How KR benefits from Formal Concept Analysis
Proposal for a half-day tutorial at KR 2023

Francesco Kriegel[1] and Barış Sertkaya[2]

[1] Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany
`francesco.kriegel@tu-dresden.de`
[2] Frankfurt University of Applied Sciences
`sertkaya@fb2.fra-uas.de`

Tutorial description

Formal Concept Analysis (FCA) [GW99] emerged from an attempt to restructuring lattice theory and making it more accessible to potential users [Wil82]. Data is represented as a formal context, which is a table showing which objects have which attributes. On the one hand, FCA can be used to identify clusters in the data, which are called concepts, and to build a hierarchy of these concepts. On the other hand, FCA helps to extract the dependencies between sets of attributes, which correspond to propositional Horn formulae. The set of all implications that hold in a formal context can be large. The *canonical base* [GD86] represents this implicative theory in a complete way and consists of a minimal number of implications.

Both the concept hierarchy and the implication base can be computed in exponential time [Gan84]. In cases where only an expert with complete knowledge is available (but no formal context), the interactive algorithm *Attribute Exploration* [Gan84] can be used to construct a minimal representation of the implicative theory. More recent algorithms are much faster and have thus enabled the usage of FCA in applications, like Fast Close-by-One [KOV10], In-Close [And18], and LinCbO [JKK21].

The idea of employing methods and tools from FCA in Description Logic (DL) has a long tradition. FCA has been used in DL for an efficient computation of an extended subsumption hierarchy of DL concepts as well as for bottom-up construction of ontologies [BM00; BS04; Ser07], for acquiring complete knowledge about an application domain as well as for enriching OWL ontologies [Baa+07; Ser07], and for mining axioms from interpretations and from knowledge graphs [BD08; Dis11; Bor14; Kri19a]. In this tutorial we will first briefly introduce basics of FCA, and later concentrate on its applications in DL.

Target audience and learning goals

The potential target audience of the tutorial are researchers working in the fields of description logic, ontologies and semantic web, as well as ontology engineers from different application domains where ontologies are employed, like life sciences and bio-medicine.

For the audience a background in logic is sufficient. Expected learning goals include:

- compact representation of implications holding in a data set,
- compact representation of axioms holding in an interpretation or in a knowledge graph,
- efficient enumeration of concepts in the data set,
- enriching OWL ontologies with missing knowledge with the help of a domain expert by using an FCA-based interactive knowledge acquisition method.

Tutorial outline

Session 1: Introduction to FCA (90 minutes)

1. Conceptual clustering with FCA *(Francesco)*

   We begin with introducing the basic notions in FCA, e.g., *formal contexts* used to represent the data and *formal concepts* that describe clusters in the data. The concepts can be partially ordered by a "is subconcept of" relation, which yields a complete lattice, i.e., an arbitrary number of formal concepts always has a most general subconcept as well as a most specific superconcept. On the one hand, this lattice reveals the conceptual hierarchy of the data, which can then be used in an application. On the other hand, it allows to visualize the data by means of a line diagram: each formal concept is displayed as a node and it is connected by an upwards directed edge to all its neighboring superconcepts. We also present algorithms for computing the concept lattice (NextClosure [Gan84], Fast Close-by-One [KOV10], InClose-5 [And18]), and demonstrate the software tool *Concept Explorer FX*.[3]

2. Extracting dependencies with FCA *(Barış)*

   Dependencies in the data are described by *implications* between attribute sets, which correspond to propositional Horn formulae. The implicative theory of a formal context is completely described by means of the *canonical base* [GD86], in the sense that every valid implication follows from it. Moreover, it consists of a minimal number of implications among all implication sets that are complete. We present complexity results related to computing the canonical base, and show algorithms for computing it (NextClosure [Gan84], Fast Close-by-One [KOV10], LinCbO [JKK21]). Furthermore, we present the the interactive algorithm *Attribute Exploration* [Gan84] and its variations, with which a minimal representation of the implicative theory can be constructed in cases where only an expert with complete knowledge is available (but no formal context).

