

Formal Concept Analysis in Knowledge Discovery: A Survey

Jonas Poelmans¹, Paul Elzinga³, Stijn Viaene^{1,2}, and Guido Dedene^{1,4}

¹ K.U.Leuven, Faculty of Business and Economics, Naamsestraat 69,
3000 Leuven, Belgium

² Vlerick Leuven Gent Management School, Vlamingenstraat 83,
3000 Leuven, Belgium

³ Amsterdam-Amstelland Police, James Wattstraat 84,
1000 CG Amsterdam, The Netherlands

⁴ Universiteit van Amsterdam Business School, Roetersstraat 11
1018 WB Amsterdam, The Netherlands

{Jonas.Poelmans, Stijn.Viaene, Guido.Dedene}@econ.kuleuven.be,
Paul.Elzinga@amsterdam.politie.nl

Abstract. In this paper, we analyze the literature on Formal Concept Analysis (FCA) using FCA. We collected 702 papers published between 2003-2009 mentioning Formal Concept Analysis in the abstract. We developed a knowledge browsing environment to support our literature analysis process. The pdf-files containing the papers were converted to plain text and indexed by Lucene using a thesaurus containing terms related to FCA research. We use the visualization capabilities of FCA to explore the literature, to discover and conceptually represent the main research topics in the FCA community. As a case study, we zoom in on the 140 papers on using FCA in knowledge discovery and data mining and give an extensive overview of the contents of this literature.

Keywords: Formal Concept Analysis (FCA), knowledge discovery, text mining, exploratory data analysis, systematic literature overview.

1 Introduction

Formal Concept Analysis (FCA) was invented in the early 1980s by Rudolf Wille as a mathematical theory (Wille, 1982). FCA is concerned with the formalization of concepts and conceptual thinking and has been applied in many disciplines such as software engineering, knowledge discovery and information retrieval during the last 15 years. The mathematical foundation of FCA is described by Ganter (1999).

A textual overview of part of the literature published until the year 2004 on the mathematical and philosophical background of FCA, some of the applications of FCA in the information retrieval and knowledge discovery field and in logic and AI was given in Priss (2006). An overview of available FCA software is provided by Tilley (2004). Carpineto (2004) provides an overview of FCA applications in information

retrieval. In Tilley (2007), an overview of 47 FCA based software engineering papers was given. The authors categorized these papers according to the 10 categories as defined in the ISO 12207 software engineering standard and visualize them in a concept lattice. In Lakhali et al. (2005), a survey on FCA-based association rule mining techniques is given.

In this paper, we describe how we used FCA to create a visual overview of the existing literature on concept analysis published between the years 2003 and 2009. The core contributions of this paper are as follows. We visually represent the literature on FCA using concept lattices, in which the objects are the scientific papers and the attributes are the relevant terms available in the title, keywords and abstract of the papers. We developed a toolset with a central FCA component that we use to index the papers with a thesaurus containing terms related to FCA research and to generate the lattices. As a case study, we zoom in on the 140 papers published between 2003 and 2009 on using FCA in knowledge discovery and data mining.

The remainder of this paper is composed as follows. In section 2 we introduce the essentials of FCA theory and the knowledge browsing environment we developed to support this literature analysis. In section 3 we describe the dataset used. In section 4 we visualize the literature using FCA lattices and we summarize the papers published in the knowledge discovery field. Section 5 concludes the paper.

2 Formal Concept Analysis

Formal Concept Analysis can be used as an unsupervised clustering technique. The starting point of the analysis is a database table consisting of rows G (i.e. objects), columns M (i.e. attributes) and crosses $I \subseteq G \times M$ (i.e. relationships between objects and attributes). The mathematical structure used to reference such a cross table is called a formal context (G, M, I) . In this paper, scientific papers (i.e. the objects) are related (i.e. the crosses) to a number of terms (i.e. the attributes); here a paper is related to a term if the title or abstract of the paper contains this term. Scientific papers containing terms from the same term-clusters are grouped in concepts. Given a formal context, FCA then derives all concepts from this context and orders them according to a subconcept-superconcept relation. This results in a line diagram (a.k.a. lattice). The details of using FCA for text mining can be found in (Poelmans 2009).

2.1 FCA Software

We developed a knowledge browsing environment to support our literature analysis process. One of the central components of our text mining environment is the thesaurus containing the collection of terms describing the different FCA-related research topics. The initial thesaurus was constructed based on expert prior knowledge and was incrementally improved by analyzing the concept gaps and anomalies in the resulting lattices. The thesaurus is a layered thesaurus containing multiple abstraction levels. The first and finest level of granularity contains the search terms of which most are grouped together based on their semantical meaning to form the term clusters at the second level of granularity.

The papers that were downloaded from the World Wide Web (WWW) were all formatted in pdf. These pdf files were converted to ordinary text and the abstract, title and keywords were extracted. The open source tool Lucene was used to index the extracted parts of the papers using the thesaurus. The result was a cross table describing the relationships between the papers and the term clusters or research topics from the thesaurus. This cross table was used as a basis to generate the lattices.

3 Dataset

This Systematic Literature Review (SLR) has been carried out by considering a total of 702 papers related to FCA published between 2003 and 2009 in the literature and extracted from the most relevant scientific sources. The sources that were used in the search for primary studies contain the work published in those journals, conferences and workshops which are of recognized quality within the research community. These sources are: *IEEE Computer Society*, *ACM Digital Library*, *Scencedirect*, *Springerlink*, *EBSCOhost*, *Google Scholar*, *Repositories of the ICFCA*, *ICCS and CLS conference*.

Other important sources such as *DBLP* or *CiteSeer* were not explicitly included since they were indexed by some of the mentioned sources (e.g. *Google Scholar*). In the selected sources we used various search strings including "Formal Concept Analysis", "FCA", "concept lattices", "Temporal Concept Analysis". We ensured that papers that appeared in multiple sources were only taken into account once.

