SmartFCA

Chronicles of Concept lattices

Unveiling Structure in the Software World

Marianne Huchard

LIRMM, Univ. Montpellier, CNRS, France

LIRMM

27th ACM International Systems and Software Product Line Conference

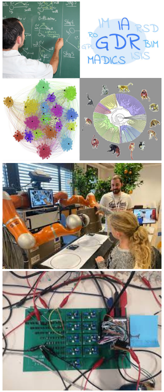
SPLC 2023

27th Systems and Software Product Line Conference

TOKYO

Table of Contents

Montpellier

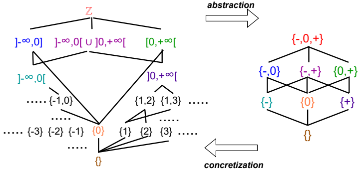


LIRMM

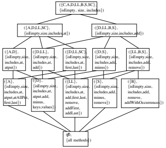
MaREL Team

Reuse, Reengineering, MDE, CBSE, IA4SE, Theorem proving, FCA

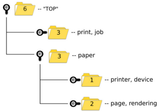
What is common to ...? (and to many others)



Domains in abstract interpretation
Cousot & Cousot, 1979



Revisiting Smalltalk classes
Godin et Mili, 1993



Concept/feature Location
Poshyvanyk et al., 2010

Within the *"Bad Smell"* category, we have the schema *Broken* super *send Chain* (shown in Figure 2). It is composed of the following elements and properties:

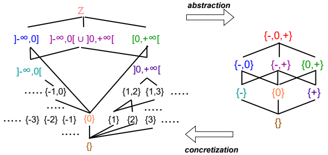
- *C invokes i via super*: {representBinaryOn:, =} are *super-called* in SortedCollection
- *i is concrete locally*: {representBinaryOn:, =} has concrete behavior in SortedCollection.
- *i is concrete in ancestor C_1 of C* : {representBinaryOn:, =} has concrete behavior in ancestor SequenceableCollection of SortedCollection.
- *i is concrete in descendant C_1 of C* : {representBinaryOn:, =} has concrete behavior in descendant SortedCollectionWithPolicy of SortedCollection.

Dependencies in OO code
Arevalo et al., 2005

$$s \wedge \neg d \Rightarrow s_2 \quad {}^F \qquad p \wedge \neg s_2 \wedge \neg i \Leftrightarrow c_p \quad {}^F$$
$$p \wedge \neg s_2 \wedge \neg d \Leftrightarrow c_p \quad {}^F \qquad n \wedge p \Leftrightarrow f_n$$
$$t \wedge p \Leftrightarrow f_t \qquad d \wedge i \Leftrightarrow s_3$$
$$\neg s_2 \wedge i \Leftrightarrow s_3 \qquad \neg s_2 \wedge d \Leftrightarrow s_3$$
$$s \wedge \neg s_2 \Leftrightarrow s_3 \qquad t \wedge \neg c_p \wedge \neg s_2 \wedge \neg i \Leftrightarrow c_t$$

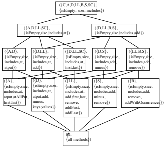
Feature model and constraints extraction, Ryssel et al., 2011

Made with Galois theory!



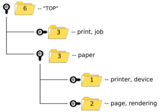
Correspondence / connection

Domains in abstract interpretation,
Cousot & Cousot, 1979



Clustering and classification

Revisiting Smalltalk classes, Godin et Mili,
1993



Clustering and classification

Concept Location, Poshyvanyk et al., 2010

Within the *"Bad Smell"* category, we have the schema *Broken* super *send Chain* (shown in Figure 2). It is composed of the following elements and properties:

- *C invokes i via super*: {representBinaryOn:, =} are *super-called* in SortedCollection
- *i is concrete locally*: {representBinaryOn:, =} has concrete behavior in SortedCollection.
- *i is concrete in ancestor C_1 of C* : {representBinaryOn:, =} has concrete behavior in ancestor SequenceableCollection of SortedCollection.
- *i is concrete in descendant C_1 of C* : {representBinaryOn:, =} has concrete behavior in descendant SortedCollectionWithPolicy of SortedCollection.

Knowledge pattern

Dependencies in OO code,
Arevalo et al., 2005

$$s \wedge \neg d \Rightarrow s_2 \quad F$$
$$p \wedge \neg s_2 \wedge \neg d \Leftrightarrow c_p \quad F$$
$$t \wedge p \Leftrightarrow f_t$$
$$\neg s_2 \wedge i \Leftrightarrow s_3$$
$$s \wedge \neg s_2 \Leftrightarrow s_3$$

$$p \wedge \neg s_2 \wedge \neg i \Leftrightarrow c_p \quad F$$
$$n \wedge p \Leftrightarrow f_n$$
$$d \wedge i \Leftrightarrow s_3$$
$$\neg s_2 \wedge d \Leftrightarrow s_3$$
$$t \wedge \neg c_p \wedge \neg s_2 \wedge \neg i \Leftrightarrow c_t$$

Logical formula

Feature model and constraints extraction,
Ryssel et al., 2011

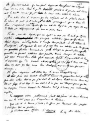
Table of Contents

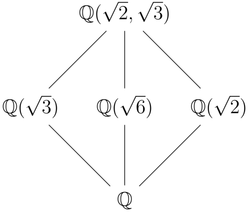
Landmarks: Concept *in* philosophy

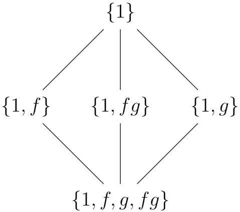
Knowledge theories, from Jean Ladrière, Encyclopedia Universalis

- Two knowledge modes
  - concrete, sensible, singular, individuals, intuition
  - abstractions, universals
- A concept can be seen as
  - a general representation
  - a mediation between the concrete and abstract levels
- Controversies
  - relations between predicates and concepts, meaning vs. denotation (Frege)
  - views on universals:
    - platonism (participating concrete objects, extension) - Plato, Carnap
    - nominalism (linguistic entities, ontology of individuals) - Goodman, Quine
    - conceptualism (individuals and properties, intension) - Aristotle, Abélard, Occam, Kant, Locke
  - some ideas are independent of the sensible experience - Kant, Descartes

Landmarks: Galois connection *in* Algebra (1830)



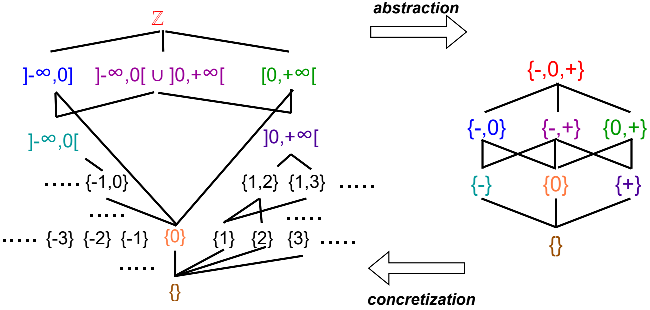


Wikipedia: Evariste Galois (1811-1932); The last page of its last manuscript; A Duel in the Bois De Boulogne, Near Paris, wood-engraving after G. Durand, Harper's Weekly (Jan. 1875); H. Lecomte, Combat de la rue de Rohan le 29 juillet 1830;