[3] https://github.com/francesco-kriegel/conexp-fx

Session 2: Applications in DL (90 minutes)

3. Acquiring complete knowledge about an application domain, enriching OWL ontologies *(Barış)*

   We present an approach for extending both the terminological and the assertional part of a DL ontology by using information provided by the ontology and by a domain expert. Our approach provides a basis for formally well-founded techniques and tools that support the knowledge engineer in checking whether a ontology contains all the relevant information about the application domain, and in extending the ontology appropriately if this is not the case. The use of techniques from Formal Concept Analysis ensures that, on the one hand, the interaction with the expert is kept to a minimum, and, on the other hand, it enables us to show that the extended ontology is complete in a certain, well-defined sense.

4. Mining axioms from interpretations and knowledge graphs *(Francesco)*

   A DL ontology is usually divided into a TBox and an ABox. The class hierarchy of the underlying domain of interest is described in the TBox: it consists of implications $C \sqsubseteq D$ between classes with the intended meaning that each object in the class C is also in the class D. In the ABox the objects are assigned to classes and linked to each other by roles. Building an ABox is as easy as collecting data. In contrast, constructing a TBox is a complex task. We present an automatic approach that axiomatizes a TBox from given data that comes in form of an interpretation or a knowledge graph (a graph in which the nodes and edges are labeled). By means of a reduction to FCA, it computes a minimal set of TBox axioms that is complete for the input data. In addition, we show variations and extensions of this approach.

5. Enumerating stable extensions in an argumentation framework using closed sets *(Barış)*

   We present an approach based on Formal Concept Analysis (FCA) for computing stable extensions of Abstract Argumentation Frameworks (AFs). To this purpose, we represent an AF as a formal context in which stable extensions of the AF are closed sets called concept intents. We make use of algorithms developed in FCA for computing concept intents in order to compute stable extensions of AFs. Experimental results show that, on AFs with a high density of the attack relation, our algorithms perform significantly better than the existing approaches. The algorithms can be modified to compute other types of extensions, in particular, preferred extensions.

Resumes of both presenters

Francesco Kriegel is a research associate in the Institute of Theoretical Computer Science at Technische Universität Dresden. He joined the research group of Franz Baader after completing his studies in Mathematics as major (with a focus on Algebra and Analysis) and Computer Science as minor. He obtained his doctoral degree in 2019 with a thesis concerned with constructing and extending Description Logic ontologies using methods of Formal Concept Analysis. From his third semester on and also during his doctoral studies he was engaged in teaching, mostly tutorials related to basic and later also advanced courses. He also gave several talks at international conferences and workshops. Currently, he is employed in a research project funded by the German Research Foundation (DFG), which is concerned with repairing Description Logic ontologies.

Barış Sertkaya is a professor of computer science at Department of Computer Science and Engineering at Frankfurt University of Applied Sciences, Germany. His research interests are centered around knowledge representation and reasoning, Description Logics, Formal Concept Analysis and applications of these fields in ontologies and the Semantic Web. In his PhD thesis he has investigated how Description Logic ontologies can benefit from Formal Concept Analysis and has developed a method for supporting ontology engineers in detecting missing knowledge in the ontology and completing it in an efficient way. His teaching experience comprises regular lectures as well as a tutorials and invited talks. He has been publishing at various AI conferences like AAAI, IJCAI, ECAI, JELIA and KR. A list of relevant publications is provided below. Barış was program chair of the International Conference on Formal Concept Analysis (ICFCA) in 2010 and the general chair in 2019.

Related publications of Francesco Kriegel

[BDK16] Daniel Borchmann, Felix Distel, and Francesco Kriegel. "Axiomatisation of General Concept Inclusions from Finite Interpretations". In: *Journal of Applied Non-Classical Logics* 26.1 (2016), pp. 1–46. DOI: 10.1080/11663081.2016.1168230.

[KB17] Francesco Kriegel and Daniel Borchmann. "NextClosures: Parallel Computation of the Canonical Base with Background Knowledge". In: *International Journal of General Systems* 46.5 (2017), pp. 490–510. DOI: 10.1080/03081079.2017.1349570.