4 Studying the Literature Using FCA

The 702 papers are grouped together according to a number of features within the scope of FCA research. The FCA lattices facilitate our exploration and analysis of the literature. The lattice in Fig. 1 contains 8 categories under which 53% of the 702 FCA papers can be categorized. Knowledge discovery is the most popular research theme covering 20% of the papers and will be analyzed in detail in section 4.1. Recently, improving the scalability of FCA to larger and complex datasets emerged as a new research topic covering 5% of the 702 FCA papers. In particular, we note that almost half of the papers dedicated to this topic work on issues in the KDD domain. Another important research topic in the FCA community is information retrieval covering 15% of the papers. 25 of the papers on information retrieval describe a combination with KDD approach and in 18 IR papers authors make use of ontologies. 15 IR papers deal with the retrieval of software structures such as software components. In 13% of the FCA papers, FCA is used in combination with ontologies or for ontology engineering. In 11% of the papers, the extension of FCA to deal with fuzzy attributes is investigated. The temporal variant of FCA received only minor attention, covering 1% of the papers. Other important topics are using FCA in software engineering (15%) and for classification (7%).

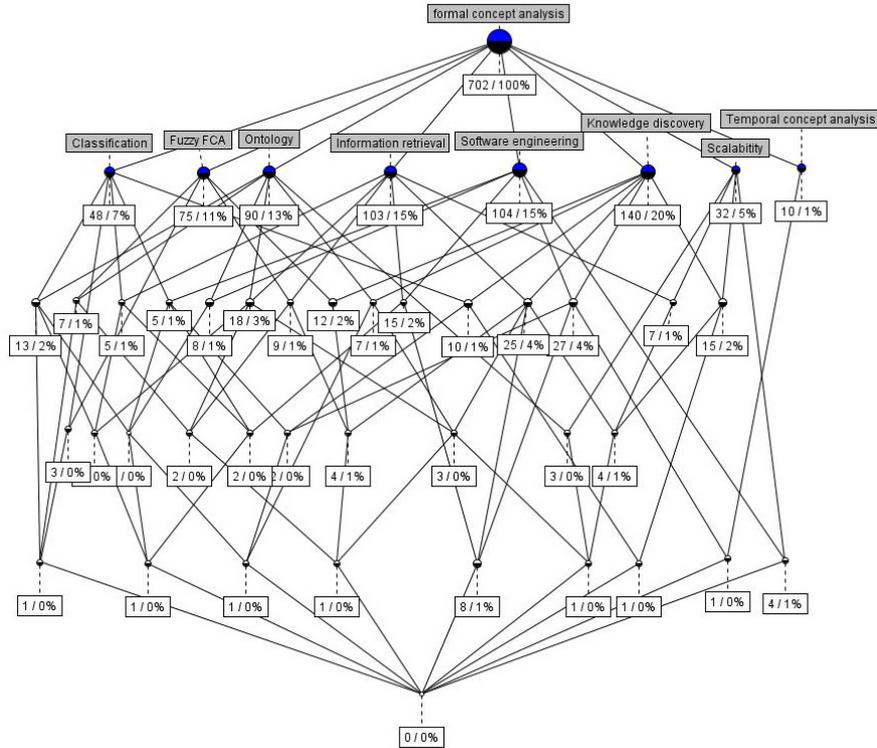


Fig. 1. Lattice containing 702 papers on FCA

4.1 Knowledge Discovery and Data Mining

In the past, the focus in knowledge discovery and data mining was on developing fully automated tools and techniques that extract new knowledge from data. Unfortunately, these techniques allowed almost no interaction between the human actor and the tool and failed at incorporating valuable expert knowledge into the discovery process (Keim 2002), which is needed to go beyond uncovering the fool's gold. These techniques assume a clear definition of the concepts available in the underlying data which is often not the case. Visual data exploration (Eidenberger 2004) and visual analytics (Thomas 2005) are especially useful when little is known about the data and exploration goals are vague. Since the user is directly involved in the exploration process, shifting and adjusting the exploration goals is automatically done if necessary.

In Conceptual Knowledge Processing (CKP) the focus lies on developing methods for processing information and knowledge which stimulate conscious reflection, discursive argumentation and human communication (Wille 2006). An important subfield of CKP is Conceptual Knowledge Discovery. FCA is particularly suited for exploratory data analysis because of its human-centeredness (Correia 2003). The generation of knowledge is promoted by the FCA representation that makes the

inherent logical structure of the information transparent. The system TOSCANA has been used as a knowledge discovery tool in various research and commercial projects (Stumme 1998).

About 74% of the FCA papers on KDD are covered by the research topics in Figure 2. In section 4.1.1 we zoom in on the 35 papers (25%) in the field of association rule mining. 19% of the KDD papers are on using FCA in the discovery of structures in software and will be described in section 4.1.2. Section 4.1.3 describes the 9% of papers on applications of FCA in web mining. Section 4.1.4 discusses some of the extensions of FCA theory for knowledge discovery (11% of papers). In section 4.1.5 we describe some of the applications of FCA in biology, chemistry and medicine covering 10% of the KDD papers. The applications on using Fuzzy FCA for KDD, covering 9% of the papers, will be discussed in section 4.1.6.

