Formal Concept Analysis

Welcome and Organizational Issues

Sebastian Rudolph

Computational Logic Group
Technische Universität Dresden

Agenda

Organization

Lectures and Exercises

lecture Thursday, 9:20 – 10:50, room INF E005

hands-on exercise Wednesday, 9:20 – 10:50, room INF E005

Course Web Page

`http://www.inf.tu-dresden.de/?node_id=3482&ln=en`

Contact

Sebastian Rudolph `sebastian.rudolph@tu-dresden.de`, room 2035, scheduled individually ...or at the Open House

Organization

hands-on exercise means

- *autonomous work* on the practice sheet in small teams of 3-4 students, under supervision
- *no general repetition* of lecture material
- *no demonstration* of the sample solution (will be provided later)

necessary for that is

- making notes during the lecture
- performing autonomous follow-up course work *before* the exercise
- bringing material and your notes to the exercise
- developing own activity

Organization

Why this exercise concept?

- active development of the lecture material is more effective
- discovering relationships in the material
- learning structured thinking and autonomous working
- learning team work
- learning to explain things
- exercise for the exams ;-)
- *You have finished your study of ... Your personal strengths include pro-activity and team work, you are communicative and willing to cooperate.* (typical job advertisement)

Agenda

Overview

Formal Concept Analysis

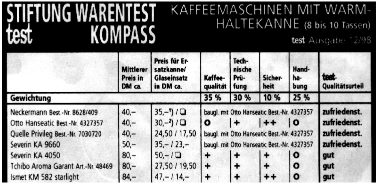
Applications

Sebastian Rudolph

Computational Logic Group
Technische Universität Dresden

History of Formal Concept Analysis

Formal Concept Analysis (FCA) originated in Darmstadt, Germany around 1980 as a mathematical theory that delivers a formalization of the concept of a "concept".



STIFTUNG WARENTEST test KOMPASS

KAFFEEMASCHINEN MIT WARMHALTEKANNE (8 bis 10 Tassen)

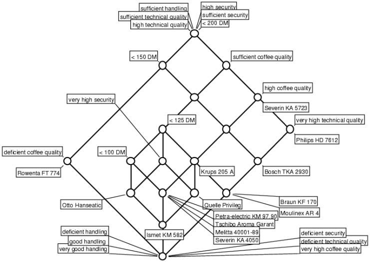
test Ausgabe 12/98

	Mittlerer Preis in DM ca.	Preis für Ersatzkanne/ Glaseinsatz in DM ca.	Kaffeequalität	Technische Prüfung	Sicherheit	Handhabung	test-Qualitätsurteil
Gewichtung			**35 %**	**30 %**	**10 %**	**25 %**	
Neckermann Best.-Nr. 8628/409	40,–	35,–[1]) / ❑	baugl. mit Otto Hanseatic Best.-Nr. 4327357				**zufriedenst.**
Otto Hanseatic Best.-Nr. 4327357	40,–	30,–[2]) / ❑	o	+	++	o	**zufriedenst.**
Quelle Privileg Best.-Nr. 7030720	40,–	24,50 / 17,50	baugl. mit Otto Hanseatic Best.-Nr. 4327357				**zufriedenst.**
Severin KA 9660	50,–	35,– / 23,–	baugl. mit Otto Hanseatic Best.-Nr. 4327357				**zufriedenst.**
Severin KA 4050	80,–	50,– / ❑	+	+	+	o	**gut**
Tchibo Aroma Garant Art.-Nr. 48469	80,–	27,50 / 19,50	+	+	+	o	**gut**
Ismet KM 582 starlight	84,–	47,– / 14,–	+	+	++	o	**gut**

Since then, FCA spread into different areas of computer science, e.g.,

- data analysis
- knowledge discovery
- software engineering

Starting from datasets, FCA derives concept hierarchies. FCA allows you to create and visualize concept hierarchies based on well-grounded mathematical foundations.

