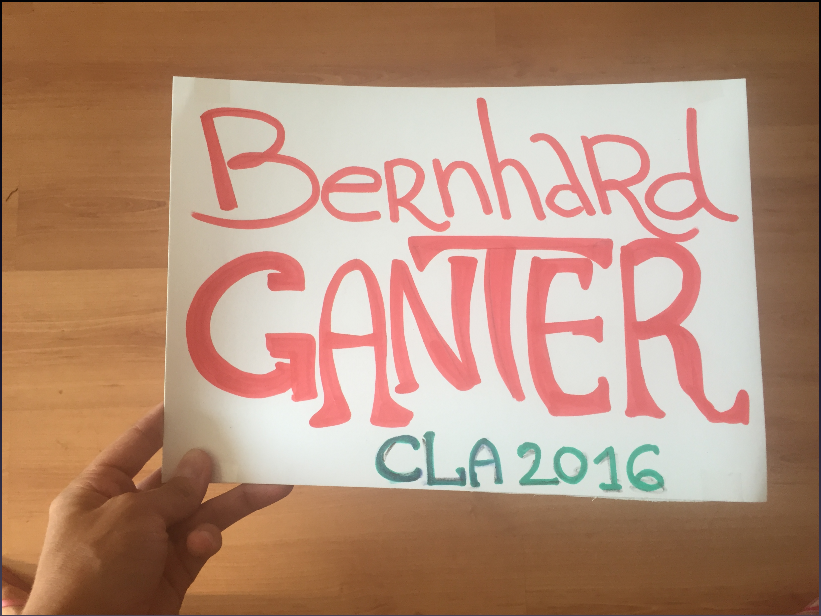

Bernhard
GANTER
CLA 2016

About our new book which has recently appeared

Bernhard Ganter

Institut für Algebra
TU Dresden

CLA Moscow, July 22, 2016

A great new book: ISBN 978-3-662-49290-1

Bernhard Ganter · Sergei Obiedkov

Conceptual Exploration

This is the first textbook on attribute exploration, its theory, its algorithms for applications, and some of its many possible generalizations. Attribute exploration is useful for acquiring structured knowledge through an interactive process, by asking queries to an expert. Generalizations that handle incomplete, faulty, or imprecise data are discussed, but the focus lies on knowledge extraction from a reliable information source.

The method is based on Formal Concept Analysis, a mathematical theory of concepts and concept hierarchies, and uses its expressive diagrams. The presentation is self-contained. It provides an introduction to Formal Concept Analysis with emphasis on its ability to derive algebraic structures from qualitative data, which can be represented in meaningful and precise graphics.

Computer Science

ISBN 978-3-662-49290-1

9 783662 492901

▸ springer.com

Ganter · Obiedkov

Conceptual Exploration

Bernhard Ganter · Sergei Obiedkov

Conceptual Exploration

Springer

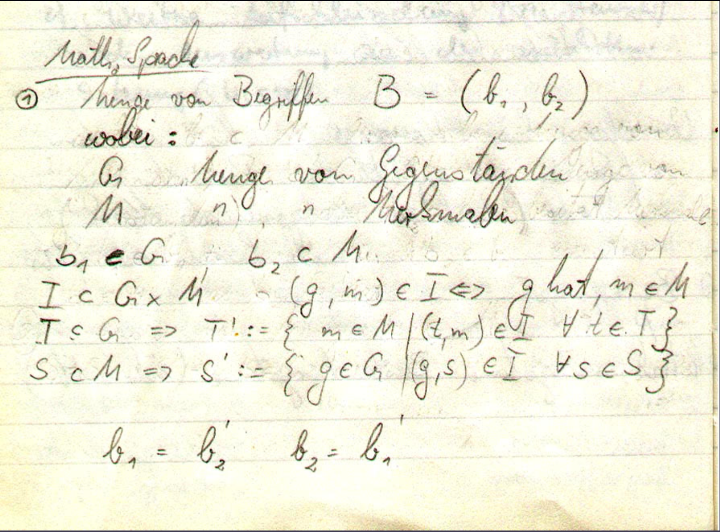
Things to do in Formal Concept Analysis

Bernhard Ganter

Institut für Algebra
TU Dresden

CLA Moscow, July 22, 2016

FCA was invented more than a generation ago



Math. Sprache

① Menge von Begriffen $B = (b_1, b_2)$

wobei:

G Menge von Gegenständen

M " " Merkmalen

$b_1 \in G \qquad b_2 \subset M$

$I \subset G \times M \qquad (g, m) \in I \Leftrightarrow g$ hat m, $m \in M$

$T \subseteq G \Rightarrow T' := \{ m \in M \mid (t, m) \in I \ \forall t \in T \}$

$S \subset M \Rightarrow S' := \{ g \in G \mid (g, s) \in I \ \forall s \in S \}$

$b_1 = b_2' \qquad b_2 = b_1'$

Seminar notes of Dec. 07, 1979, defining a formal concept.
(From K.E. Wolff's Ordnung, Wille, und Begriff.)