The Galois Fundamental theorem:
An example of the **1-1 correspondence** between the lattice of its intermediate fields and the lattice of the subgroups of its Galois group

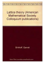
Landmarks: Galois connection *in* Lattice theory

An example with abstract interpretation (Cousot & Cousot)
Reason on signs of variables whose values are integer sets



A set of signed integers is associated to the set of their signs ∪ {0}
A set of signs ∪ {0} is associated to the set of all integers having these signs

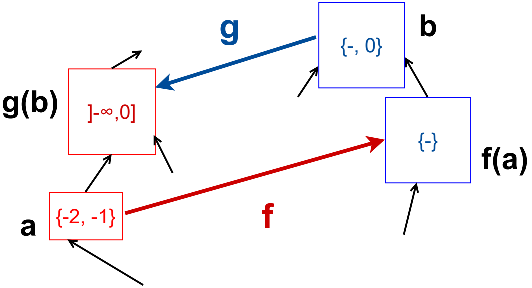
Landmarks: Galois connection *in* Lattice theory (1940-1944)


Georges David Birkhoff, 1940; Øystein Ore, 1944

A monotone Galois connection between $(A, \leq)$ and $(B, \leq)$ is a pair (f, g) s.t. f, g monotone and $f : A \to B$ and $g : B \to A$ $\forall a \in A, b \in B$, $a \leq g(b) \Leftrightarrow f(a) \leq b$

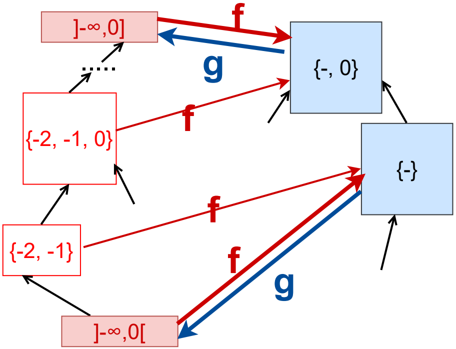
$(A, \leq)$
Sets of integers ordered by inclusion



$(B, \leq)$
Set of signs ordered by inclusion

Landmarks: Galois connection *in* Lattice theory (1940-1944)

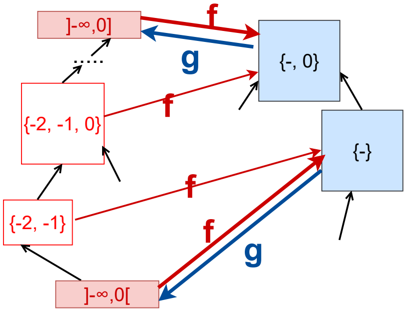
(f, g) Galois connection *implies* $f \circ g$ and $g \circ f$ are closure operators
e.g. for $f \circ g$: isotone ($X \leq Y \Rightarrow f \circ g(X) \leq f \circ g(Y)$), extensive ($X \leq f \circ g(X)$), and idempotent ($f \circ g(f \circ g(X)) = f \circ g(X)$)



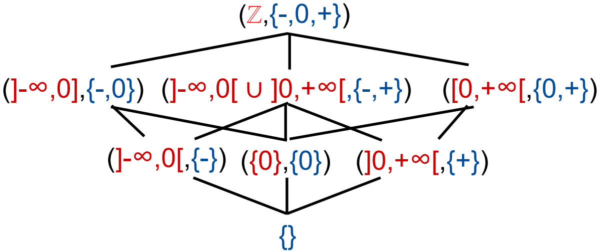
Closed elements for $g \circ f$
$g \circ f(x) = x$
Ex. $]-\infty, 0]$
CtEx. $\{-2, -1, 0\}$
$g \circ f(\{-2, -1, 0\}) =]-\infty, 0]$

Closed elements for $f \circ g$
$f \circ g(x) = x$
Ex. $\{-, 0\}$

Landmarks: Galois lattices *in* Lattice theory (1940-1944)



A few corresponding closed elements



The Galois lattice is the ordered set of pairs/assemblies of corresponding closed elements

Landmarks: Galois lattices *in* Lattice theory (1940-1944)

Particular case: Galois connection associated with a binary relation

O objects, A attributes, $R \subseteq O \times A$

	flies	nocturnal	feathered	migratory	red-bill	elytra	sea-habitat	wood-habitat	six-legged	eats-fish	water-habitat
ladybird	×					×			×		
bat	×	×									
ostrich			×								
greater-flamingo	×		×	×			×			×	×
silver-gull	×		×		×		×			×	×
little-tern	×		×	×			×			×	×
great-auk	×		×				×			×	×
wood-pecker	×		×					×			
giant-otter										×	×
arctic-tern	×		×	×	×		×			×	×

Landmarks: Galois lattices *in* Lattice theory (1940-1944)

f associates an object set with their shared attributes

$$f : P(O) \rightarrow P(A) \qquad X \longmapsto f(X) = \{y \in A \mid \forall x \in X, (x, y) \in R\}$$

	flies	nocturnal	feathered	migratory	red-bill	elytra	sea-habitat	wood-habitat	six-legged	eats-fish	water-habitat
ladybird	×					×			×		
bat	×	×									
ostrich			×								
greater-flamingo	×		×	×			×			×	×
silver-gull	×		×		×		×			×	×
little-tern	×		×	×			×			×	×
great-auk	×		×				×			×	×
wood-pecker	×		×					×			
giant-otter										×	×
arctic-tern	×		×	×	×		×			×	×

Landmarks: Galois lattices *in* Lattice theory (1940-1944)

g associates an attribute set with the objects sharing them

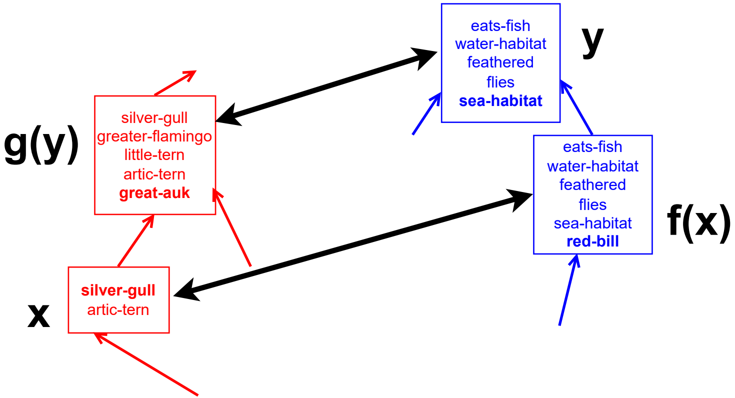
$$g : P(A) \rightarrow P(O) \qquad Y \longmapsto g(Y) = \{x \in O \mid \forall y \in Y, (x, y) \in R\}$$

	flies	nocturnal	feathered	migratory	red-bill	elytra	sea-habitat	wood-habitat	six-legged	eats-fish	water-habitat
ladybird	×					×			×		
bat	×	×									
ostrich			×								
greater-flamingo	×		×	×			×			×	×
silver-gull	×		×		×		×			×	×
little-tern	×		×	×			×			×	×
great-auk	×		×				×			×	×
wood-pecker	×		×					×			
giant-otter										×	×
arctic-tern	×		×	×	×		×			×	×

