certain support and confidence levels), their number is still prohibitive. A much smaller implication subfamily is defined using the notion of equivalence class on $\wp(\mathcal{A})$ and generators, which still withholds the information encoded in the entire family.

Equivalence classes are induced on both powersets of a context by $''$: two sets A and B are equivalent if they share the same closure, *i.e.* $A'' = B''$. Each equivalence classes on has a unique maximum, e.g. in $\wp(\mathcal{A})$, it is a concept intent. Conversely, there are one or more minimal elements, called *generators*. Now, the most informative implication rules have the form $G \rightarrow M \setminus G$ where M is an intent and G one of its generators [10]. FCA methods exist [19] allowing these rules to be efficiently generated.

FCA reveals conceptual abstractions on objects by factoring out shared attributes. RCA [17] extends it by enabling the factoring in of relational information. Its input format is thus a relational context family which is made of a set of contexts $\mathcal{K}_i = (\mathcal{O}_i, \mathcal{A}_i, \mathcal{I}_i)$ and a set of binary relations $r_{i,j} \subseteq \mathcal{O}_i \times \mathcal{O}_j$ linking a pair of context $\mathcal{K}_i$ (domain) and $\mathcal{K}_j$ (range).

$\mathcal{K}_1$	*sko*	*cst*	*smL*	*tcL*
12			×	×
13	×	×		
14	×	×	×	
15		×	×	

g	P_1	P_2	P_3
12	×		×
13	×		
14		×	×
15			×

$\mathcal{K}_2$	*t5*	*stp*	*qlt*	*mld*	*cost*
P_1	×	×		×	×
P_2	×		×		×
P_3		×	×	×	

Figure 3. Relational Context Family Example

Consider the context family in Fig. 3: Context $\mathcal{K}_1$, on the left-hand side, represents data on machined parts, i.e. objects 12, 13, 14 and 15. The respective attribute set comprises inadequate thickness (*sko*), part to be melted again and recast (*cst*), thickness under lower threshold (*smL*) and pressure under the lower threshold (*tcL*). $\mathcal{K}_2$, on the right-hand side, gathers data on problems (objects P_1, P_2, and P_3). Its attributes areas follows: time to solve less than 5 min (*t*5), machine stopped (*stp*), quality problem (*qlt*), mould defect (*mld*), and medium financial impact (*cost*).

A relation g, for 'generates', links machined part to generated problems (given in the center of Fig. 3). While a relation is much alike a formal context, there is a key difference: While in a context, rows and columns correspond to objects and attributes, respectively, in the cross-table of a relation they both represent objects, the ones from the domain context (rows) and from the range context (columns).

To become shareable, inter-object links are transformed into new *relational* attributes of the contexts in the family (actually of extended version thereof). The corresponding processing, called *relational scaling*, uses abstractions from the range context $\mathcal{K}_j$ as targets for relational attributes in $\mathcal{K}_i$ (one per abstraction). Absent a prior hierarchy of such abstractions on $\mathcal{K}_i$, the RCA method builds and then exploits the concept lattice of that context. A typical attribute over a relation r is structured much alike a role restriction in description logics (DL), i.e. as $qr : c$ where q is a quantifier ($\{\exists, \forall, \forall\exists, \ldots\}$) and c is a concept over $\mathcal{K}_j$. The way $\mathcal{I}_i$ is extended for each new $qr : c$ depends on q and c (see [17] for details).

The overall RCA method is iterative: Each iteration alternates lattice construction with scaling steps. After each construction, the versions of the contexts are enhanced with additional attributes created by scaling over the differential set of concepts w.r.t. to the previous iteration. The global set of concepts is monotonously non decreasing since adding new attributes to a context preserves all its concepts (as far as their extents are concerned) while possibly creating new ones. Thus, the global iterative process necessarily ends at a fix-point [17] which corresponds to a family of relational lattices with a number of inter-concept links via relational attributes. Eventually, these links appear in the Hasse diagrams of the fix-point lattices, as seen in Fig. 4 and Fig. 5.