"Mittagsseminar", Dec. 13, 1979

13.12 Rudolf: Verbandstheorie als Begriffsalgebra

(Bezug: DIN 2331)

G Menge von "Gegenständen"

M Menge von "Merkmalen"

$I \subseteq G \times M$, $gIm :\Leftrightarrow g$ hat m

$A' := \{ m \in M : gIm \ \forall g \in A \}$ für $A \subseteq G$

$B' := \{ g \in G : gIm \ \forall m \in B \}$ für $B \subseteq M$

Satz: $'$ liefert Galoiskorrespondenz

d.h. $A_1 \subseteq A_2 \Rightarrow A_1' \supseteq A_2$

$B_1 \subseteq B_2 \Rightarrow B_1' \supseteq B_2$

$A \subseteq A''$, $B \subseteq B''$.

Begriff (A, B), falls $A' = A$, $B' = A$

A "Umfang", B "Inhalt" des Begriffes

$\mathfrak{B}(G, M, I)$ Menge aller Begriffe von (G, M, I)

$(A_1, B_1) \leq (A_2, B_2) :\Leftrightarrow A_1 \subseteq A_2$

$(\Leftrightarrow B_1 \supseteq B_2)$

(Unterbegriff $\leftrightarrow$ Oberbegriff)

(damit erhält man einen Verband)

Bsp: Klassifikation von Musikinstrumenten (vorgeführt an 8 mitgebrachten Flöten)

Satz: V vollst. Verband $\Rightarrow V \cong \mathfrak{B}(V, V, \leq)$

Zusatz: V Verband

$\Rightarrow \mathfrak{B}(V, V, \leq)$ Dedekind-McNeillsche Hülle

Begriffsart: $\mathfrak{A} \subseteq \mathfrak{B}(G, M, I)$ mit

$\{ A : (A, B) \in \mathfrak{A} \}$ Klasseneinteilung von G

$\mathfrak{A}_1 \wedge \mathfrak{A}_2 := \{ (A_1 \cap A_2, (A_1 \cap A_2)') : (A_1, A_1') \in \mathfrak{A}_1, (A_2, A_2') \in \mathfrak{A}_2$ und $A_1 \cap A_2 \neq \emptyset \}$

$\mathfrak{A}_1 \vee \mathfrak{A}_2 = \ldots$

Let's face it . . .

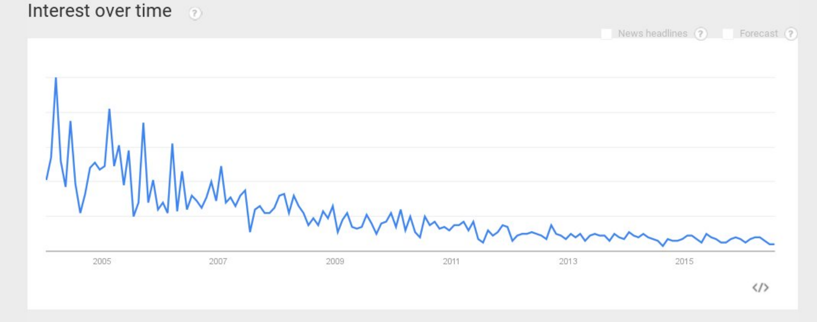
- the first generation of FCA researchers is leaving the stage

Let's face it . . .

- the first generation of FCA researchers is leaving the stage
- today's young researchers were not even born when FCA emerged,

Let's face it . . .

- the first generation of FCA researchers is leaving the stage
- today's young researchers were not even born when FCA emerged,
- and the search term frequency for FCA is fading.



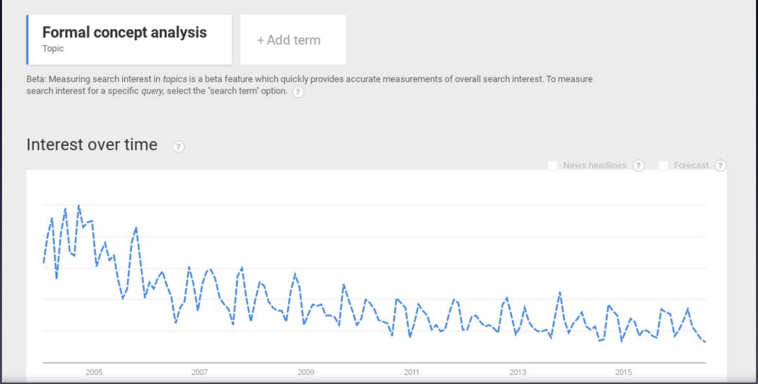
Research areas come and go.

On the other hand . . .

- We see bright people with innovative ideas on great conferences like this one,

On the other hand . . .

- We see bright people with innovative ideas on great conferences like this one,
- the interest in the topic FCA is not fading as quickly as the number of searches for the term FCA,



. . . and moreover, . . .

- FCA really has something to offer,

. . . and moreover, . . .