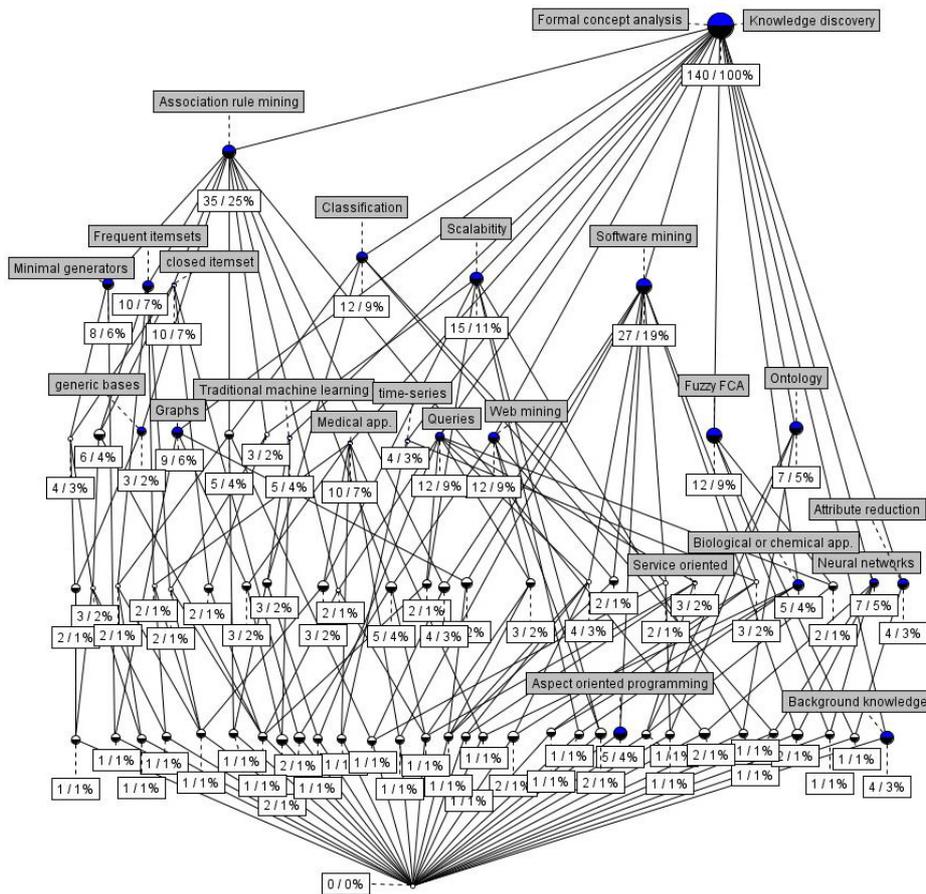


Fig. 2. Lattice containing 140 papers on using FCA in KDD

4.1.1 Association Rule Mining

Association rule mining covers 25% of the papers using FCA for KDD. Association rule mining from a transaction database requires the detection of frequently occurring patterns called frequent itemsets (FI). Recent approaches for FI mining use the closed itemset paradigm to limit the mining effort to the subset of frequent closed itemsets (FCIs). The intent of a concept C is called a closed itemset and consists of the maximum set of attributes that characterizes C . Several FCA-based algorithms were developed for mining frequent closed itemsets including CLOSE, PASCAL, CHARM, CLOSET and TITANIC (Stumme 2002) which mines frequent closed itemsets by constructing an iceberg concept lattice. In Qi et al (2004) an algorithm Closearcher is proposed based on FCA for mining frequent patterns.

The minimal generators for a concept C are the minimal subsets of C 's intent which can similarly characterize C . Nehmé (2005) proposes a novel method for computing the minimal generator family. Tekaya et al. (2005) propose an algorithm called GenAll to build an FCA lattice in which each concept is decorated by its minimal generators with the aim to derive generic bases of association rules. Generic bases constitute reduced sets of association rules and preserve the most relevant rules without loss of information. The GenAll algorithm further improves on the algorithm presented by Nourine et al. (1999). In Hamrouni, (2005), the extraction of reduced size generic bases of association rules is discussed to decrease the overwhelming number of association rules resulting from ARM. Hamrouni (2005a) proposes an algorithm called PRINCE which builds a minimal generator lattice from which the derivation of the generic association rules becomes straightforward. Dong et al. (2005) introduce the succinct system of minimal generators (SSMG) as a minimal representation of the minimal generators of all concepts, and gives an efficient algorithm for mining SSMGS. The SSMGS are also used for losslessly reducing the size of the representation of all minimal generators. Hamrouni et al. (2007) present a new sparseness measure for formal contexts using the framework of SSMGS. Their measure is an aggregation of two complementary measures, namely the succinctness and compactness measures of each equivalence class, induced by the closure operator. This is important for the performance of frequent closed itemset mining algorithms which is closely dependent on the type of handled extraction context, i.e. sparse or dense. Hermann et al. (2008) investigate the computational complexity of some of the problems related to generators of closed itemsets. The authors also present an incremental polynomial time algorithm that can be used for computing all minimal generators of an implication-closed set. In Yahia et al. (2004), inference axioms are presented for deriving all association rules from generic bases.

Valtchev et al. (2004), discuss the existing FCA-based data association rule mining techniques and provide guidelines for the design of novel ones to be able to apply FCA in a larger set of situations. They also propose two online methods for computing the minimal generators of a closure system. Gupta et al. (2005) discuss how classification rules based on association rules can be generated using concept lattices. Valtchev et al. (2008) show how FCIs can be mined incrementally yet efficiently whenever a new transaction is added to a database whose mining results are available. In Quan et al (2009), a new cluster-based method is proposed for mining conceptual association rules. Maddouri (2005) discusses the discovery of association rules and proposes a new approach to mine interesting itemsets as the

optimal concepts covering a binary table. Maddouri et al. (2006) summarizes many of the statistical measures introduced for selecting pertinent formal concepts. Maddouri et al (2009) present a method for building only a part of the lattice including the best concepts, which are used as classification rules.

Wollbold et al (2008) make use of FCA to construct a knowledge base consisting of a set of rules such that reasoning over temporal dependencies within gene regulatory networks is possible. Zhou et al (2005) use FCA to mine association rules from web logs, which can be used for online applications such as web recommendation and personalization. Richards et al. (2003) explores the possibilities of using FCA for mining knowledge and reorganizing this knowledge into an abstraction hierarchy and to discover higher-level concepts in the knowledge. Richards et al (2003a) discuss the discovery of multi-level knowledge from rule bases which is important to allow queries at and across different levels of abstraction. FCA is used to develop an abstraction hierarchy and the approach is applied to knowledge bases from the domain of chemical pathology. In Zarate et al. (2009), FCA is used to extract and represent knowledge in the form of a non-redundant canonical rule base with minimal implications from a trained ANN.

4.1.2 Software Mining

Software mining covers 19% of the 140 KDD papers and describes how FCA can be used to gain insight in amongst others software source code. In Cole et al. (2005), FCA is used to conceptually analyse relational structures in software source code and to detect unnecessary dependencies between software parts. In Cellier et al. (2008), FCA is used in combination with association rules for fault localization in software source code. Wermelinger (2009) uses FCA lattices to visualize the relations between the software artefacts and the developers who should fix the bugs in them. In Eisenbarth (2003), a technique is presented for reconstructing the mapping for features that are triggered by the user to the source code of the system. Mens et al (2005) use FCA to delve a system's source code for relevant concepts of interest: what concerns are addressed in the code, what patterns, coding idioms and conventions have been adopted and where and how are they implemented.