[Kri16a] Francesco Kriegel. "Axiomatization of General Concept Inclusions from Streams of Interpretations with Optional Error Tolerance". In: *Proceedings of the 5th International Workshop "What can FCA do for Artificial Intelligence?" (FCA4AI 2016) co-located with the European Conference on Artificial Intelligence (ECAI 2016).* Ed. by Sergei Kuznetsov, Amedeo Napoli, and Sebastian Rudolph. Vol. 1703. CEUR Workshop Proceedings. The Hague, The Netherlands: CEUR-WS.org, 2016, pp. 9–16. URL: http://ceur-ws.org/Vol-1703/paper2.pdf.

[Kri16b] Francesco Kriegel. "NextClosures with Constraints". In: *Proceedings of the 13th International Conference on Concept Lattices and Their Applications (CLA 2016)*. Ed. by Marianne Huchard and Sergei Kuznetsov. Vol. 1624. CEUR Workshop Proceedings. Moscow, Russia: CEUR-WS.org, 2016, pp. 231–243. URL: http://ceur-ws.org/Vol-1624/paper18.pdf.

[Kri16c] Francesco Kriegel. "Parallel Attribute Exploration". In: *Proceedings of the 22nd International Conference on Conceptual Structures (ICCS 2016)*. Ed. by Ollivier Haemmerlé, Gem Stapleton, and Catherine Faron-Zucker. Vol. 9717. Lecture Notes in Computer Science. Annecy, France: Springer-Verlag, 2016, pp. 91–106. DOI: 10.1007/978-3-319-40985-6_8.

[Kri17] Francesco Kriegel. "Probabilistic Implication Bases in FCA and Probabilistic Bases of GCIs in $\mathcal{EL}^{\bot}$". In: *International Journal of General Systems* 46.5 (2017), pp. 511–546. DOI: 10.1080/03081079.2017.1349575.

[Kri18] Francesco Kriegel. "Acquisition of Terminological Knowledge in Probabilistic Description Logic". In: *KI 2018: Advances in Artificial Intelligence - 41st German Conference on AI, Berlin, Germany, September 24-28, 2018, Proceedings*. Ed. by Frank Trollmann and Anni-Yasmin Turhan. Vol. 11117. Lecture Notes in Computer Science. Berlin, Germany: Springer, 2018, pp. 46–53. DOI: 10.1007/978-3-030-00111-7_5.

[Kri19a] Francesco Kriegel. "Constructing and Extending Description Logic Ontologies using Methods of Formal Concept Analysis". Doctoral Thesis. Dresden, Germany: Technische Universität Dresden, 2019. URL: http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa2-360998.

[Kri19b] Francesco Kriegel. "Joining Implications in Formal Contexts and Inductive Learning in a Horn Description Logic". In: *15th International Conference on Formal Concept Analysis, ICFCA 2019, Frankfurt, Germany, June 25-28, 2019, Proceedings*. Ed. by Diana Cristea, Florence Le Ber, and Barış Sertkaya. Vol. 11511. Lecture Notes in Computer Science. Springer, 2019, pp. 110–129. DOI: 10.1007/978-3-030-21462-3_9.

[Kri19c] Francesco Kriegel. "Learning Description Logic Axioms from Discrete Probability Distributions over Description Graphs". In: *16th European Conference on Logics in Artificial Intelligence, JELIA 2019, Rende, Italy, May 7-11, 2019, Proceedings*. Ed. by Francesco Calimeri, Nicola Leone, and Marco Manna. Vol. 11468. Lecture Notes in Computer Science. Springer, 2019, pp. 399–417. DOI: 10.1007/978-3-030-19570-0_26.

[Kri20a] Francesco Kriegel. "Constructing and Extending Description Logic Ontologies using Methods of Formal Concept Analysis: A Dissertation Summary". In: *KI - Künstliche Intelligenz* 34 (2020), pp. 399–403. DOI: 10.1007/s13218-020-00673-8.

[Kri20b] Francesco Kriegel. "Most Specific Consequences in the Description Logic $\mathcal{EL}$". In: *Discrete Applied Mathematics* 273 (2020), pp. 172–204. DOI: 10.1016/j.dam.2019.01.029.

Related publications of Barış Sertkaya

[BS04] Franz Baader and Barış Sertkaya. "Applying Formal Concept Analysis to Description Logics". In: *Concept Lattices, Second International Conference on Formal Concept Analysis, ICFCA 2004, Sydney, Australia, February 23-26, 2004, Proceedings*. Ed. by Peter W. Eklund. Vol. 2961. Lecture Notes in Computer Science. Springer, 2004, pp. 261–286. DOI: 10.1007/978-3-540-24651-0_24.