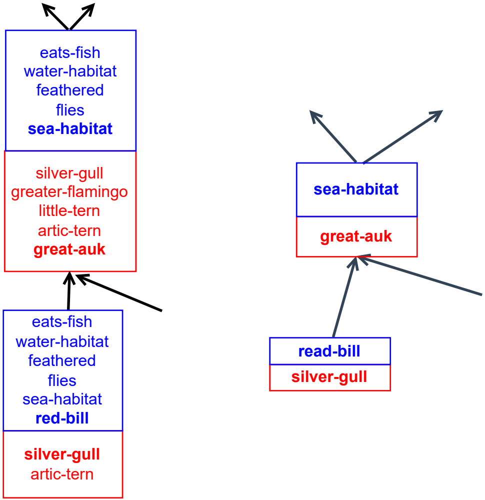
Landmarks: Galois lattices *in* Lattice theory (1940-1944)

(f, g) is a Galois connection between $(2^O, \subseteq)$ and $(2^A, \supseteq)$

Closed sets are maximal sets of objects sharing maximal set of attributes (and reversely)



Landmarks: Galois lattices *in* Lattice theory (1940-1944)

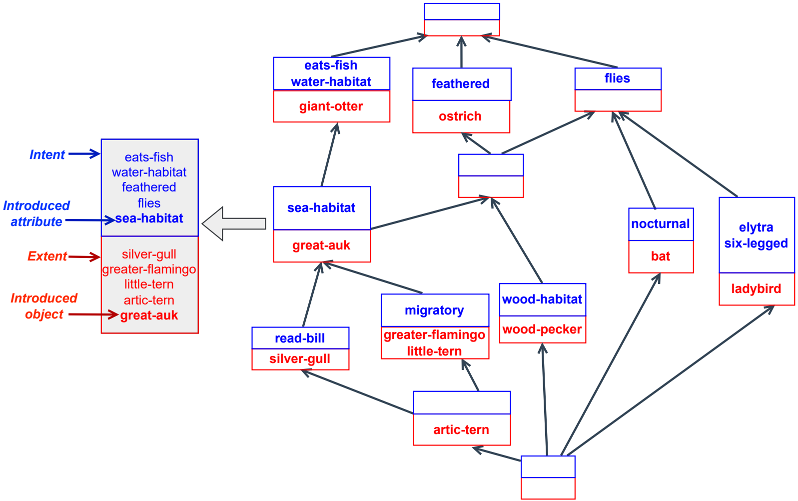


Detail of the Galois lattice of animals

Elements are assemblies of closed sets of objects and attributes

Simplified view (detail): Top-down inherited attributes, bottom-up inherited objects are removed

Landmarks: Galois lattices *in* Lattice theory (1940-1944)



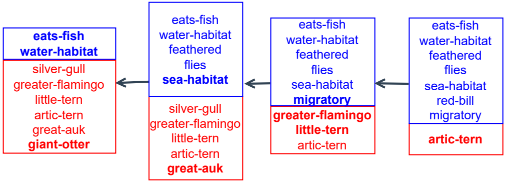
Landmarks: Galois lattices / connection *in* math. & social sciences (1970)



Marc Barbut & Bernard Monjardet, 1970

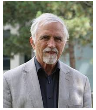
Qualitative multivariate analysis of questionnaires

Guttman scale
(chain in the lattice):
a single dimension
re-arranging
subjects and questions

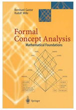


	eats-fish	water-habitat	feathered	flies	sea-habitat	migratory	red-bill
giant-otter	×	×					
great-auk	×	×	×	×	×		
greater-flamingo	×	×	×	×	×	×	
little-tern	×	×	×	×	×	×	
arctic-tern	×	×	×	×	×	×	×

Landmarks: FCA *from* philosophy *to* math. (1982-1999)







Rudolf Wille, 1982, Bernhard Ganter & Rudolf Wille, 1999

Formal Concept Analysis

The decisive turning point

Gathering all lines of thinking, popularizing the approach

Developing many theoretical tools for applications, including the logical perspective

(O, A, R)	Formal context
$f \quad g \quad f \circ g \quad g \circ f$	$' \quad ' \quad '' \quad ''$
Closed set of objects	Extension, extent
Closed set of attributes	Intension, intent
Assembly of corresponding closed sets	Concept
Partial order on assembled closed sets	Partial order subconcept / superconcept
Galois lattice	Concept lattice

Applications (A few domains)

- Environment, biology, chemistry, health
- Linguistics, Text understanding
- Software engineering
- Communities, social network

Part of them in: Jonas Poelmans, Dmitry I. Ignatov, Sergei O. Kuznetsov, Guido Dedene: Formal concept analysis in knowledge processing: A survey on applications. Expert Syst. Appl. 40(16): 6538-6560 (2013) + significant new work since

Tools

More than 50 web applications, downloadable software or plugins at Uta Priss webpage `https://upriss.github.io/fca/fcasoftware.html`

- Algorithms: Conexp family, ToscanaJ, fcaR, GALACTIC
- Extensions: RCA software (Galicia, RCAexplore, FCA4J), Fuzzy (fcaR), polyadic (FCA Tools Bundle)
- Search/Query engines and IR: Credo family, Search Sleuth family, Camelis, Sparklis family
- Visualization+navigation: Latviz, RV-xplorer, ConceptCloud, RCAviz

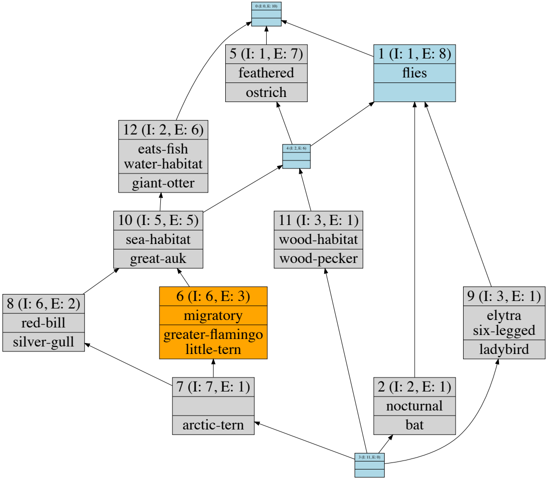
Workshops on tools @ICFCA or @CLA

Table of Contents

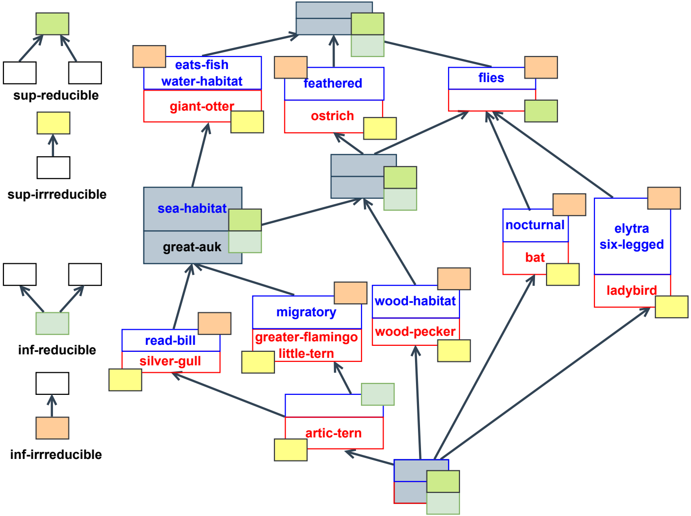
Unveiling conceptual hierarchical structure