4. RCA-Based Knowledge Discovery

RCA discovers conceptual abstractions inherent to the data not originally present in the data schema[19]. Moreover, its iterative computing method allows further abstractions to be discovered on top of previously generated ones. It is also worth mentioning that RCA can produce rich relations thanks to the diverse quantifiers that can be used and

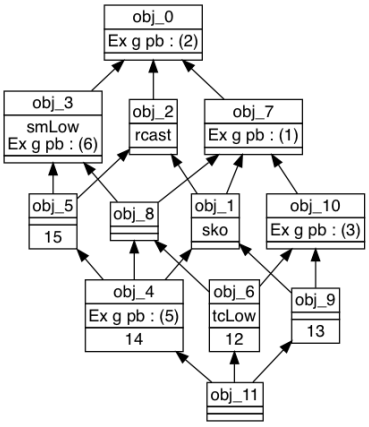

Figure 4. Machined Parts Lattice

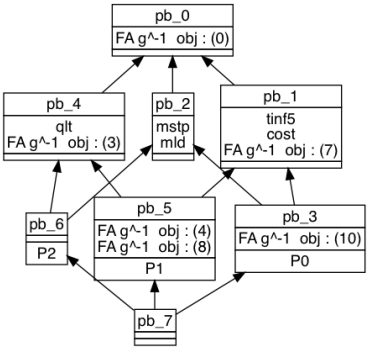

Figure 5. Problems Lattice

combined altogether in a given association rule. The down-side of such an expressive format is some RCA intents, and thus the corresponding rules, might present terminological cycles, which makes them hard to interpret. For instance, in Fig. 4 and Fig. 5 the concept obj_2 will yield the rule $rcast \rightarrow \exists g : pb(2)$ which could be interpreted by "if a part presents the attribute *rcast*, then there exists a related problem described by concept pb_2". If the description of pb_2 is inserted to resolve the reference, that would yield "if a part present attribute rcast then there exists a co-occuring problem that has mstp mld and $\forall g^{-1} : obj(0)$". However, $\forall g^{-1} : obj(0)$ refers to a concept whose attribute refers back to pb_2. Therefore the interpretation would be entrapped in a cycle.

We designed a technique to disentangle the description of a relational concept intent (subject of a separate publication). It exploits the iterative nature of the generations process which is inherently cycle-free : at the creation of a node, this one can only have two types of attributes : non relational ones and relational that refer to a node that was created in a previous iteration. Hence, in a concept any relational attribute $\rho r : C$ targeting a concept C by the relation r and the scaling operator ρ can be interpreted in $\rho r : (A_1, ..., A_n)$ where $(A_1, ..., A_n)$ is the intent of C at its creation. Recursively, if any attribute of A_i is relational it is replaced by its interpretation.

Using this principle, we can generate lattices without references as in figure 6 from which we can extract immediately interpretable rules such as :

$$\forall g^{-1} obj : (sko, cst, smL) \rightarrow \forall g^{-1} obj : (\exists g\ pb : (t5, cost), \exists g\ pb : (stp, mld))$$

This states that any problem generated by all machine parts that are sko, cst and smL are also generated by machine parts that generates a t5 and cost problem or an stp and mld problem.

In the next section, we discuss the results obtained by applying RCA on an industrial relational dataset linking produces machined parts to generated problems.

5. Experimental Study

We present below our study, dataset and experimental settings, and discuss its outcome.