- FCA really has something to offer,
- and the word of FCA being useful has spread outside the CLA/ICFCA communities.

. . . and moreover, . . .

- FCA really has something to offer,
- and the word of FCA being useful has spread outside the CLA/ICFCA communities.

From a paper which I recently had to review:

"Formal Concept Analysis (FCA) has [. . .] proven useful in a wide range of application areas such as medicine, psychology, sociology, linguistics, archaeology, anthropology, biology, chemistry, civil engineering, electrical engineering, information science, library science, information technology, software engineering, computer science and even mathematics itself."

Does FCA deserve a future?

Personally, I am convinced that FCA has unfolded some fundamental truths, and has worked out supporting theory. This will remain, but not necessarily as a field of research.

Does FCA deserve a future?

Personally, I am convinced that FCA has unfolded some fundamental truths, and has worked out supporting theory. This will remain, but not necessarily as a field of research.

But come on, you all have invested at least some time in FCA and thus should hope for it to play a good role in the <u>near</u> future. So what can we do?

Does FCA deserve a future?

Personally, I am convinced that FCA has unfolded some fundamental truths, and has worked out supporting theory. This will remain, but not necessarily as a field of research.

But come on, you all have invested at least some time in FCA and thus should hope for it to play a good role in the near future. So what can we do?

There is no master plan.

Does FCA deserve a future?

Personally, I am convinced that FCA has unfolded some fundamental truths, and has worked out supporting theory. This will remain, but not necessarily as a field of research.

But come on, you all have invested at least some time in FCA and thus should hope for it to play a good role in the near future. So what can we do?

There is no master plan. Let us try to make one.

Does FCA deserve a future?

Personally, I am convinced that FCA has unfolded some fundamental truths, and has worked out supporting theory. This will remain, but not necessarily as a field of research.

But come on, you all have invested at least some time in FCA and thus should hope for it to play a good role in the near future. So what can we do?

There is no master plan. Let us try to make one.

I shall give here some thoughts on a strategy for FCA, knowing well that these will be insufficient. I believe that in order to be successful, we must better understand what the weaknesses and strengths of our field are.

What was so exciting right from the start?

Immediately after the discoveries of 1979, the Darmstadt group started the FCA project with great enthusiasm. We knew that we had something important in our hands.

And that was: a strikingly simple and convincing interpretation of lattices as hierarchies of concepts.

What was so exciting right from the start?

Immediately after the discoveries of 1979, the Darmstadt group started the FCA project with great enthusiasm. We knew that we had something important in our hands.

And that was: a strikingly simple and convincing interpretation of lattices as hierarchies of concepts.

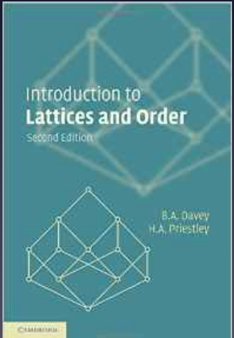
The impact of this discovery was obvious:

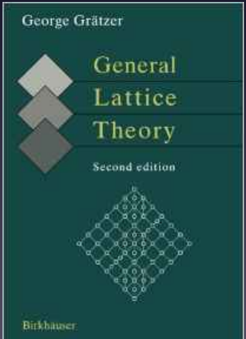
- it gave new meaning to the mathematical theory of complete lattices, and
- it provided mathematical tools for the age-old philosophical theory of concepts.

Lattice Theory was not amused

We were astonished to learn that the lattice theory community did not share our enthusiasm. They were (until today) largely uninterested, sometimes even hostile towards FCA.

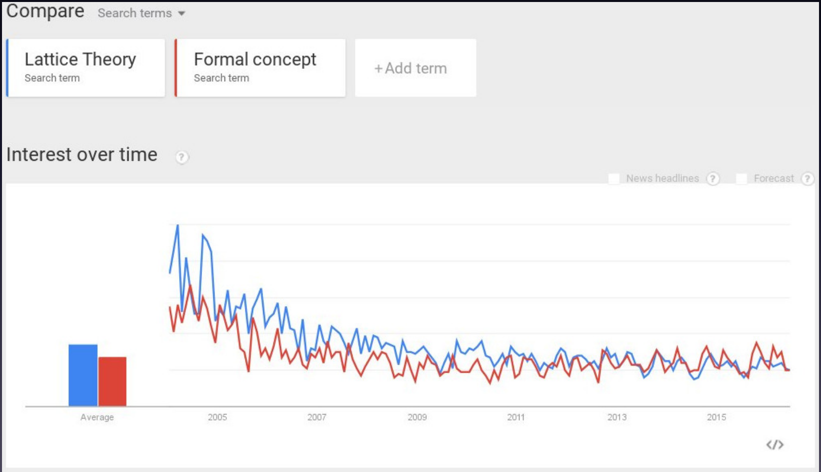
With some notable exceptions, see the books by Davey+Priestley and by Grätzer.





Should we better integrate into LT?

Probably not, since it would not help much. Lattice theory is not mainstream either.