Extent/intent view (duality)

- Ontology, conceptual model
- Visualization, conceptual navigation
- Querying (classified answers)
- Recommendation



Unveiling essential elements (sup/meet irreducible - data structure skeleton)



Unveiling logic structure (intensional view)

"What can we remember and gather from the data forgetting the objects" V. Duquenne, 1987

Intensional view

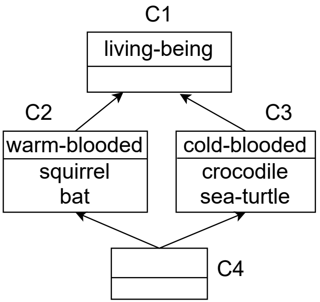
- Implication rules
- Association rules
- Other logical constraints

Data mining

six-legged => flies, elytra
wood-habitat => flies, feathered
elytra => flies, six-legged
nocturnal => flies
red-bill => flies, feathered, sea-habitat, eats-fish, water-habitat
migratory => flies, feathered, sea-habitat, eats-fish, water-habitat
sea-habitat => flies, feathered, eats-fish, water-habitat
feathered, eats-fish, water-habitat => flies, sea-habitat
flies, eats-fish, water-habitat => feathered, sea-habitat
water-habitat => eats-fish
eats-fish => water-habitat

Unveiling logic structure: implications, co-occurrences, mutex, or, xor

An example with the complex constraint Xor



Extents of C2, C3 form a partition of C1 Extent

$Extent(C1) = Extent(C2) \cup Extent(C3)$ and $Extent(C2) \cap Extent(C3) = \emptyset$

$$\boxed{living\text{-}being \implies warm\text{-}blooded \oplus cold\text{-}blooded}$$

Universality

Universality of the descriptions

- Any Galois connection between finite lattices induces a Galois lattice (the theory extends thus to complex descriptions set, far beyond tabular data)

Universality of the structure

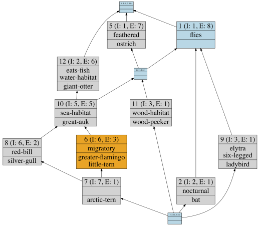
- Any lattice can be labelled by objects and attributes to give rise to a concept lattice

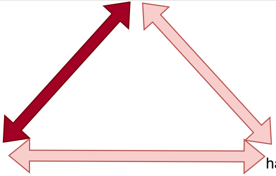
Universality of the underlying logic formula

- Any formula in propositional logic can give rise to a concept lattice (adding an interpretation)

Universality / The magic triangle

	flies	nocturnal	feathered	migratory	red-bill	elytra	sea-habitat	wood-habitat	six-legged	eats-fish	water-habitat
ladybird	×					×			×		
bat	×	×									
ostrich			×								
greater-flamingo	×		×	×			×			×	×
silver-gull	×		×		×		×			×	×
little-tern	×		×	×			×			×	×
great-auk	×		×				×			×	×
wood-pecker	×		×					×			
giant-otter										×	×
arctic-tern	×		×	×	×		×			×	×





<0> flies,feathered,sea-habitat,wood-habitat,eats-fish,water-habitat => nocturnal,migratory,red-bill,elytra,six-legged
(...)
<0> flies,nocturnal,elytra,six-legged => feathered,migratory,red-bill,sea-habitat,wood-habitat,eats-fish,water-habitat
(...)
<1> nocturnal => flies\\
(...)
<5> flies,eats-fish,water-habitat => feathered,sea-habitat\\(...)

Table of Contents

Embedding complex data and data models

My data are not tabular!

Embedding complex data and data models

Remedies

- Multi-valued attributes : integers, double, terms, structures, symbolic objects, etc. (Ganter et Wille, Diday, Polaillon, ...)
- Value taxonomies (Godin et al., Carpineto et Romano, ...)
- Logical description (Chaudron et al., Ferré et al., ...)
- Graphs (Ganter and Kuznetsov, Liquière, Prediger et Wille, Kötters et al., Graph-FCA Ferré et al....)
- Multi-relational, RCA (Priss, Rouane et al., ...);
- Polyadic (Sacarea, Tronca et al.)
- Sequences (Boukhetta, Demko, Bertet et al., Buzmakov et al.)
- Temporal data (Wolff et al., Nica, Braud, Dolques, Le Ber et al., Boukhetta, Demko, Bertet et al.)
- Pattern Structures (Ganter et al., Kuznetsov, Napoli, Buzmakov et al.)

Embedding complex data and data models: Pattern Structures

Pattern structures, Ganter and Kuznetsov 2001

In the following framework:

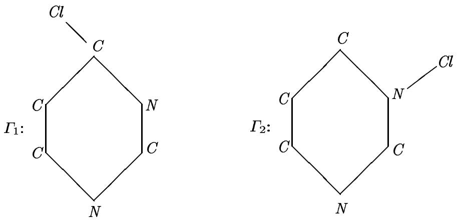
- a set of objects G
- a semi-lattice of descriptions $(D, \sqcap)$, $\sqcap$ similarity/intersection operator
- a partial order on D: $a \sqsubseteq b$ iff $a \sqcap b = a$ (a is subsumed by b)
- a map which associates an object to its description $\delta : G \rightarrow D$

A Galois connection (f, g) can be defined:

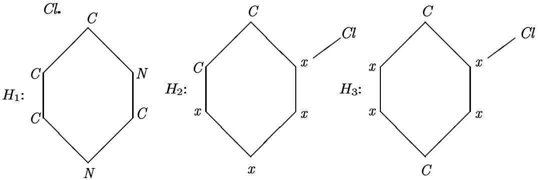
- $\forall A \subseteq G, f(A) = \sqcap_{g \in A} \delta(g)$
- $\forall d \in D, g(d) = \{g \in G | d \sqsubseteq \delta(g)\}$

Embedding complex data and data models: Graph Pattern Structures [Ganter&Kuznetsov, 2001]

G1

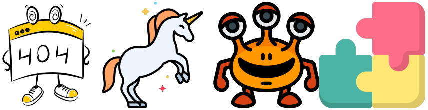


G2



Similarity ('intersection') operator: $G1 \sqcap G2$

Incomplete, imprecise, noisy data



My data are not perfect!