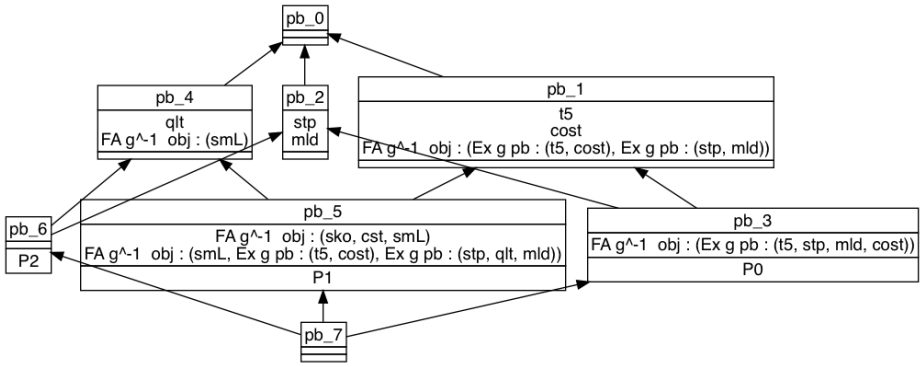

Figure 6. Lattice of observed problems

5.1. Dataset

We started our experimental study with a dataset that covers a single month of production. It comprises three sorts of triples: Some triples represent machined parts and their properties, while others describe observed problems. The remaining triples in the dataset describe the relationships between parts and the generated problems, if any.

Overall, the dataset covers 58000 parts described by 25 features. These include date and shift work of the manufacturing, machined part type, its thickness, volume and conformity, also position, speed and pressure of the piston in the pump during the three phases of the process (respectively when aluminum is not in the mould, when it is entering the mould and when it is compressed after filling the mould). Other features are test results on the part w.r.t low and high thresholds on the mentioned features.

Next, 16 problems are tracked. They are described by 23 features divided into three categories : financial losses incurred, duration until resolution, and problem's nature (quality of the piece, mechanical wear, machine calibration...).

5.2. Experimental Settings

In order to find correlations between machined parts characteristics, i.e. categories of faulty parts, generic types of problems (rather than individual problems), analyzing the co-occurrences of those in the dataset seems a natural starting point. Thus, RCA association rules constitute a reasonable first target.

We set up a relational context family made of two contexts, one for machined parts and one for problems, and a single 'generates' relation (with its reverse relation being 'is generated by') that links both contexts. A preprocessing step encodes input triples, as described above, into cross tables: It splits each categorical feature over multiple attributes, one per possible mode, while numerical ones are split over a set of intervals. In our case, the machined parts context contains 160 attributes and the problem one 28.

In the first stage of our study, industrial partners were looking for insights about how product thickness relates to quality problems, as well as the link between these problems and piston course. Thus, we focused on rules associating machined part features to the thickness or, alternatively, piston course to generic problem categories.

Based on the need to extract the occurrence of any problem for a specific machined parts characteristic combination, the existential scaling operator was chosen. Also, we

only kept a one-way interpretation of the above relation, namely *generates* since the need was to link part characteristics to problem, and not to link part characteristics to others parts that would have generated an equivalent problem.

Finally, in a post-processing step, we selected the implication rules which contained only thickness or piston characteristics for machined parts.

5.3. Experimental Outcome

Below, we illustrate the result extracted by RCA: We provide a small number of rules and explain their significance.

5.3.1. Invalid Thickness

Among the output rules we found the following set related to a single concept :

$$\begin{cases} (1)\ \text{min-sm}\ (<\text{LimLow}) \Rightarrow \text{sm-1-ko} \\ (2)\ \text{sm-1-ko} \Rightarrow \text{min-sm}\ (<\text{LimLow}) \\ (3)\ \text{sm-1-ko} \Rightarrow \exists\text{generates problems : (micro-stop, sm-quality, time}(<5\text{min}), \text{cost-medium)} \\ (4)\ \text{sm-1-ko} \Rightarrow \text{recasting} \end{cases}$$

The above set can be interpreted as follows: The definition for a product to have defective thickness (sm-1-ko) is that the thickness of the product is below the lower threshold limit (1 and 2) and so, will be (4) recast (melt again and reprocessed completely). When that happens, (3) a micro-stop problem about thickness quality (sm-quality), that took less than < 5 min to solve also occurs. This problem created a medium financial impact (cost-medium).