[CBS19] Diana Cristea, Florence Le Ber, and Barış Sertkaya, eds. *Formal Concept Analysis - 15th International Conference, ICFCA 2019, Frankfurt, Germany, June 25-28, 2019, Proceedings*. Vol. 11511. Lecture Notes in Computer Science. Springer, 2019. DOI: 10.1007/978-3-030-21462-3.

[DS11] Felix Distel and Barış Sertkaya. "On the complexity of enumerating pseudo-intents". In: *Discret. Appl. Math.* 159.6 (2011), pp. 450–466. DOI: 10.1016/j.dam.2010.12.004.

[Gui+21] Ricardo Guimarães, Ana Ozaki, Cosimo Persia, and Barış Sertkaya. "Mining EL Bases with Adaptable Role Depth". In: *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*. AAAI Press, 2021, pp. 6367–6374. URL: https://ojs.aaai.org/index.php/AAAI/article/view/16790.

[Gui+23] Ricardo Guimarães, Ana Ozaki, Cosimo Persia, and Barış Sertkaya. "Mining EL Bases with Adaptable Role Depth". In: *Journal of Artificial Intelligence Research (JAIR)* 76 (2023). To appear.

[KS10] Léonard Kwuida and Barış Sertkaya, eds. *Proceedings of the 8th International Conference on Formal Concept Analysis (ICFCA 2010)*. Vol. 5986. Lecture Notes in Computer Science. Springer-Verlag, 2010.

[OSZ19] Sergei Obiedkov, Barış Sertkaya, and Denis Zolotukhin. "Probably Approximately Correct Completion of Description Logic Knowledge Bases". In: *Proceedings of the 32nd International Workshop on Description Logics, Oslo, Norway, June 18-21, 2019*. Ed. by Mantas Simkus and Grant E. Weddell. Vol. 2373. CEUR Workshop Proceedings. CEUR-WS.org, 2019. URL: http://ceur-ws.org/Vol-2373/paper-22.pdf.

[PS10] Rafael Peñaloza and Barış Sertkaya. “On the Complexity of Axiom Pinpointing in the $\mathcal{EL}$ Family of Description Logics”. In: *Proceedings of the Twelfth International Conference on Principles of Knowledge Representation and Reasoning, (KR 2010).* Ed. by Fangzhen Lin, Ulrike Sattler, and Miroslaw Truszczynski. AAAI Press, 2010, pp. 280, 289.

[PS17] Rafael Peñaloza and Barış Sertkaya. “Understanding the complexity of axiom pinpointing in lightweight description logics”. In: *Artif. Intell.* 250 (2017), pp. 80–104. DOI: 10.1016/j.artint.2017.06.002.

[Ser07] Barış Sertkaya. “Formal Concept Analysis Methods for Description Logics”. Doctoral Thesis. Dresden, Germany: Technische Universität Dresden, 2007. URL: http://nbn-resolving.de/urn:nbn:de:bsz:14-ds-1215598189927-85390.

[Ser10] Barış Sertkaya. “A Survey on how Description Logic Ontologies Benefit from Formal Concept Analysis”. In: *Proceedings of the 2010 International Conference on Concept Lattices and their Applications, (CLA 2010).* Ed. by Marzena Kryszkiewicz and Sergei Obiedkov. Vol. 672. CEUR Workshop Proceedings. 2010, pp. 2–21. URL: http://ceur-ws.org/Vol-672/paper2.pdf.

Other references

[And18] Simon Andrews. “A New Method for Inheriting Canonicity Test Failures in Close-by-One Type Algorithms”. In: *Proceedings of the Fourteenth International Conference on Concept Lattices and Their Applications, CLA 2018, Olomouc, Czech Republic, June 12-14, 2018.* Ed. by Dmitry I. Ignatov and Lhouari Nourine. Vol. 2123. CEUR Workshop Proceedings. CEUR-WS.org, 2018, pp. 255–266. URL: http://ceur-ws.org/Vol-2123/paper21.pdf.