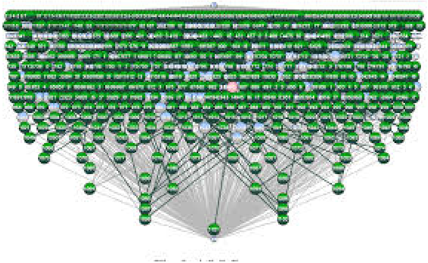
Incomplete, imprecise, noisy data

Remedies

- Fuzzy frameworks (Belohlavek et al., Cabrera, Cordero, Enciso, Mora, Lòpez-Ròdriguez, Ojeda-Aciego et al., Cornejo, Medina et al., Yahia et al., Dubois, Prade, Boffa et al for Fuzzy RCA)
- Attribute exploration (Ganter, Wille, Rudolph, Obiedkov)
  Successive presentation of implications to experts + gathering answer Yes/no and counterexamples that complete/consolidate the dataset

- Is it true that red-bill => flies, feathered, sea-habitat, eats-fish, water-habitat?
- No! Quelea verifies: red-bill, feathered, flies, savanna-habitat, eats-cereal

Scalability: result size



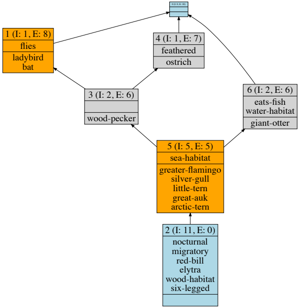
The lattice can be huge!
#concepts $< 2^{min(|A|,|O|)}$

The Rijksmuseum collection

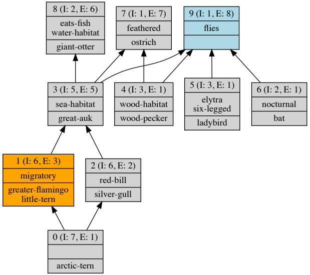
100,000 objects; 1,716 attributes; 994,967 concepts; computed with FCbO update algorithm [Wray et al., 2016]

Scalability: result size

Remedies: clarify data, restrict to irreducible elements, to frequent concepts, to introducer concepts, partial construction with incremental algorithms



Iceberg: Restriction to concepts with frequent intent or frequent extent



AOC-poset: Restriction to introducer concepts

Scalability: result size

There are many possible logical constraints!

Scalability: result size

Remedy 1: compute implication bases with particular properties

Sound, complete, non redundant implication sets

- Basis of Duquennes-Guigues [Duquennes&Guigues, 1986]
  For a systematic study, see [Bertet&Monjardet, 2010]
- Left-minimal direct basis of implications [Cordero 2013]
- Basis of proper premises [Reppe, 2008; Ryssel, 2014]
- and others ...

Remedy 2: eliminate accidental constraints

Using additional information (e.g. ontologies)

Scalability: result complexity

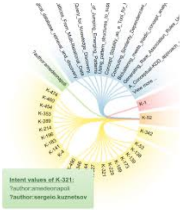
"what kind of alien really reads the Figure 3 concept lattice ?"

an anonymous reviewer

Scalability: result complexity

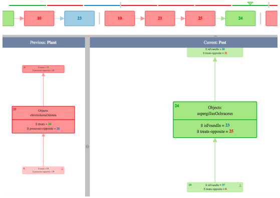
Remedy 1: Rephrasing results in terms of the domain experts

Remedy 2: Visualization and interactive exploration tools



Circular view
Alam et al. 2015

Tag clouds
Greene et al. 2015



Local views
Muller et al. 2022

Table of Contents

FCA in Software Engineering at a glance

- Mostly well formatted, complete, complex data: conceptual models, specifications, source code, traces, call graphs, git actions ...
- Software engineers are proficient to understand and exploit the results
- A large range of applications (non exhaustive references hereafter; and other references included in the SPLE section)
- *Most of the approaches convert data into tabular form*
- There is space to update the survey *“Thomas Tilley, Richard Cole, Peter Becker, Peter W. Eklund: A Survey of Formal Concept Analysis Support for Software Engineering Activities. Formal Concept Analysis 2005: 250-27”*!

FCA in Software Engineering at a glance

Galois connection

- Abstract interpretation [Cousot & Cousot, POPL 1979]

Patterns

- Analyzing method call schemes in OO languages [Arevalo et al. ASE, 2003]

Clustering

- Migration from procedural to OO paradigm [Sahraoui et al. ASE, 1997]
- Modularizing [Lindig & Snelting, ICSE 1997; Siff and Reps, IEEE TSE 1999]

Rules

- Fault Localization [Cellier et al., ICFCA 2008]
- Learning model transformation from examples [Saada et al., MODELS 2012]

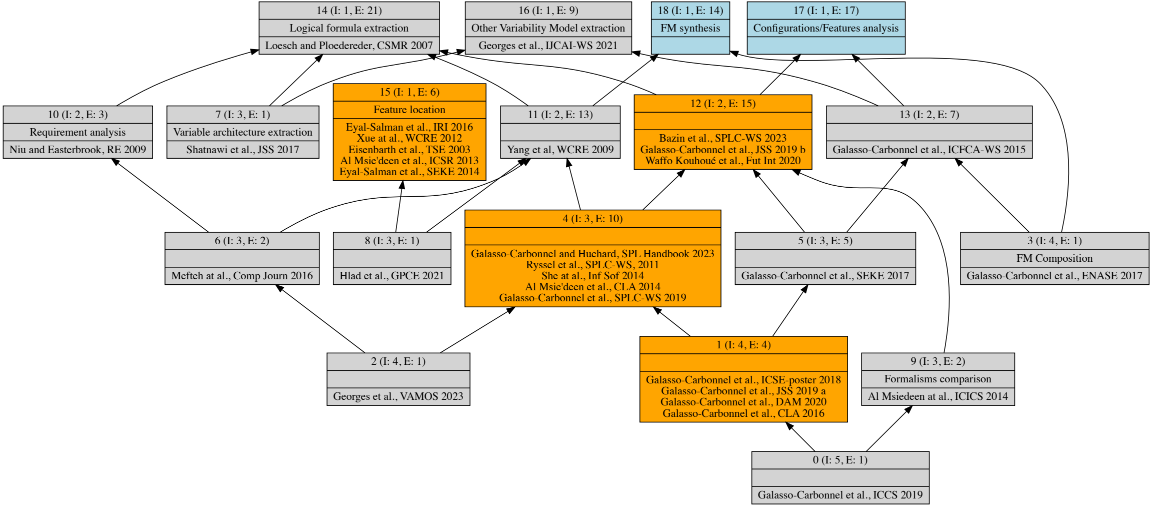
FCA in Software Engineering at a glance

Classification, exploration

- Refactoring of class models / Class hierarchies from:
  - artefact description [Godin & Mili OOPSLA, 1993 (classes); Dao et al. ICCS, 2004 (classes+associations)]
  - artefact usage [Snelting & Tip TOPLAS, 2000]
- Concept location [Koschke, Eisenbarth et al., ASE, 2005, IEEE TSE 2003, Poshyvanyk et al., TOSEM, 2010]
- Libraries of repositories of software artefacts: classes, components, web services [Aboud et al., ECSA 2009 & JOT 2019; Azmeh et al. ICWS, 2011]
- Indexing, retrieval of software artefacts [Park, JSS 2000]
- Exploration of software version control repositories [Greene et al. 2017]