Crosscutting concerns, i.e. functionalities that are not assigned to a single modular unit in the implementation are one of the major problems in software evolution. Aspect Oriented Programming offers mechanisms to factor them out into a modular unit, called an aspect. In Tonella et al. (2004), aspect identification in existing code is supported by means of dynamic code analysis. Execution traces are generated for the use cases that exercise the main functionalities of a given application. The relationship between execution traces and executed computational units is subjected to concept analysis. In the resulting lattice, potential aspects are detected. Yang et al (2008), discuss an aspect-mining approach in which execution profiles of legacy systems are analyzed using concept lattices to identify the invoked computational units that traverse system's use case models. They can be abstracted into early-aspects for re-engineering of the legacy system with AOSD. Qu (2007) also discusses the use of FCA for aspect mining to identify crosscutting concerns in a system thereby improving the system's comprehensibility and enabling migration of existing (object-oriented) programs to aspect-oriented ones. Breu et al (2006) mined aspects from Eclipse by analyzing where developers added code to the program over time. In Del

Grosso et al. (2007), an approach is proposed to identify from database-oriented applications, pieces of functionality to be potentially exported as services.

Role Based Access Control (RBAC) is a methodology for providing users in an IT system with specific permissions like read or write. Molloy et al. (2008) use FCA for mining roles from user-permission and user-attribute information to complement the costly top-down approaches for RBAC. Dau et al. (2009) apply FCA in combination with Description Logics to capture the RBAC constraints and for deriving additional constraints.

4.1.3 Web Mining

Web mining and improving the quality of web search results is investigated in 8% of the KDD papers. Periodic web personalization aims to recommend the most relevant resources to a user during a specific time period by analyzing the periodic access patterns of the user from web usage logs. Beydoun et al. (2007) introduce a system which captures user trails as they search the internet. They construct a semantic web structure from the trails and this semantic web structure is expressed as a conceptual lattice guiding future searches. Beydoun (2009) further investigates the possibilities of FCA for processing students virtual surfing trails to express and exploit the dependencies between visited web-pages to yield subsequent and more effective focused search results. He (2007) also proposes a method for automatically mining and acquiring web user profiles using FCA. Okubo et al. (2006) show how FCA can be used for the conceptual clustering of web documents and to provide a conceptual meaning for each document cluster. Myat et al. (2005) use FCA for conceptual document clustering to manage the information published on the World Wide Web. Wang et al. (2008) give a method for using FCA for developing a topic-specific web crawler for use in web data mining.

Du et al. (2009) present a method based on FCA for mining association rules that can be used to match user queries with web pages to avoid returning irrelevant web pages for search engine results. Hsieh et al. (2007) propose a knowledge acquisition system which dynamically constructs the relationships and hierarchy of concepts in a query-based ontology to provide answers for user's queries. Kim et al. (2007) discuss a novel approach using FCA to build a contextualized folksonomy and concept hierarchies from tags of blogosphere.

4.1.4 Extending FCA for Data Mining

In the last years, multiple extensions have been introduced into the literature that improve traditional FCA theory's applicability to knowledge discovery problems. Belohlavek et al. (2009) emphasizes the need for taking into account background knowledge in FCA. They present an approach for modeling background knowledge that represents user's priorities regarding attributes and their relative importance. Only those concepts that are compatible with user's priorities are considered as relevant and extracted from the data. In Pogel et al. (2008), FCA is used in combination with a tag context for formally incorporating important kinds of background knowledge. The results are Generalized Contingency Structures and Tagged Contingency structures which can be used for data summarization in epidemiology. In Poelmans et al. (2009), FCA is used in combination with Emergent Self Organising Maps for detecting domestic violence in the unstructured text of police reports.

In Besson et al. (2006), FCA is extended to cope with faults and to improve formal concepts towards fault tolerance. Pfaltz (2007) extends FCA to deal with numerical values. Valverde-Albacete et al. (2006) introduced a generalization of FCA for data mining applications called *K-Formal Concept Analysis*. This idea was further developed in Valverde-Albacete et al. (2007) where the lattice structure for such generalized contexts was introduced. This research topic was further investigated in Valverde-Albacete et al (2008). Hashemi (2004) proposes a method for efficiently creating a new lattice from an already existing one when the data granularity is changed. Lei (2007) introduces the notion of extended many-valued context to avoid the generation of a large one-valued context in document knowledge discovery. In Deogun et al (2003) FCA is complemented with Bacchus probability logic, which makes use of statistical and propositional probability inference. The authors introduce a new type of concept called "previously unknown and potentially useful" and formalize KDD as a process to find such concepts.

In its classical form FCA considers attributes as a non-ordered set. When attributes of the context are partially ordered to form a taxonomy, conceptual scaling allows the taxonomy to be taken into account by producing a context completed with all attributes deduced from the taxonomy. In Cellier et al (2008) an algorithm is proposed to learn concept-based rules in the presence of a taxonomy.

Another FCA research topic is attribute reduction. Shao et al. (2008) show how to remove redundant attributes from real set formal contexts without any loss of knowledge. Wu et al. (2009) discuss the application of viewing data at different levels of granularity to construct a granular data structure which can be used for knowledge reduction in FCA. Wang et al (2008) deal with approaches to generalized attribute reduction in a consistent decision formal context. Ganter et al (2008) describe how scaled many-valued contexts of FCA may make feature selection easier.

4.1.5 FCA Mining Applications in Biology and Medicine

10% of the KDD papers describe applications of FCA in biology, chemistry or medicine. In Sato et al (2007), FCA is used to cluster time-series medical data and to analyze these clusters. Sklenar (2005) used FCA to evaluate epidemiological questionnaire physical activity data to find dependencies between demographic data and degree of physical activity. In Kaytone et al. (2009), FCA is used for mining and clustering gene expression data. Fu (2006) applies FCA as a tool for analysis and visualization of data in a digital ecosystem. Maddouri (2004) outlines a new incremental learning approach based on FCA that supports incremental concept formation and applies it to the problem of cancer diagnosis. The incremental approach has the advantage of handling both the problem of data addition, data deletion, data update, etc.