[Baa+07] Franz Baader, Bernhard Ganter, Barış Sertkaya, and Ulrike Sattler. “Completing Description Logic Knowledge Bases Using Formal Concept Analysis”. In: *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007.* Ed. by Manuela M. Veloso. 2007, pp. 230–235. URL: http://ijcai.org/Proceedings/07/Papers/035.pdf.

[BD08] Franz Baader and Felix Distel. “A Finite Basis for the Set of $\mathcal{EL}$-Implications Holding in a Finite Model”. In: *Formal Concept Analysis, 6th International Conference, ICFCA 2008, Montreal, Canada, February 25-28, 2008, Proceedings.* Ed. by Raoul Medina and Sergei A. Obiedkov. Vol. 4933. Lecture Notes in Computer Science. Springer, 2008, pp. 46–61. DOI: 10.1007/978-3-540-78137-0_4.

[BM00] Franz Baader and Ralf Molitor. “Building and Structuring Description Logic Knowledge Bases Using Least Common Subsumers and Concept Analysis”. In: *Conceptual Structures: Logical, Linguistic,* *and Computational Issues, 8th International Conference on Conceptual Structures, ICCS 2000, Darmstadt, Germany, August 14-18, 2000, Proceedings*. Ed. by Bernhard Ganter and Guy W. Mineau. Vol. 1867. Lecture Notes in Computer Science. Springer, 2000, pp. 292–305. DOI: 10.1007/10722280_20.

[Bor14] Daniel Borchmann. "Learning terminological knowledge with high confidence from erroneous data". Doctoral thesis. Technische Universität Dresden, 2014. URL: https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa-152028.

[Dis11] Felix Distel. "Learning description logic knowledge bases from data using methods from formal concept analysis". Doctoral thesis. Technische Universität Dresden, 2011. URL: https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa-70199.

[Gan10] Bernhard Ganter. "Two Basic Algorithms in Concept Analysis". In: *Formal Concept Analysis, 8th International Conference, ICFCA 2010, Agadir, Morocco, March 15-18, 2010. Proceedings*. Ed. by Léonard Kwuida and Barış Sertkaya. Vol. 5986. Lecture Notes in Computer Science. Springer, 2010, pp. 312–340. DOI: 10.1007/978-3-642-11928-6_22.

[Gan84] Bernhard Ganter. *Two Basic Algorithms in Concept Analysis*. FB4-Preprint 831. Republished as [Gan10]. Darmstadt, Germany: Technische Hochschule Darmstadt, 1984.

[GD86] Jean-Luc Guigues and Vincent Duquenne. "Famille minimale d'implications informatives résultant d'un tableau de données binaires". In: *Mathématiques et Sciences Humaines* 95 (1986), pp. 5–18.

[GW99] Bernhard Ganter and Rudolf Wille. *Formal Concept Analysis - Mathematical Foundations*. Springer, 1999. DOI: 10.1007/978-3-642-59830-2.

[JKK21] Radek Janoštík, Jan Konečný, and Petr Krajča. "LinCbO: Fast algorithm for computation of the Duquenne-Guigues basis". In: *Inf. Sci.* 572 (2021), pp. 223–240. DOI: 10.1016/j.ins.2021.04.104.

[KOV10] Petr Krajča, Jan Outrata, and Vilém Vychodil. "Advances in Algorithms Based on CbO". In: *Proceedings of the 7th International Conference on Concept Lattices and Their Applications, Sevilla, Spain, October 19-21, 2010*. Ed. by Marzena Kryszkiewicz and Sergei A. Obiedkov. Vol. 672. CEUR Workshop Proceedings. CEUR-WS.org, 2010, pp. 325–337. URL: http://ceur-ws.org/Vol-672/paper29.pdf.

[Wil09] Rudolf Wille. "Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts". In: *Formal Concept Analysis, 7th International Conference, ICFCA 2009, Darmstadt, Germany, May 21-24, 2009, Proceedings*. Ed. by Sébastien Ferré and Sebastian Rudolph. Vol. 5548. Lecture Notes in Computer Science. Springer, 2009, pp. 314–339. DOI: 10.1007/978-3-642-01815-2_23.

[Wil82] Rudolf Wille. “Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts”. In: *Ordered Sets*. Ed. by Ivan Rival. Republished as [Wil09]. Dordrecht: Springer Netherlands, 1982, pp. 445–470.