I don't believe in such graphics anyway.

Wehrung's paper was rejected by JAMS in 2007

F. Wehrung's celebrated solution of a 50-year-old problem in LT was rejected for publication by the JAMS for its "lack of interaction with other parts of mathematics."

Doron Zeilberger sarcastically comments this in his "Opinion 81":

> Lattice Theory is old hat. Dilworth and Birkhoff, and even Paul Cohen, are out-of-style. Maybe try to submit your stuff to a golden-oldies radio station, but not to a mainstream contemporary rock station . If you want to be considered for publication in JAMS, you need some good buzz words: "Kazhdan-Lusztig", "Langlands program", "Ricci flow", etc. etc. In fact, if the editors of JAMS would have been computer savvy, they could have saved themselves lots of trouble by writing a short Sed or Awk or Perl script eliminating submissions by key-words.

Applied lattice theory

FCA can be understood as applied lattice theory. But there are different approaches to that term. Compare!

- From Birkhoff's Lattices and their applications (1938):

lattice theory, is an essential preliminary to the full understanding of logic, set theory, probability, functional analysis, projective geometry, the decomposition theorems of abstract algebra, and many other branches of mathematics.

- Wille's first version of the basic theorem (1979):

Bsp: Klassifikation von Musikinstrumenten
(vorgeführt an 8 mitgebrachten Flöten)

Satz: V vollst. Verband $\Rightarrow V \cong \mathfrak{B}(V, V, \leq)$

Zusatz: V ~~vollst.~~ Verband

$\Rightarrow \mathfrak{B}(V, V, \leq)$ Dedekind-McNeillsche Hülle

Where the beef is

Understanding the key value of FCA is crucial for making the best steps in the future:

Where the beef is

Understanding the key value of FCA is crucial for making the best steps in the future:

The essential importance of FCA is that it establishes a connection between an area of mathematics and the fundamental cognitive notion of concepts and their hierarchies.

- The emergence of FCA was not an important breakthrough <u>inside mathematics</u>.
- FCA provides a model for the <u>doctrine of concepts</u>. It is strongly simplifying, but has the advantage of mathematical reliability and a powerful theory.
- The <u>pragmatics</u> of FCA is not yet generally accepted, in spite of a large number of convincing examples.

Dos and Don’ts

So how can we forward FCA?

I try to give some “dos and don’ts”.

Feel invited to add your own ones.

Don'ts

- Don't try to convince mathematicians. Forget it. FCA does not fit into their business plan. Be mathematical in the sense that your language is formal and precise, and your results are based on good theorems with proofs.
- Don't try to teach FCA <u>theory</u> to non-mathematicians. Instead, convince them by providing powerful methodology, which they can handle and control.
- Don't attempt to make FCA a big thing in Computer Science. Small is beautiful. But be aware of the standards and trends in information technology.
- Don't compete with techniques of big data mining. They have powerful methods, many of which are fundamentally flawed, because they rely on using metric methods for non-metric data.

Dos

I believe that the following activities should be beneficial for FCA. I thus shall go into these in more detail.

- Make usefulness explicit.
- Improve the usability.
- Broaden.
- Conquer.
- Standardize.
- Modernize.
- Find better forms of teaching FCA.
- Deepen.

Making the usefulness explicit

FCA with its “human centered approach” has a different goal than most numeric methods. Rather than producing a <u>result</u> which can be expressed in a few bytes, it offers <u>expressive and reliable diagrams</u> which are of great help for human data interpretation.

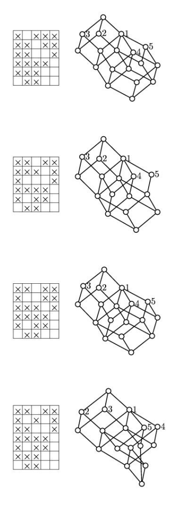
Making the usefulness explicit

FCA with its "human centered approach" has a different goal than most numeric methods. Rather than producing a <u>result</u> which can be expressed in a few bytes, it offers <u>expressive and reliable diagrams</u> which are of great help for human data interpretation.

But are they really? My impression is that not many people can read concept lattice diagrams well enough to find them really useful. One indication is that even diagrams produced by members of the FCA community are often ugly and even faulty.

We need to make explicit why lattice diagrams are useful, how they empower the user to draw conclusions that are difficult to find otherwise.

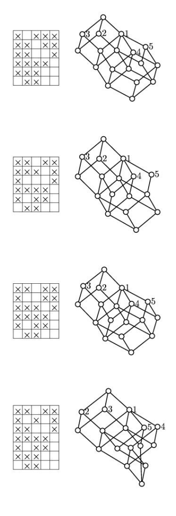
Improving the usability: computer programs



Christian Zschalig

Many computer programs are available for FCA. But my impression is that CONEXP is still the most popular one, although it is outdated and rather limited in its functionality. CONEXP is intuitive and easy to understand.