4.1.6 Fuzzy FCA in KDD

In Fuzzy FCA, each table entry contains a truth degree to which an attribute applies to an object. 9% of the papers use Fuzzy FCA for KDD. In Chou et al. (2008), Fuzzy FCA is used for tag mining, i.e. to analyze the relationships between semantic tags and Web APIs. In Fenza et al. (2008), Fuzzy FCA is used for the discovery of semantic web services. Zhou et al. (2006) use Fuzzy FCA to construct a user behaviour model from web usage logs to identify the resources that the user is most

likely interested in during a given period. Fenza et al (2009) present a system which uses fuzzy FCA for supporting the user in the discovery of semantic web services. Through a concept-based navigation mechanism, the user discovers conceptual terminology associated to the web resources and uses it to generate an appropriate service request. Yan (2007) uses Fuzzy Set Theory to extend the many-valued context from FCA. This fuzzy many-valued context can then be used for document knowledge discovery.

5 Conclusion

We found FCA to be an interesting instrument to explore the literature on concept analysis. Over 700 papers have been published over the past 7 years on FCA and 140 zoomed in on applying FCA in KDD. The main research topics in this area are association rule mining, software mining, web mining, KDD in medicine and biology, using Fuzzy FCA for KDD and to extend the capabilities of traditional FCA theory for data mining. In the future, we will host the references and links to the article on a public interface. Further research consists of doing this same analysis for information retrieval, scalability and ontology construction.

Acknowledgements

Jonas Poelmans is aspirant of the "Research Foundation - Flanders" or "Fonds voor Wetenschappelijk Onderzoek - Vlaanderen".

References

1. Belohlavek, R., Vychodil, V.: Formal Concept Analysis With Background Knowledge: Attribute Priorities. *IEEE Trans. Syst., man, & cyb. - C: App. & rev.* 39(4) (2009)
2. Besson, J., Robardet, C., Boulicaut, J.F.: Mining a New Fault-Tolerant Pattern Type as an Alternative to Formal Concept Discovery. In: Schärfe, H., Hitzler, P., Øhrstrøm, P. (eds.) *ICCS 2006. LNCS (LNAI)*, vol. 4068, pp. 144–157. Springer, Heidelberg (2006)
3. Beydoun, G.: Using Formal Concept Analysis towards Cooperative E-Learning. In: Richards, D., Kang, B.-H. (eds.) *PKAW 2008. LNCS (LNAI)*, vol. 5465, pp. 109–117. Springer, Heidelberg (2009)
4. Beydoun, G., Kultchitsky, R., Manasseh, G.: Evolving semantic web with social navigation. *Expert Systems with Applications* 32, 265–276 (2007)
5. Breu, S., Zimmermann, T., Lindig, C.: Mining Eclipse for Cross-Cutting Concerns. In: *MSR 2006, Shanghai, China, May 22-23 (2006)*
6. Carpineto, C., Romano, G.: *Concept data analysis: Theory and applications*. John Wiley & Sons, Chichester (2004)
7. Cellier, P., Ferré, S., Ridoux, O., Ducasse, M.: A parameterized algorithm to explore formal contexts with a taxonomy. *Int. J. Found. of Comp. Sc.* 19(2), 319–343 (2008)
8. Chang, Y.H.: Automatically constructing a domain ontology for document classification. In: *6th Int. Conf. on Machine Learning and Cybernetics, Hong Kong (2007)*
9. Chou, C.Y., Mei, H.: Analyzing Tag-based Mashups with Fuzzy FCA. In: *IEEE Int. Symposium on Service-Oriented System Engineering (2008)*

10. Cole, R., Becker, P.: Navigation Spaces for the Conceptual Analysis of Software Structure. In: Ganter, B., Godin, R. (eds.) ICFCA 2005. LNCS (LNAI), vol. 3403, pp. 113–128. Springer, Heidelberg (2005)
11. Cole, R., Becker, P.: Navigation Spaces for the Conceptual Analysis of Software Structure. In: Ganter, B., Godin, R. (eds.) ICFCA 2005. LNCS (LNAI), vol. 3403, pp. 113–128. Springer, Heidelberg (2005)
12. Cole, R., Eklund, P., Stumme, G.: Document retrieval for e-mail search and discovery using Formal Concept Analysis. *App. Art. Intel.* 17, 257–280 (2003)
13. Cole, R., Tilley, T., Ducrou, J.: Conceptual Exploration of Software Structure: A Collection of Examples. In: Belohlavek, R., Snasel, V. (eds.) CLA, pp. 135–148 (2005)
14. Correira, J.H., Stumme, G., Wille, R., Wille, U.: Conceptual knowledge discovery - a human-centered approach. *Applied Artificial Intelligence* 17, 281–302 (2003)
15. Dau, F., Knechte, M.: Access Policy Design Supported by FCA Methods. In: Rudolph, S., Dau, F., Kuznetsov, S.O. (eds.) ICCS 2009. LNCS (LNAI), vol. 5662, pp. 141–154. Springer, Heidelberg (2009)
16. Del Grosso, C., Penta, M.D., Guzman, I.G.R.: An approach for mining services in database-oriented applications. In: 11th IEEE Eur. Conf. on Software Maintenance and Reeng. (2007)
17. Deogun, J., Jiang, L., Xie, Y., Raghavan, V.: Probability Logic Modeling of Knowledge Discovery in Databases. In: Zhong, N., Raś, Z.W., Tsumoto, S., Suzuki, E. (eds.) ISMIS 2003. LNCS (LNAI), vol. 2871, pp. 402–407. Springer, Heidelberg (2003)
18. Dong, G., Jiang, C., Pei, J., Li, J., Wong, L.: Mining Succinct Systems of Minimal Generators of Formal Concepts. In: Zhou, L.-z., Ooi, B.-C., Meng, X. (eds.) DASFAA 2005. LNCS, vol. 3453, pp. 175–187. Springer, Heidelberg (2005)
19. Du, Y.J., Li, H.M.: Strategy for Mining Association Rules for Web Pages Based on Formal Concept Analysis. Elsevier, Amsterdam (2009), doi:10.1016/j.asoc.2009.09.007
20. Ducrou, J.: DVDSleuth: A Case Study in Applied Formal Concept Analysis for Navigating Web Catalogs. In: Priss, U., Polovina, S., Hill, R. (eds.) ICCS 2007. LNCS (LNAI), vol. 4604, pp. 496–500. Springer, Heidelberg (2007)
21. Eidenberger, H.: Visual Data Mining. In: Proceedings SPIE Optics East Conf., Philadelphia, October 26–28, vol. 5601, pp. 121–132 (2004)
22. Eisenbarth, T., Koschke, R., Simon, D.: Locating Features in Source Code. *IEEE Transactions on Software Engineering* 29(3) (March 2003)
23. Eklund, P., Ducrou, J.: Navigation and Annotation with Formal Concept Analysis. In: Richards, D., Kang, B.-H. (eds.) PKAW 2008. LNCS (LNAI), vol. 5465, pp. 118–121. Springer, Heidelberg (2009)
24. Eklund, P., Wille, R.: Semantology as Basis for Conceptual Knowledge Processing. In: Kuznetsov, S.O., Schmidt, S. (eds.) ICFCA 2007. LNCS (LNAI), vol. 4390, pp. 18–38. Springer, Heidelberg (2007)
25. Fenza, G., Loia, V., Senatore, S.: Concept Mining of Semantic Web Services By Means of Extended Fuzzy Formal Concept Analysis (FFCA). In: *Int. Conf. Syst., Man & Cyb.* (2008)
26. Fenza, G., Senatore, S.: Friendly web services selection exploiting fuzzy formal concept analysis. In: *Soft Computing*. Springer, Heidelberg (2009)
27. Fu, H.: Formal Concept Analysis for Digital Ecosystem. In: *Proc. of the 5th Int. Conf. or Machine Learning and Applications (ICMLA 2006)*. IEEE, Los Alamitos (2006)
28. Ganter, B., Kuznetsov, S.O.: Formalizing Hypotheses with Concepts. In: Ganter, B., Mineau, G.W. (eds.) ICCS 2000. LNCS (LNAI), vol. 1867, pp. 342–356. Springer, Heidelberg (2000)