Table of Contents

FCA in Software Product Lines



FCA as a variability framework

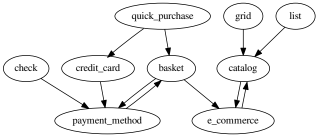
A framework for logical variability expression

- In correspondence with propositional/predicate/description logics background
- Sound and complete extraction capabilities
- Canonical and exhaustive constructions
- Graphical representation and exploration capabilities
- Extensible to complex descriptions

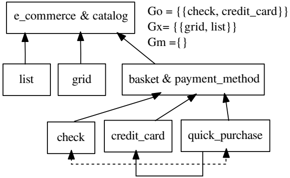
Credits to Jessie Galasso-Carbonnel PhD work

Structuring framework

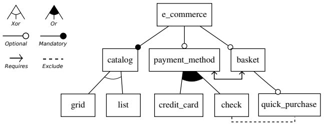
- Positioning relatively to other (intensional) formalisms



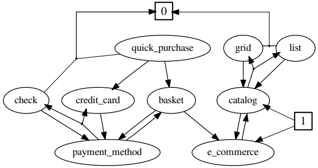
Binary implication graph



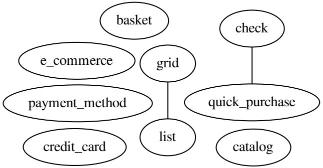
Feature graph



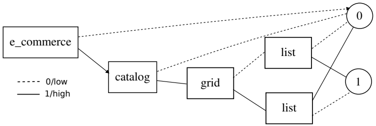
Feature model



Feature hypergraph

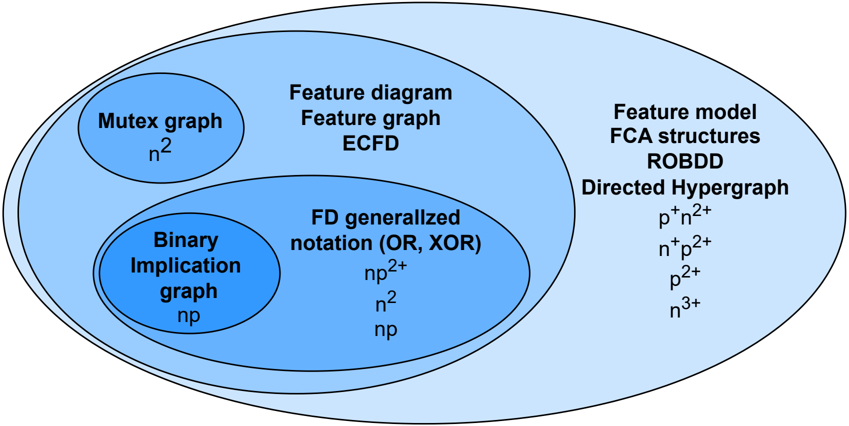


Mutex graph



ROBDD

Structuring framework: one of the sound and complete representations

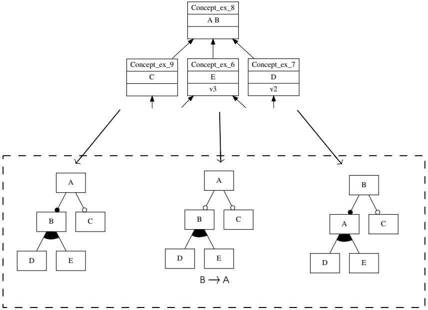


CNF, *n* negative litteral, *p* positive litteral

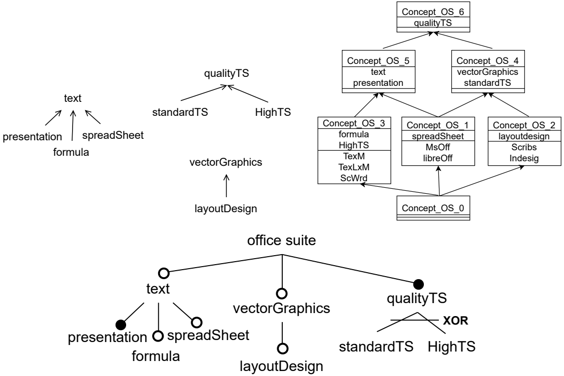
J. Galasso-Carbonnel et al., ICCS 2019

Structuring framework

- Canonical structure: The concept lattice associated with a propositional logic formula is unique (up to isomorphism)
- All FDs with the same logical semantics are embedded in the lattice



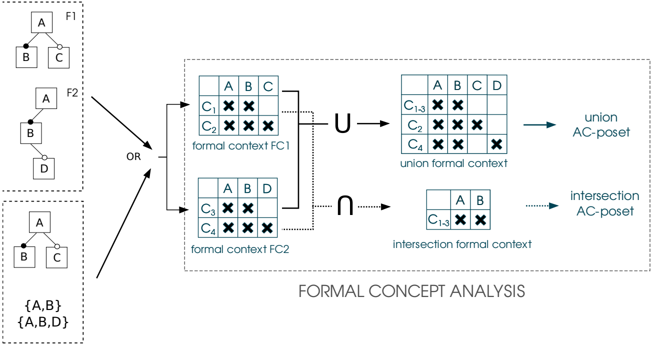
J. Galasso-Carbonnel et al. 2019 (JSS a)