Improving the usability: computer programs



Christian Zschalig

Many computer programs are available for FCA. But my impression is that CONEXP is still the most popular one, although it is outdated and rather limited in its functionality. CONEXP is intuitive and easy to understand.

It is however not at all easy to draw good diagrams with CONEXP, and the program offers little help for interpretations.

We need intuitive computer programs producing good lattice diagrams whenever possible. These should also offer good label positioning and hints for interpretations.

Algorithms for diagram layout

There are several promising approaches to automated diagram drawing, and several interesting implementations. A major problem is that we do not know what makes a diagram good.

Algorithms for diagram layout

There are several promising approaches to automated diagram drawing, and several interesting implementations. A major problem is that we do not know what makes a diagram good.

Minimizing the number of edge crossings seems not to help very much.

Zschalig combined the techniques of "force directed placement", "additivity" and "conflict distance" do generate diagrams of all 14664 concept lattices with five attributes. His results are remarkable, but can be improved.

Algorithms for diagram layout

There are several promising approaches to automated diagram drawing, and several interesting implementations. A major problem is that we do not know what makes a diagram good.

Minimizing the number of edge crossings seems not to help very much.

Zschalig combined the techniques of “force directed placement”, “additivity” and “conflict distance” do generate diagrams of all 14664 concept lattices with five attributes. His results are remarkable, but can be improved.

In many cases it is a bad idea to put labels inside a diagram. A default positioning of labels outside would be welcome.

Improving the usability: Recipes

We have an impressive record of documented applications of FCA. Here are some of many:

- A survey of formal concept analysis support for software engineering activities (Tilley et al.)
- Linguistic applications of formal concept analysis (Priss)
- Learning concept hierarchies from text corpora using formal concept analysis (Hotho et al.)
- Formal concept analysis in information science (Priss)
- Formal concept analysis in knowledge discovery: a survey (Poelmans et al.)

But all that seems not to be sufficient. Perhaps "recipes", i.e., standardized procedures for basic data analysis tasks, can help.

Broadening the approach

At the end of the 1980ies it became apparent that the theoretical kernel of FCA was essentially completed. It was time to embed FCA into a larger framework.
Many variants have been discussed, among them

- topological formal contexts,
- triadic formal contexts,
- logical information systems and pattern structures,
- fuzzy concept analysis,
- iceberg concept lattices, faulty data,
- concept analysis for relational data.

Topological formal contexts

Concept lattices are complete lattices, and according to the basic theorem, every complete lattice is isomorphic to a concept lattice.

But not all lattices are complete. Thus there are (infinite) lattices, which are not isomorphic to concept lattices.

Topological formal contexts

Concept lattices are complete lattices, and according to the basic theorem, every complete lattice is isomorphic to a concept lattice.

But not all lattices are complete. Thus there are (infinite) lattices, which are not isomorphic to concept lattices.

These can be represented as lattices of topological formal concepts of formal context with suitable topologies (Hartung, Sacarea).

Topological formal contexts

Concept lattices are complete lattices, and according to the basic theorem, every complete lattice is isomorphic to a concept lattice.

But not all lattices are complete. Thus there are (infinite) lattices, which are not isomorphic to concept lattices.

These can be represented as lattices of topological formal concepts of formal context with suitable topologies (Hartung, Sacarea).

This seems, however, not to be of interest to many.

Triadic formal contexts

It is straightforward to generalize the definition of a formal context and a formal concept to n dimensions ("Polyadic concept analysis", Voutsadakis 2002).

Lehmann and Wille have investigated the triadic case (i.e., n=3) in 1995. Wille argued that higher values of n are much less important, according to C. S. Peirce.

Triadic formal contexts

It is straightforward to generalize the definition of a formal context and a formal concept to n dimensions ("Polyadic concept analysis", Voutsadakis 2002).

Lehmann and Wille have investigated the triadic case (i.e., n=3) in 1995. Wille argued that higher values of n are much less important, according to C. S. Peirce.

The mathematical theory of Triadic Concept Analysis turns out to be difficult, and many of the basic questions are yet unanswered.

Why triadic contexts are important?

Triadic formal contexts

It is straightforward to generalize the definition of a formal context and a formal concept to n dimensions ("Polyadic concept analysis", Voutsadakis 2002).

Lehmann and Wille have investigated the triadic case (i.e., n=3) in 1995. Wille argued that higher values of n are much less important, according to C. S. Peirce.

The mathematical theory of Triadic Concept Analysis turns out to be difficult, and many of the basic questions are yet unanswered.

Why triadic contexts are important?
They are the FCA data type corresponding to RDF!

LIS and pattern structures

“Logical Information Systems” (Ferre and Ridoux) and “Pattern Structures” (Kuznetsov at. al) are not generalizations in the strict sense, since they also result in complete lattices.

But these approaches offer alternative language for concept analysis, and are better suited in certain situations, and therefore are successful and welcome facets of Formal Concept Analysis.

LIS and pattern structures

"Logical Information Systems" (Ferre and Ridoux) and "Pattern Structures" (Kuznetsov at. al) are not generalizations in the strict sense, since they also result in complete lattices.