29. Ganter, B., Kuznetsov, S.O.: Scale Coarsening as Feature Selection. In: Medina, R., Obiedkov, S. (eds.) ICFCA 2008. LNCS (LNAI), vol. 4933, pp. 217–228. Springer, Heidelberg (2008)
30. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical foundations. Springer, Heidelberg (1999)
31. Gupta, A., Kumar, N., Bhatnagar, V.: Incremental Classification Rules Based on Association Rules Using Formal Concept Analysis. In: Perner, P., Imiya, A. (eds.) MLDM 2005. LNCS (LNAI), vol. 3587, pp. 11–20. Springer, Heidelberg (2005)
32. Hamrouni, T., Yahia, S.B., Nguifo, E.M.: Towards a Finer Assessment of Extraction Contexts Sparseness. In: 18th Int. Workshop on Database and Expert Systems Applications (2007)
33. Hamrouni, T., Yahia, S.B., Slimani, Y.: Avoiding the itemset closure computation pitfall. In: Belohlavek, R., Snasel, V. (eds.) CLA 2005, pp. 46–59 (2005)
34. Hamrouni, T., Yahia, S.B., Slimani, Y.: Prince: An Algorithm for Generating Rule Bases Without Closure Computations. In: Tjoa, A.M., Trujillo, J. (eds.) DaWaK 2005. LNCS, vol. 3589, pp. 346–355. Springer, Heidelberg (2005a)
35. Hashemi, R.R., De Agostino, S., Westgeest, B., Talburt, J.R.: Data Granulation and Formal Concept Analysis. In: Processing NAFIPS 2004, IEEE Annual Meeting of the Fuzzy Information, vol. 1, pp. 79–83 (2004)
36. He, H., Hai, H., Rujing, W.: FCA –Based Web User Profile Mining for Topics of Interest. In: Int. Conf. on Integation Technology, Shenzhen, China, March 20-24 (2007)
37. Hermann, M., Sertkaya, B.: On the Complexity of Computing Generators of Closed Sets. In: Medina, R., Obiedkov, S. (eds.) ICFCA 2008. LNCS (LNAI), vol. 4933, pp. 158–168. Springer, Heidelberg (2008)
38. Hsieh, T.C., Tsai, K.H., Chen, C.L., Lee, M.C., Chiu, T.K., Wang, T.: Query-based ontology approach for semantic search. In: 6th Int. Conf. on Machine Learning & Cyb. (2007)
39. Kaytone, M., Duplessis, S., Kuznetsov, S.O., Napoli, A.: Two FCA-Based Methods for Mining Gene Expression Data. In: Ferre, S., et al. (eds.) ICFCA. LNCS (LNAI), vol. 5548, pp. 251–266. Springer, Heidelberg (2009)
40. Keim, D.A.: Information visualization and visual data mining. IEEE Transactions on Visualization and Computer Graphics 8(I) (2002)
41. Kim, H.L., Hwang, S.H., Kim, H.G.: FCA-based Approach for Mining Contextualized Folksonomy. In: SAC 2007, Seoul, Korea, March 11-15 (2007)
42. Lakhil, L., Stumme, G.: Efficient Mining of Association Rules Based on Formal Concept Analysis. In: Ganter, B., Stumme, G., Wille, R. (eds.) Formal Concept Analysis. LNCS (LNAI), vol. 3626, pp. 180–195. Springer, Heidelberg (2005)
43. Lei, Y., Cao, B., Yu, J.: A formal description-based approach to extended many-valued context analysis. In: 4th Conf. Fuzzy Systems and Knowledge Discovery, vol. 1, pp. 545–549 (2007)
44. Lim, W.C., Lee, C.S.: Knowledge Discovery Through Composited Visualization, Navigation and Retrieval. In: Hoffmann, A., Motoda, H., Scheffer, T. (eds.) DS 2005. LNCS (LNAI), vol. 3735, pp. 377–379. Springer, Heidelberg (2005)
45. Maddouri, M.: Towards a machine learning approach based on incremental concept formation. Intelligent Data Analysis 8, 267–280 (2004)
46. Maddouri, M.: A Formal Concept Analysis Approach to Discover Association Rules from Data. In: Belohlavek, R., Snasel, V. (eds.) CLA, pp. 10-21 (2005)