Ontologie+lattice = FM

Composition: union / intersection of variability models

J. Galasso-Carbonnel et al. 2017 (ENASE 2017)



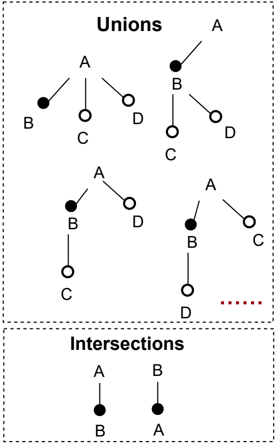
	A	B	C
C_1	✖	✖	
C_2	✖	✖	✖

	A	B	D
C_3	✖	✖	
C_4	✖	✖	✖

	A	B	C	D
$C_{1\text{-}3}$	✖	✖		
C_2	✖	✖	✖	
C_4	✖	✖		✖

	A	B
$C_{1\text{-}3}$	✖	✖

FORMAL CONCEPT ANALYSIS



Product family analysis: summarization with implication bases

E.g. with the Duquenne-Guigues implication basis

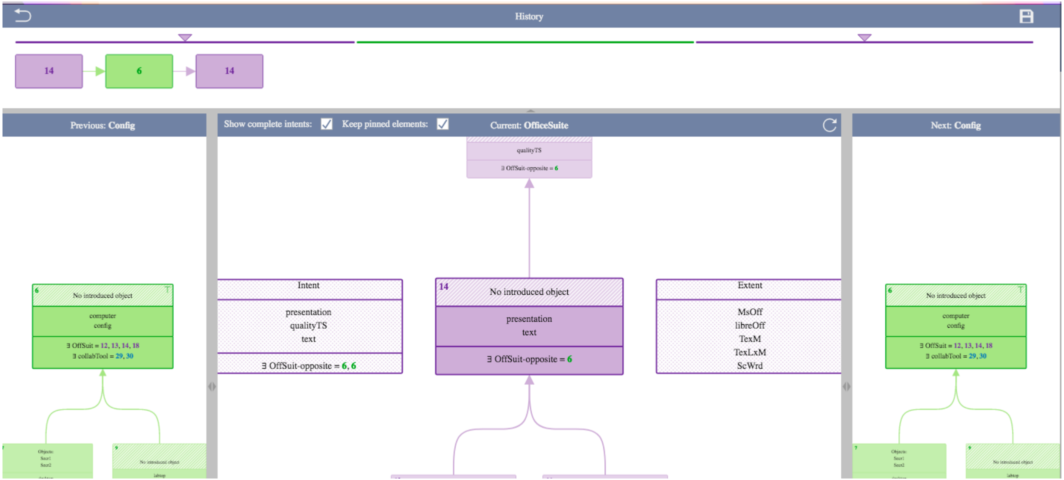
<7> $\Rightarrow$ qualityTS
<3> qualityTS,HighTS $\Rightarrow$ text,formula,presentation
<4> qualityTS,StandardTS $\Rightarrow$ vectorGraphics
<5> presentation,qualityTS $\Rightarrow$ text
<4> vectorGraphics,qualityTS $\Rightarrow$ StandardTS
<2> layoutDesign,qualityTS $\Rightarrow$ vectorGraphics,StandardTS
<3> formula,qualityTS $\Rightarrow$ text,presentation,HighTS
<5> text,qualityTS $\Rightarrow$ presentation
<2> text,vectorGraphics,presentation,qualityTS,StandardTS $\Rightarrow$ spreadSheet
<2> spreadSheet,qualityTS $\Rightarrow$ text,vectorGraphics,presentation,StandardTS
<0> spreadSheet,text,layoutDesign,vectorGraphics,presentation,qualityTS,StandardTS $\Rightarrow$ formula,HighTS
<0> spreadSheet,text,formula,vectorGraphics,presentation,qualityTS,StandardTS,HighTS $\Rightarrow$ layoutDesign

And all is said!

Exploring the product family (navigation, recommendation)

S. Ferré, 2014, *Conceptual Navigation*

A. Bazin et al., ICFCA 2019 *On-demand algorithm* E. Muller et al., CLA 2021 *RCAviz*



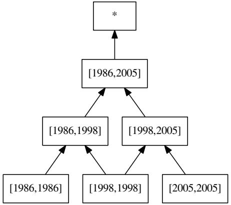
Extending the framework: Complex descriptions

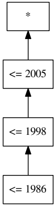
Multi-valued attributes, Cardinalities (for feature or feature group)

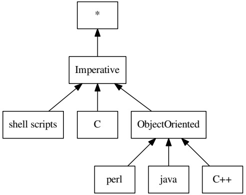
J. Galasso-Carbonnel et al. 2019 (JSS b)

- Leverages the FCA variant of Pattern structures
- Extraction of conceptual structures, implications, co-occurrences and mutex
- Applied to 30 wikipedia PCM on software, Robocode and 3 JHipster excerpts (500, 1000, 2000 products)

Systems	Creator	Last Stable Version	Release Date	License	Implementation language(s)	Server Operating System	Databas backen
Act! CRM	Pat Sullivan and Mike Muhney	22.1	June 2020	Proprietary or SaaS	C#, .NET	Windows, Web, iOS, Android	MS S
Adempiere	Jorg Janke	3.9.3	December 2019	GPLv2	Java	Windows, Linux, Unix, Mac	
Base CRM	Base CRM	2016	2009	SaaS	Ruby on Rails, Python	Cloud Computing, iOS, Android, Windows	MyS





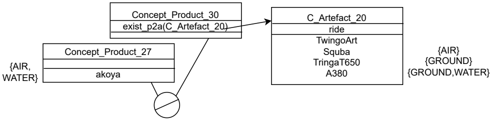


Extending the framework: Complex descriptions

Assist Feature location in Software Products Lines

Hlad et al. GPCE 2021

- Feature location: find the software artefacts (e.g. code) corresponding to a feature ; consider feature interaction
- Associate a group of artefacts with a group of features shared by products
- Infer: variability model, components and code annotations




```
public class Vehicle{
 //#if (AIR || GROUND) && ! ( AIR  &&  WATER )
  public void ride_(){...}
 //#endif
...
}
```

Extending the framework

Consider multiple dimensions with polyadic concept analysis

A few examples:

- Variability in space and time: configurations, features, time/release
  A. Bazin et al., Varivolution - SPLC 2023
- Variability in user stories: configurations, roles, features
  A. Bazin, T. Georges et al., ongoing work

Challenges

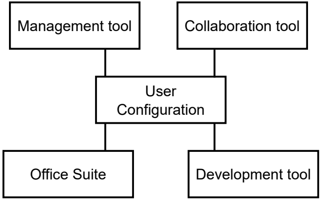
- Representation of conceptual structures
- Rule interpretation requires subtlety

Extending the framework

Consider interconnected product families

J. Galasso-Carbonnel et al., Workshop REVE - SPLC 2019

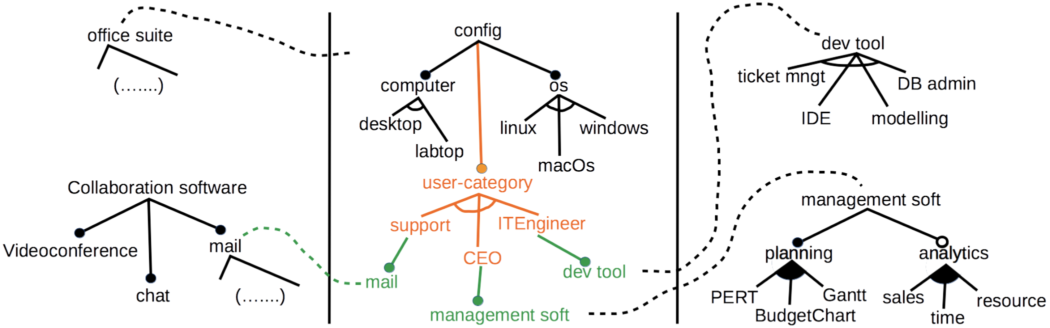
- Set of (Configurations + Features) - one per family
- Relational Concept Analysis (RCA) = Set of lattices + links between lattices



Interconnected product lines

Guiding interconnected FM construction

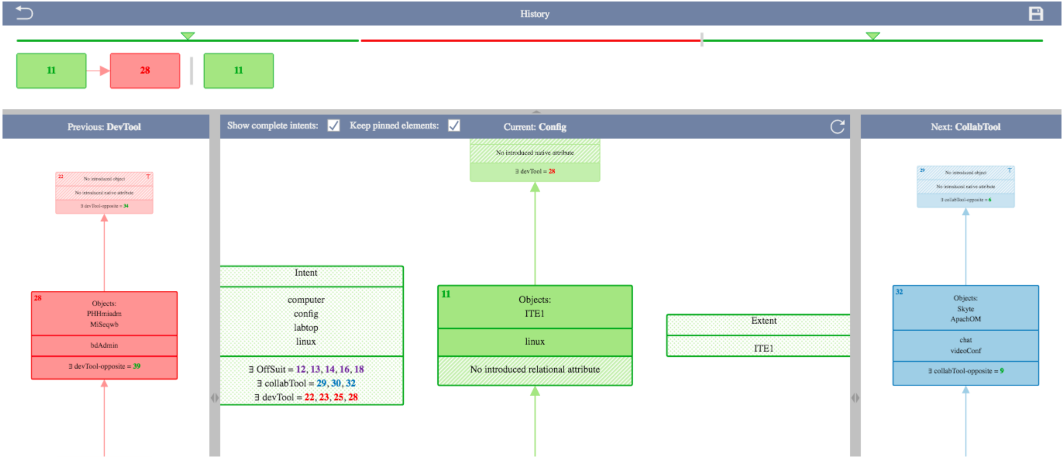
Using the lattices and the lattice interconnexions created by RCA