But these approaches offer alternative language for concept analysis, and are better suited in certain situations, and therefore are successful and welcome facets of Formal Concept Analysis.

One must however keep in mind that the different versions translate into each other, and should avoid doubling the effort.

Fuzzy concept analysis and faulty data

Remarkable research has been accumulated on the topic of fuzzy concept analysis. As in the case of LIS and Pattern Structures, most of the theory is not a generalization in the strict sense.

But it is beyond doubt that fuzzy concept analysis meanwhile has developed a substantial amount of theory and can be regarded as an area of its own.

From the beginning, there was demand for a "fuzzy" version of FCA, because real data tends to be noisy, imprecise, unreliable, or faulty.

Fuzzy concept analysis and faulty data

My impression is that the pragmatic aspect of fuzzy concept analysis is much less developed than that of FCA.

It seems that some users follow a strategy which may be caricatured as follows: "My data is faulty, so I must use a method that can deal with that. But it does not really matter which method I use and how, because my data is faulty anyway."

Fuzzy concept analysis and faulty data

My impression is that the pragmatic aspect of fuzzy concept analysis is much less developed than that of FCA.

It seems that some users follow a strategy which may be caricatured as follows: "My data is faulty, so I must use a method that can deal with that. But it does not really matter which method I use and how, because my data is faulty anyway."

I would appreciate to see an introduction that carefully distinguishes between fuzzy, probabilistic, and fault tolerant conceptual data analysis. It should help the users to apply the methods which are meaningful for the respective analyses.

Contextual Logic

"Contextual Logic" was Wille's main project of extending FCA to relational data. His plan was to provide a formalization of Aristoteles' and Kant's doctrine of "Concept – Judgment – Conclusion", in which FCA would formalize the "concept" part.

For formalizing "Judgments" he used a mathematically correct variant of Sowa's "Conceptual Graphs", and the inference logic of these formalized "Conclusion".

Contextual Logic

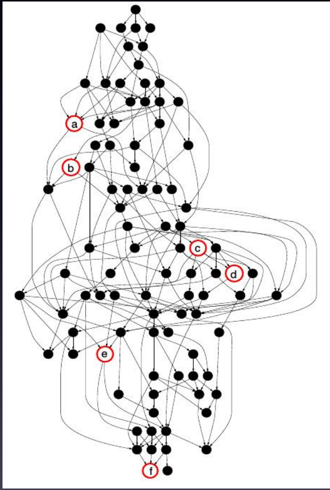
"Contextual Logic" was Wille's main project of extending FCA to relational data. His plan was to provide a formalization of Aristoteles' and Kant's doctrine of "Concept – Judgment – Conclusion", in which FCA would formalize the "concept" part.

For formalizing "Judgments" he used a mathematically correct variant of Sowa's "Conceptual Graphs", and the inference logic of these formalized "Conclusion".

Wille's approach, although far developed, has not been well received by the FCA community, probably due to its cumbersome language. Serveral other approaches to relational data are in use. We need unification!

Including other areas

For example, the remarkably successful theory of Knowledge Spaces, developed by Doignon and Falmagne, is in its basic definitions very close to FCA and can certainly be nicely expressed in FCA language.



FCA, with its broader theoretical background, most likely can substantially add to Knowledge Space theory. This has already been started for a special branch "Competence based Knowledge Space Theory" (CbKST), using the technique of Boolean matrix decomposition.

KST

Knowledge Space Theory is about learning to solve certain “tasks”, which represent the atoms of a field of study.

Tasks are ordered by the “prerequisite” order. The “knowledge state” of a learner consists of all tasks this learner can master. Such states necessarily are downsets of the prerequisite order. The set of all such downsets is the “knowledge space” of the field of study.

KST

Knowledge Space Theory is about learning to solve certain "tasks", which represent the atoms of a field of study.

Tasks are ordered by the "prerequisite" order. The "knowledge state" of a learner consists of all tasks this learner can master. Such states necessarily are downsets of the prerequisite order. The set of all such downsets is the "knowledge space" of the field of study.

Recall that the downsets are precisely the concept extents of the contra-ordinal scale, i.e., of the formal context $(\mathrm{P}, \mathrm{P}, \not\geq)$, where $(\mathrm{P}, \leq)$ is the prerequisite order.

CbKST

Competence based Knowledge Space Theory assumes that there are reasons called "competences" that enable a learner to master a task.

This leads to three formal contexts

(Learners, Tasks, masters)
(Learners, Competences, has)
(Competences, Tasks, enables),

the first one of which is observable, while the other two are part of pedagogical theory.

The assumption is that a learner masters a task iff (s)he has at least one competence for that. This leads to boolean product decomposition of formal contexts.