5.3.2. Invalid Piston Course

Piston is monitored through three variables: C1, its position when aluminum gets from the injector to the mould entrance, C2, its position when aluminum just filled the mould and CC, its position after compression. Among rules in the RCA output, we detected several independent sets of rules following the template below:

$$\begin{cases} \text{C1=X, CC=Y} \Rightarrow \text{C2=Z} \\ \text{C2=Z, CC=Y} \Rightarrow \text{C1=X} \\ \text{C1=X, C2=Z} \Rightarrow \text{CC=Y} \\ \text{C1=X, C2=Z} \Rightarrow \exists\ \text{generates problems : (sm=14, predictive_1, quality)} \end{cases}$$

Each occurrence of the above template carries its own combination of X, Y and Z values. However, experts noticed that a linear combination of the values is quasi-invariant.

5.3.3. Impact

The domain expert concluded that the above set of rules on thickness reflected the presence of hardened aluminum residue in the machine. Such residue grows with time while reducing the space available for the part in production and altering the piston course. It typically leads to costly machine stoppages as the produced parts end up violating the

admissible variations and have to be recast. As a result of this observation, a reduced cost and less time consuming solution could be found : calibrating sensors to respond automatically to the piston course and send alerts to operators to clean the machine in time before producing a faulty part.

6. Related Work

Traditional data mining methods have been designed to work on a single table of a database. All of them can be applied to a LD dataset provided a prior flattening, or feature extraction, is performed to bring it down to that format [16].

Logic-based LD miners such as DL-Learner [11] avoid flattening. DL-Learner is a supervised concept learner for DL classes, hence it only discovers plausible descriptions for *existing* abstractions already in the schema. RCA also borrows attribute constructors from DL, yet, unlike DL-Learner, its discovery process is extension-bound: It stops when no new extent are discovered (at a fix-point). DL-based association miners for LD such as [9] are also intentionally guided. Moreover, in the wider field of MRDM [7], a logic-based association rule miners is described in [5]. An ad-hoc approach, called contextual association rule mining, is proposed in [6] that amounts to dealing with multi-relational datasets by focusing exclusively on pairs of entity types and a single relation type. As the notion of closure and generator are missed in all these approaches, their output is inherently redundant. Pure pattern miners in the field are presented in [4] (as valid/relevant queries in DATALOG) and in [18].

Alternatively, associations can be mined from LD seen as graphs using classical graph pattern mining algorithms. Among those admitting labels on both vertices and edges, a necessary condition to fit the LD format, arguably the most widely used are gSpan [24] and Gaston [14]. Yet graph pattern offer a single view of the shared substructure between data objects: They lack the expressive power of logic-based associations, e.g. miss the equivalent of a universal quantifier (as seen in Fig. 6).

Beside RCA, alternative attempts at extending FCA to graph data are proposed in [8, 23]. However, they do not seem to easily adapt to association rule mining.

7. Conclusion

RCA combines the mathematical rigor of FCA with expressive power of linked data to support KDD in MRDM mode. Scaling creates relational attributes in RCA that abstract from inter-object links and then factors in those attributes into concept intents. In this paper, we show how associations can be defined in RCA. To bring those to a easily-readable –and explainable– form, we resolve concept references in attributes using intent substitution. Moreover, to avoid the pit of circular references, only minimal generating parts of the intents are used in substitution.

To demonstrate the utility of the discovered relational associations, we applied our method to a dataset representing machined parts together with observed problems. Preliminary results are encouraging: Some rules revealing non trivial relationships between respective features of parts and problems were rated as highly valuable by the expert. A larger evaluation effort involving more experts and a variety of datasets is ongoing.

At a next stage, we plan to test various scenarios of encoding relations in RCA: bi-directional parts-to-problems, mixing several quantifiers, etc. In the longer run, we plan to compare RCA association rules to those discovered by competing LD miners.

References

[1] R. Agrawal et al. Fast discovery of association rules. *Advances in knowledge discovery and data mining*, 12(1):307–328, 1996.

[2] J. Besson et al. Mining a new fault-tolerant pattern type as an alternative to formal concept discovery. In *International Conference on Conceptual Structures*, pages 144–157. Springer, 2006.