Interconnected product lines

Exploring the product family through conceptual structures

E. Muller et al., CLA 2022 *RCAviz*



Interconnected product lines

Exploring the product family through implications

L. Musslin et al., ongoing work *RCAvizIR*

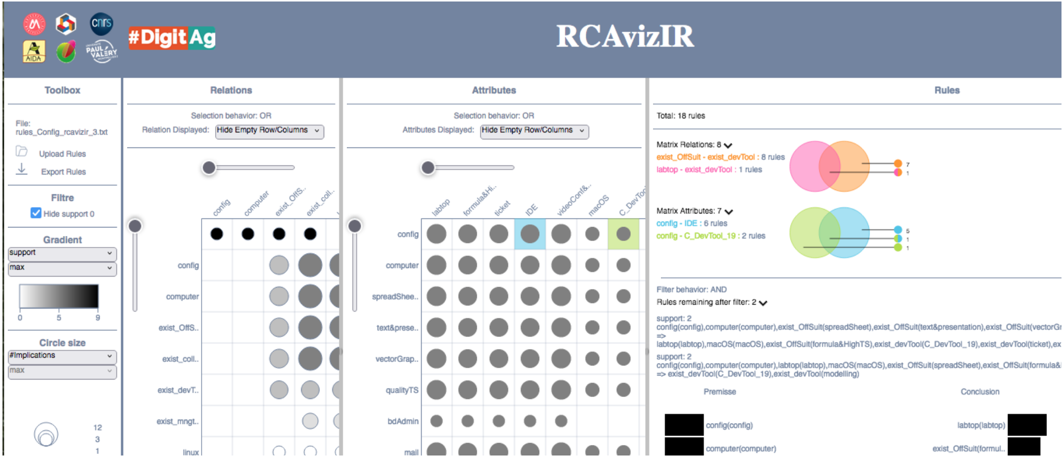
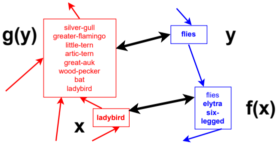
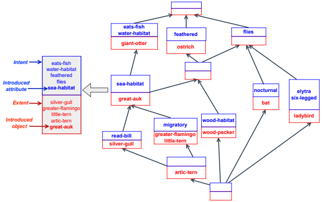


Table of Contents

Synthesis on FCA

Concepts in philosophy + Algebra ⇒ Galois / FCA ; Structures / Rules





six-legged => flies,elytra
red-bill => flies,feathered,sea-habitat,eats-fish,water-habitat
feathered,eats-fish,water-habitat => flies,sea-habitat
eats-fish => water-habitat

A concept is a medium between concrete and abstract levels

FCA / Galois structures in the Artificial Intelligence landscape

"A Human-centered approach for explainable complex data analysis"
J.Hirth, T. Hanika, 2023

- Symbolic (logics based) vs. subsymbolic
- White box vs. black box: explainable
- Closed word assumption
- Supervised and unsupervised
- Structured vs. unstructured data
- Complex and heterogeneous data
- Robustness vs. depending data distribution or algorithms behavior
- Multi-hierarchical, canonical classification vs. single, non canonical

What's next for FCA?

- Theory: e.g. analyzing structures, connecting the different FCA trends, connecting FCA to other KD approaches
- Algorithms: e.g. local/incremental, parallel
- Methodology: e.g. user interaction, detecting and correcting anomalies, classifying rules, generalizing from applications
- Address big data and deluge of patterns challenges
- Using FCA/RCA as data complexity measuring framework
- Hybrid AI/KD systems, Neuro-symbolic AI; Explainable AI
- Integration in data science workflows (e.g. in Orange or Scikit-learn with subsymbolic ML, in Jupyter Notebooks)

Hybrid IA/KD/DS systems

Recurring work connecting FCA and other KR/ML techniques

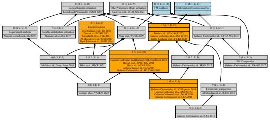
- Decision trees, Bi-clustering
- Description logics, Propositionalization, Ontologies
- Constraint Programming (CSP)
- and others ...

Recent tracks

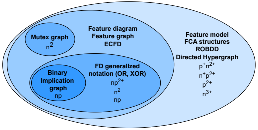
- Decision quivers (E. Dudyrev et al.)
- Alternative classifiers with explanations (“competitive for explanation”, S. Kuznetsov et al.)
- Providing conceptual views; reasons for classification in ML methods; abductive learning of comprehensible rules from neurons (Endres and Foldiak, A. Sangrova et al., A. Tomat, Khatri et al., J. Hirth et al.)

Synthesis on FCA + SPLE

Applications in SPLE, in particular for (complex) variability structuring



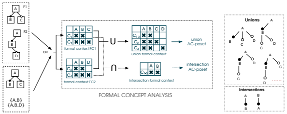
A variety of applications



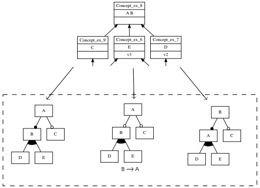
Canonical model

$\Rightarrow$ qualityTS
qualityTS,HighTS $\Rightarrow$ text,formula,presentation
qualityTS,StandardTS $\Rightarrow$ vectorGraphics
presentation,qualityTS $\Rightarrow$ text
vectorGraphics,qualityTS $\Rightarrow$ StandardTS
layoutDesign,qualityTS $\Rightarrow$
vectorGraphics,StandardTS

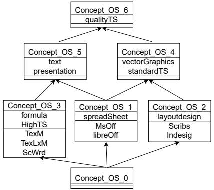
Summarization with implication bases



Variability models combination



Reference model



Combine Intensional/extensional views

What's next for SPLE with FCA?

- Complex variability models
  - Multi-dimensional
  - Multi-product line
  - Predicates / Description logics
- Deepen equivalences and bridges between existing variability models
- Develop FCA-based components dedicated to SPLE
- Methodology
  - Define typical patterns for encoding SPLE problems in FCA
  - Design processes including FCA-dedicated components
  - Combine with other AI approaches (hybrid AI)
    - For result consolidation
    - For explanation
    - As part of processing workflows

Supported by the ANR SmartFCA project. Grant ANR-21-CE23-0023 of the French National Research Agency anr®

Thank you for listening!

This reflexion could not have happened without ...

icons by Flaticon - word cloud by wordclouds.com