Boolean decomposition of formal contexts

$$\mathrm{l} \mathrel{\square} \mathrm{q} \iff \exists_{\mathrm{C}\in\mathcal{C}} \, (\mathrm{l} \circ \mathrm{C} \text{ and } \mathrm{C} \models \mathrm{q}),$$

is symbolised by

$$(\mathrm{L}, \mathrm{Q}, \square) = (\mathrm{L}, \mathcal{C}, \circ) \cdot (\mathcal{C}, \mathrm{Q}, \models).$$

Of course, $\mathrm{l} \mathrel{\square} \mathrm{q}$ is interpreted as "learner l masters task q", $\mathrm{l} \circ \mathrm{C}$ as "learner l has competence C" and $\mathrm{C} \models \mathrm{q}$ reads as "competence C suffices for mastering question q".

Boolean decomposition of formal contexts

$$\mathrm{l} \mathrel{\Box} \mathrm{q} \iff \exists_{\mathrm{C}\in\mathcal{C}} (\mathrm{l} \circ \mathrm{C} \text{ and } \mathrm{C} \models \mathrm{q}),$$

is symbolised by

$$(\mathrm{L}, \mathrm{Q}, \Box) = (\mathrm{L}, \mathcal{C}, \circ) \cdot (\mathcal{C}, \mathrm{Q}, \models).$$

Of course, $\mathrm{l} \mathrel{\Box} \mathrm{q}$ is interpreted as "learner l masters task q", $\mathrm{l} \circ \mathrm{C}$ as "learner l has competence C" and $\mathrm{C} \models \mathrm{q}$ reads as "competence C suffices for mastering question q".

It was discovered by Belohlavek and others that such boolean decompositions correspond to converings of the incidence relation by formal concepts. These are in turn known to be equivalent to embeddings into of the concept lattice of the complimentary context into boolean lattices.

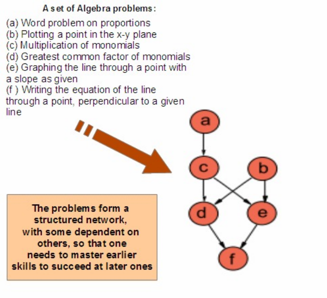
Learning from other areas

FCA can probably learn a lot from KST's elaborated stochastic tools, which they have developed for their practical applications.

Learning from other areas

FCA can probably learn a lot from KST's elaborated stochastic tools, which they have developed for their practical applications.

And perhaps, we could also learn from their successful business model.



Standardizing language and appearance

That Rudolf Wille's research group did not find much resonance among other math researchers was somewhat disappointing. But it had the big advantage that they could work out their theory in a systematic and unified language.

Wille's philosophical training resulted in strange terms such as "formal concept" and "conceptual scaling", which his mathematically trained students did not like very much. Meanwhile, all this has become best practice.

As part of its corporate identity, Formal Concept Analysis benefits from its systematic, meaningful and mathematically precise terminology and its well-defined graphical representations.

Modernizing the research

Almost ten years ago, Rudolph, Krötzsch, and Hitler criticized the "non-impact of conceptual structures on the semantic web". Conceptual Graphs and Formal Concept Analysis were missing opportunities for showing their effectiveness for problems of this booming area. Unfortunately, this still applies.

$\mathcal{EL}^+$ OWL RDF TRIPLE STORE HTML6

Finding better forms of teaching FCA

I have heard rumor that here at HSE they work on up-to-date course material for FCA. Great!

Finding better forms of teaching FCA

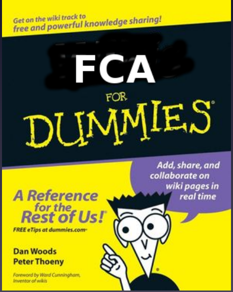
I have heard rumor that here at HSE they work on up-to-date course material for FCA. Great!

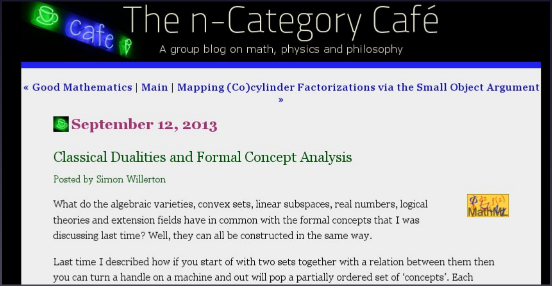
But today even university students demand other presentation forms than university courses only. We need materials in the language of the present time.

Finding better forms of teaching FCA

I have heard rumor that here at HSE they work on up-to-date course material for FCA. Great!

But today even university students demand other presentation forms than university courses only. We need materials in the language of the present time.





Deepening the theory

Although the core theory of FCA was worked out in the 1980ies, there is still room for basic research.

Mathematical theories take a long time to mature.

Deepening the theory

Although the core theory of FCA was worked out in the 1980ies, there is still room for basic research.

Mathematical theories take a long time to mature.

Among the discoveries of the last years are

- the FCA interpretation of Boolean matrix factorization and
- the characterization of large concept lattices via contra-nominal scales.

Deepening the theory

Although the core theory of FCA was worked out in the 1980ies, there is still room for basic research.

Mathematical theories take a long time to mature.