[3] P. Cellier et al. Formal concept analysis enhances fault localization in software. In *International Conference on Formal Concept Analysis*, pages 273–288. Springer, 2008.

[4] L. Dehaspe and H. Toivonen. Discovery of frequent datalog patterns. *Data Mining and knowledge discovery*, 3(1):7–36, 1999.

[5] L. Dehaspe and H. Toivonen. Discovery of relational association rules. In *Relational data mining*, pages 189–212. Springer, 2001.

[6] W. Dong et al. A general framework to encode heterogeneous information sources for contextual pattern mining. In *21st ACM Intl. Conf. on Information and Knowledge Management*, pages 65–74, 2012.

[7] S. Džeroski. Multi-relational data mining: an introduction. *ACM SIGKDD Explorations Newsletter*, 5(1):1–16, 2003.

[8] S. Ferré. A proposal for extending formal concept analysis to knowledge graphs. In *International Conference on Formal Concept Analysis*, pages 271–286. Springer, 2015.

[9] J. Józefowska et al. On reducing redundancy in mining relational association rules from the semantic web. In *International Conference on Web Reasoning and Rule Systems*, pages 205–213. Springer, 2008.

[10] M. Kryszkiewicz. Concise Representations of Association Rules. In *Pattern Detection and Discovery*, volume 2447, pages 92–109. Springer Berlin Heidelberg, 2002.

[11] J. Lehmann. Dl-learner: learning concepts in description logics. *Journal of Machine Learning Research*, 10(Nov):2639–2642, 2009.

[12] J. Lijffijt et al. Pn-rminer: A generic framework for mining interesting structured relational patterns. *International Journal of Data Science and Analytics*, 1(1):61–76, 2016.

[13] V. Narasimha et al. Liddm: A data mining system for linked data. In *Workshop on Linked Data on the Web. CEUR Workshop Proceedings*, volume 813, page 108, 2011.

[14] S. Nijssen and N. Kok. Frequent graph mining and its application to molecular databases. In *2004 IEEE International Conference on Systems, Man and Cybernetics*, volume 5, pages 4571–4577. IEEE, 2004.

[15] R. Pensa and J.-F. Boulicaut. Towards fault-tolerant formal concept analysis. In *Congress of the Italian Association for Artificial Intelligence*, pages 212–223. Springer, 2005.

[16] P. Ristoski et al. Mining the Web of Linked Data with RapidMiner. *Web Semantics: Science, Services and Agents on the World Wide Web*, 35:142–151, 2015.

[17] M. Rouane-Hacene et al. Relational concept analysis: mining concept lattices from multi-relational data. *Annals of Mathematics and Artificial Intelligence*, 67(1):81–108, 2013.

[18] E. Spyropoulou and T. De Bie. Interesting multi-relational patterns. In *Data Mining (ICDM), 2011 IEEE 11th International Conference on*, pages 675–684. IEEE, 2011.

[19] L. Szathmary et al. A fast compound algorithm for mining generators, closed itemsets, and computing links between equivalence classes. *Annals of Maths and Artificial Intelligence*, 70(1):81–105, 2014.

[20] P.-N. Tan et al. *Introduction to Data Mining*. Pearson Education Limited, 2019.

[21] P. Valtchev et al. Formal Concept Analysis for Knowledge Discovery and Data Mining: The New Challenges. In *Proc. of ICFCA 2004*, volume 2961 of *LNCS*, pages 352–371, 2004.

[22] R. Wille. *Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts*, pages 445–470. Springer Netherlands, Dordrecht, 1982.

[23] R. Wille. Existential concept graphs of power context families. In *International Conference on Conceptual Structures*, pages 382–395. Springer, 2002.

[24] X. Yan and J. Han. gSpan: Graph-based substructure pattern mining. In *2002 IEEE International Conference on Data Mining, 2002. Proceedings.*, pages 721–724. IEEE, 2002.