47. Maddouri, M., Kaabi, F.: On Statistical Measures for Selecting Pertinent Formal Concepts to Discover Production Rules from Data. In: 6th Int. Conf. Data Mining Workshops (2006)
48. Meddouri, N., Maddouri, M.: Boosting Formal Concepts to Discover Classification Rules. In: Chien, B.-C., Hong, T.-P., Chen, S.-M., Ali, M. (eds.) IEA/AIE 2009. LNCS (LNAI), vol. 5579, pp. 501–510. Springer, Heidelberg (2009)
49. Mens, K., Tourwe, T.: Delving source code with formal concept analysis. *Computer Languages. Systems & Structures* 3, 183–197 (2005)
50. Molloy, I., Chen, H., Li, T., Wang, Q., Li, N., Bertino, E., Calo, S., Lobo, J.: Mining Roles with Semantic Meanings, Estes Park, Colorado, USA, June 11-13 (2008)
51. Myat, N.N., Hla, K.H.S.: A combined approach of formal concept analysis and text mining for concept based document clustering. In: Int. Conf. on Web Intelligence (2005)
52. Nehme, K., Valtchev, P., Rouane, M.H., Godin, R.: On Computing the Minimal Generator Family for Concept Lattices and Icebergs. In: Ganter, B., Godin, R. (eds.) ICFCFA 2005. LNCS (LNAI), vol. 3403, pp. 192–207. Springer, Heidelberg (2005)
53. Nourine, L., Raynaud, O.: A fast algorithm for building lattices. In: *Information Processing Letters*, pp. 199–214 (1999)
54. Okubo, Y., Haraguchi, M.: Finding Conceptual Document Clusters with Improved Top-N Formal Concept Search. In: Int. Conf. on Web Intelligence (2006)
55. Pfaltz, J.C.: Representing numeric values in concept lattices. In: CLA (2007)
56. Poelmans, J., Elzinga, P., Viaene, S., Dedene, G.: A case of using formal concept analysis in combination with emergent self organizing maps for detecting domestic violence. In: Perner, P. (ed.) *Advances in Data Mining. Applications and Theoretical Aspects*. LNCS, vol. 5633, pp. 247–260. Springer, Heidelberg (2009)
57. Pogel, A., Ozonoff, D.: Contingency Structures and Concept Analysis. In: Medina, R., Obiedkov, S. (eds.) ICFCFA 2008. LNCS (LNAI), vol. 4933, pp. 305–320. Springer, Heidelberg (2008)
58. Priss, U.: Formal Concept Analysis in Information Science. In: Blaise, C. (ed.) *Annual Review of Information Science and Technology, ASIST*, vol. 40 (2006)
59. Priss, U., Old, L.J.: Conceptual Exploration of Semantic Mirrors. In: Ganter, B., Godin, R. (eds.) ICFCFA 2005. LNCS (LNAI), vol. 3403, pp. 21–32. Springer, Heidelberg (2005)
60. Qi, H., Liu, D.Y., Hu, C.Q., Lu, M., Zhao, L.: Searching for closed itemset with Formal Concept Analysis. In: 3rd Int. Conf. on Machine Learning and Cybernetics, Shanghai (2004)
61. Qu, L., Liu, D.: Extending Dynamic Aspect Mining Using Formal Concept Analysis. In: Proc. of 4th Conf. on Fuzzy Systems and Knowledge Discovery (2007)
62. Quan, T.T., Ngo, L.N., Hui, S.C.: An Effective Clustering-based Approach for Conceptual Association Rules Mining. In: *Conf. Comp. & Comm. Techn.*, pp. 1–7 (2009)
63. Richards, D., Malik, U.: Multi level knowledge discovery from rule bases. *Applied Artificial Intelligence* 17, 181–205 (2003a)
64. Richards, D.: Addressing the Ontology Acquisition Bottleneck through Reverse Ontological Engineering. *Knowledge and Information Systems* 6, 402–427 (2004)
65. Richards, D.: Ad-Hoc and Personal Ontologies: A Prototyping Approach to Ontology Engineering. In: Hoffmann, A., Kang, B.-h., Richards, D., Tsumoto, S. (eds.) PKAW 2006. LNCS (LNAI), vol. 4303, pp. 13–24. Springer, Heidelberg (2006)
66. Richards, D., Malik, U.: Mining Propositional Knowledge Bases to Discover Multi-level Rules. In: Zaïane, O.R., Simoff, S.J., Djeraba, C. (eds.) MDM/KDD 2002 and KDMCD 2002. LNCS (LNAI), vol. 2797, pp. 199–216. Springer, Heidelberg (2003)