Among the discoveries of the last years are

- the FCA interpretation of Boolean matrix factorization and
- the characterization of large concept lattices via contra-nominal scales.

The area with the most challenging open mathematical problems seems to be that of triadic FCA.

Challenges

Most urgent:

Challenges

Most urgent:

Bring peace to Concept Analysis of relational data.

Challenges

Most urgent:

Bring peace to Concept Analysis of relational data.

There are too many different approaches to the same thing, and it seems that what is natural to one group is not acceptable to the other.

We need to converge to a natural, logically and philosophically sound version.

Programming challenges

Many activities are underway. LatViz and FCART2 are remarkable examples. I mention an additional few possible ones.

- Design an algorithm that, without much ado, draws nice diagrams for all lattices with up to 20 elements. Find a default way to label diagrams.
- Write a modern diagram editor.
- Allow diagrams as graphical input.
- Modernize the R package for FCA, and make standard techniques of FCA accessible for R users.

Some books to write

Many activities are underway. “Conceptual Exploration” is a recent example. I mention an additional few possible ones.

- Write an introduction to FCA with a focus on what you can do with FCA. Definitions and theorems come last.
- Write the book “Formal Concept Analysis: Pragmatic Foundations”.
- Write the book “Formal Concept Analysis: Philosophical Foundations”.
- Write the book “Knowledge Spaces from an FCA perspective”.

Mathematical challenges

FCA was criticized for not generating enough interesting math problems. Here are some:

- What is the maximal possible number of triconcepts of an $a \times b \times c$ triadic context?
- What are the triadic arrow relations?
- What is triadic distributivity?
- Is there a way to predict if a formal context has moderately many or horribly many formal concepts?
- Is there an effective way to generate reduced formal contexts with 10 objects at random?
- What are residuated formal contexts?

Mathematical challenges

- What is the maximal possible number of triconcepts of an a × b × c triadic context?

Obviously, a (dyadic) formal context can have at most

$$2^{\min\{\#\text{objects},\#\text{attributes}\}}$$

formal concepts, and this bound is always sharp.

I am not aware of a similarly simple bound for the number of triconcepts of a triadic context.

Mathematical challenges

- What are the triadic arrow relations?

In FCA, the arrow relations are of crucial importance. They encode, how join- and meet-irreducible elements are connected in congruence relations.

I am not aware of a corresponding notion for triadic contexts. I do not even have a guess for the shape of such arrows.

Mathematical challenges

- What is triadic distributivity?

The smallest non-trivial equational class of lattices is that of distributive lattices. These are the sublattices of the powerset lattices, and of products of chains.

There are powerset trilattices, and there are trichains. But I do not have a definition of a distributive trilattice.

Mathematical challenges

- Is there a way to predict if a formal context has moderately many or horribly many formal concepts?

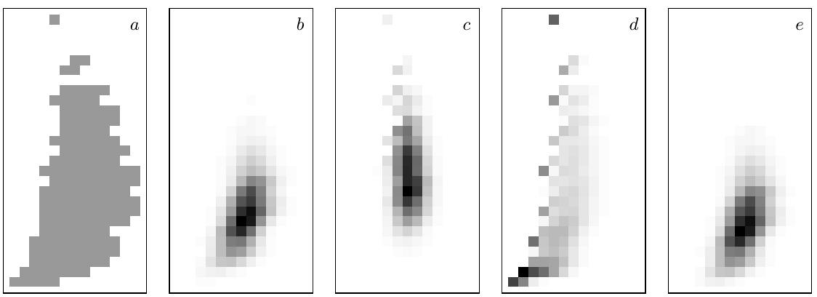
Kuznestov has shown that predicting the number of concepts of a formal context is $\#\mathcal{P}$-hard.

Albano and Chernomaz have shown that a formal context not containing certain substructures cannot have very many concepts.

So how precisely can we predict the number in reasonable time?

Mathematical challenges

- Is there an effective way to generate reduced formal contexts with 10 objects at random (equally distributed)?



Mathematical challenges

- Is there an effective way to generate reduced formal contexts with 10 objects at random (equally distributed)?

It is easy to generate reduced formal contexts with 6 objects at random, since there are only 75 973 751 474 of them. Make a list, and then choose numbers between 1 and 75973751474 at random.

Mathematical challenges

- Is there an effective way to generate reduced formal contexts with 10 objects at random (equally distributed)?

It is easy to generate reduced formal contexts with 6 objects at random, since there are only 75 973 751 474 of them. Make a list, and then choose numbers between 1 and 75973751474 at random.

This does not work for n = 7, because there are 14 087 648 235 707 352 472 of them, which are too many.

Can we nevertheless generate reduced contexts with n = 10 objects at random?

Mathematical challenges

- What are residuated formal contexts?

(Complete) residuated lattices are important for fuzzy set theory. They are lattices with additional algebraic structure (monoid, residuation).

In some other cases, concept lattices with extra operations can be naturally obtained from formal contexts with extra operations.

Is there something similar for residuated lattices?

The End