67. Rudolph, S.: Exploring Relational Structures Via FLE. In: Wolff, K.E., Pfeiffer, H.D., Delugach, H.S. (eds.) ICCS 2004. LNCS (LNAI), vol. 3127, pp. 196–212. Springer, Heidelberg (2004)
68. Rudolph, S.: Acquiring Generalized Domain-Range Restrictions. In: Medina, R., Obiedkov, S. (eds.) ICFCA 2008. LNCS (LNAI), vol. 4933, pp. 32–45. Springer, Heidelberg (2008)
69. Rudolph, S., Völker, J., Hitzler, P.: Supporting Lexical Ontology Learning by Relational Exploration. In: Priss, U., Polovina, S., Hill, R. (eds.) ICCS 2007. LNCS (LNAI), vol. 4604, pp. 488–491. Springer, Heidelberg (2007)
70. Sato, K., Okubo, Y., Haraguchi, M., Kunifuji, S.: Data Mining of Time-Series Medical Data by Formal Concept Analysis. In: Apolloni, B., Howlett, R.J., Jain, L. (eds.) KES 2007, Part II. LNCS (LNAI), vol. 4693, pp. 1214–1221. Springer, Heidelberg (2007)
71. Shao, M.W., Guo, Y.L.: Attribute reduction of large crisp-real concept lattices. In: Proc. of the 7th Int. Conf. on Machine Learning and Cybernetics, Kunming (2008)
72. Shi, B.S., Shen, X.J., Liu, Z.T.: A Knowledge discovery technique for heterogeneous data sources. In: Proc. of IEEE 2nd Int. Conf. on Machine Learning and Cybernetics, Xian (2003)
73. Sklenar, V., Zacpal, J., Sigmund, E.: Evaluation of IPAQ questionnaire by FCA. In: Belohlavek, R., Snasel, V. (eds.) CLA, pp. 60–69 (2005)
74. Stumme, G., Bestride, Y., Taouil, R., Lakhal, L.: Computing Iceberg Concept Lattices with TITANIC. *Data and Knowledge Engineering* 42(2), 189–222 (2002)
75. Stumme, G., Wille, R., Wille, U.: Conceptual knowledge discovery in databases using Formal Concept Analysis Methods. In: Żytkow, J.M. (ed.) PKDD 1998. LNCS, vol. 1510, pp. 450–458. Springer, Heidelberg (1998)
76. Tekaya, S.B., Yahia, S.B., Slimani, Y.: GenAll Algorithm: Decorating Galois lattice with minimal generators. In: Belohlavek, R., Snasel, V. (eds.) CLA, pp. 166–178 (2005)
77. Thomas, J., Cook, K.: Illuminating the path: research and development agenda for visual analytics. National Visualization and Analytics Ctr. (2005)
78. Tilley, T.: Tool support for FCA. In: Eklund, P. (ed.) ICFCA 2004. LNCS (LNAI), vol. 2961, pp. 104–111. Springer, Heidelberg (2004)
79. Tilley, T., Eklund, P.: Citation analysis using Formal Concept Analysis: A case study in software engineering. In: 18th Int. Conf. on Database and Expert Systems Applications (2007)
80. Tonella, P., Ceccato, M.: Aspect Mining through the Formal Concept Analysis of Execution Traces. In: Proc. of the 11th Working Conf. on Reverse Engineering (2004)
81. Valtchev, P., Missaoui, R., Godin, R.: Formal Concept Analysis for Knowledge Discovery and Data Mining: The New Challenges. In: Eklund, P. (ed.) ICFCA 2004. LNCS (LNAI), vol. 2961, pp. 352–371. Springer, Heidelberg (2004)
82. Valtchev, P., Missaoui, R., Godin, R.: A framework for incremental generation of closed itemsets. *Discrete Applied Mathematics* 156, 924–949 (2008)
83. Valverde-Albacete, F.J., Pelaez-Moreno, C.: Galois Connections Between Semimodules and Applications in Data Mining. In: Kuznetsov, S.O., Schmidt, S. (eds.) ICFCA 2007. LNCS (LNAI), vol. 4390, pp. 181–196. Springer, Heidelberg (2007)
84. Valverde-Albacete, F.J., Pelaez-Moreno, C.: Spectral Lattices of $R_{\max,+}$ -Formal Contexts. In: Medina, R., Obiedkov, S. (eds.) ICFCA 2008. LNCS (LNAI), vol. 4933, pp. 124–139. Springer, Heidelberg (2008)
85. Valverde-Albacete, F.J., Pelaez-Moreno, C.: Towards a generalization of Formal concept analysis for data mining purposes. In: Missaoui, R., Schmidt, J. (eds.) ICFCA 2006. LNCS (LNAI), vol. 3874, pp. 161–176. Springer, Heidelberg (2006)

86. Volker, J., Rudolph, S.: Fostering Web Intelligence by Semi-automatic OWL Ontology Refinement. In: *Int. Conf. on Web Intelligence and Intelligent Agent Technology* (2008)
87. Volker, J., Rudolph, S.: Lexico-Logical Acquisition of OWL DL Axioms: An Integrated Approach to Ontology Refinement. In: Medina, R., Obiedkov, S. (eds.) *ICFCA 2008. LNCS (LNAI)*, vol. 4933, pp. 62–77. Springer, Heidelberg (2008)
88. Wang, H., Zhang, W.X.: Generalized attribute reduction in consistent decision formal context. In: *7th Int. Conf. on Machine Learning and Cybernetics*, Kunming, July 12-15 (2008)
89. Wermelinger, M., Yu, Y., Strohmaier, M.: Using Formal Concept Analysis to Construct and Visualize Hierarchies of Socio-Technical Relations. In: *ICSE 2009*, Vancouver (2009)
90. Wille, R.: Methods of conceptual knowledge processing. In: Missaoui, R., Schmidt, J. (eds.) *ICFCA 2006. LNCS (LNAI)*, vol. 3874, pp. 1–29. Springer, Heidelberg (2006)
91. Wille, R.: Restructuring lattice theory: an approach based on hierarchies of concepts. In: Rival, I. (ed.) *Ordered sets*, pp. 445–470. Reidel, Dordrecht (1982)
92. Wollbold, J., Guthke, R., Ganter, B.: Constructing a Knowledge Base for Gene Regulatory Dynamics by Formal Concept Analysis Methods. In: Horimoto, K., Regensburger, G., Rosenkranz, M., Yoshida, H. (eds.) *AB 2008. LNCS*, vol. 5147, pp. 230–244. Springer, Heidelberg (2008)
93. Wu, W.Z., Leung, Y., Mi, J.S.: Granular Computing and Knowledge Reduction in Formal Contexts. *IEEE Transactions on Knowledge & Data Engineering* 21(10) (October 2009)
94. Yahia, S.B., Xguifo, E.M.: Revisiting Generic Bases of Association Rules. In: Kambayashi, Y., Mohania, M., Wöß, W. (eds.) *DaWaK 2004. LNCS*, vol. 3181, pp. 58–67. Springer, Heidelberg (2004)
95. Yan, W., Baoxiang, C.: Fuzzy Many-Valued Context Analysis Based on Formal Description. In: *8th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing* (2007)
96. Yang, S.Z., Hou, X.W., Zhang, M.Q.: Approach on Aspect-Oriented Software Reverse Engineering at Requirements Level. In: *Int. Conf. on Computer Science and Software Engineering*. IEEE, Los Alamitos (2008)
97. Yang, Y., Du, Y., Sun, J., Hai, Y.: A Topic-Specific Web Crawler with Concept Similarity Context Graph Based on FCA. In: Huang, D.-S., Wunsch II, D.C., Levine, D.S., Jo, K.-H. (eds.) *ICIC 2008. LNCS (LNAI)*, vol. 5227, pp. 840–847. Springer, Heidelberg (2008)
98. Zarate, L.E., Dias, S.M.: Qualitative behavior rules for the cold rolling process extracted from trained ANN via the FCANN method. *Engineering Applications of Artificial Intelligence* 22, 718–731 (2009)
99. Zhou, B., Hui, S.C., Chang, K.: A formal concept analysis approach for web usage mining. *Intelligent Information Processing II* 163, 437–441 (2005)
100. Zhou, B., Hui, S.C., Fong, A.C.M.: An Effective Approach for Periodic Web Personalization. In: *Proc. of the IEEE/WIC/ACM Int. Conf. on Web Intelligence* (2006)