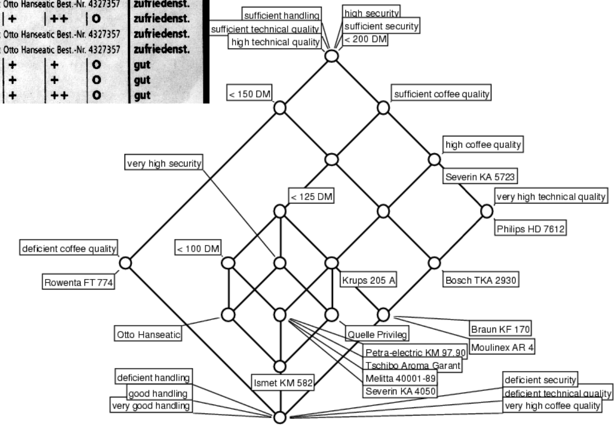


Example: Comparison of Coffee Machines

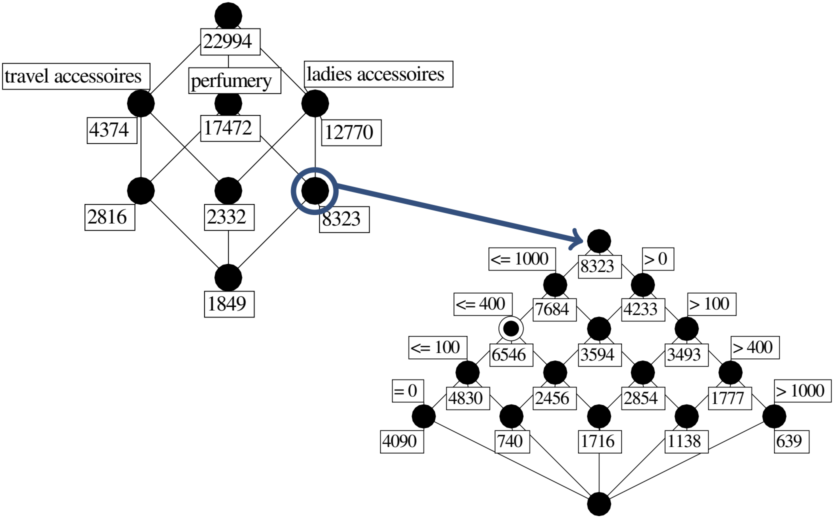
STIFTUNG WARENTEST test KOMPASS — KAFFEEMASCHINEN MIT WARMHALTEKANNE (8 bis 10 Tassen)

test Ausgabe 12/98

	Mittlerer Preis in DM ca.	Preis für Ersatzkanne/ Glaseinsatz in DM ca.	Kaffeequalität	Technische Prüfung	Sicherheit	Handhabung	test-Qualitätsurteil
Gewichtung			**35 %**	**30 %**	**10 %**	**25 %**	
Neckermann Best.-Nr. 8628/409	40,–	35,–[1]) / ❑	baugl. mit Otto Hanseatic Best.-Nr. 4327357				**zufriedenst.**
Otto Hanseatic Best.-Nr. 4327357	40,–	30,–[2]) / ❑	**O**	**+**	**++**	**O**	**zufriedenst.**
Quelle Privileg Best.-Nr. 7030720	40,–	24,50 / 17,50	baugl. mit Otto Hanseatic Best.-Nr. 4327357				**zufriedenst.**
Severin KA 9660	50,–	35,– / 23,–	baugl. mit Otto Hanseatic Best.-Nr. 4327357				**zufriedenst.**
Severin KA 4050	80,–	50,– / ❑	**+**	**+**	**+**	**O**	**gut**
Tchibo Aroma Garant Art.-Nr. 48469	80,–	27,50 / 19,50	**+**	**+**	**+**	**O**	**gut**
Ismet KM 582 starlight	84,–	47,– / 14,–	**+**	**+**	**++**	**O**	**gut**

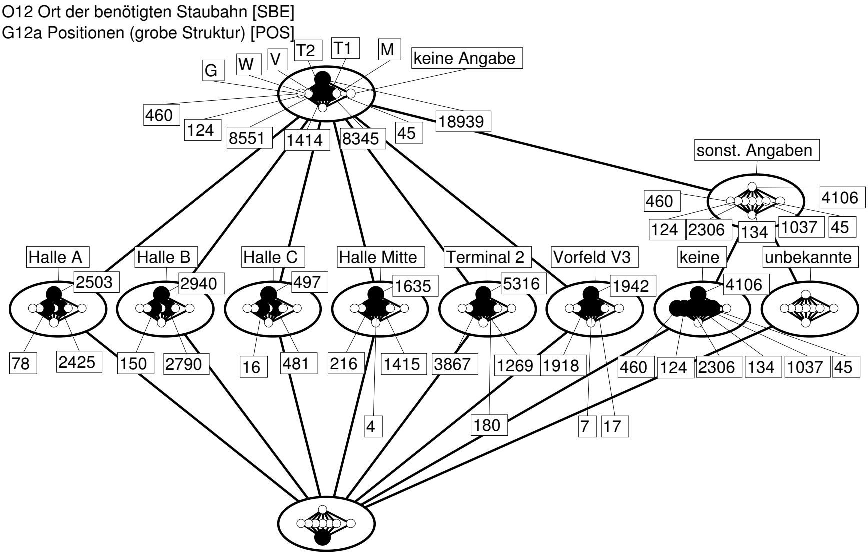


Database Marketing at Jelmoli AG, Zurich[1]



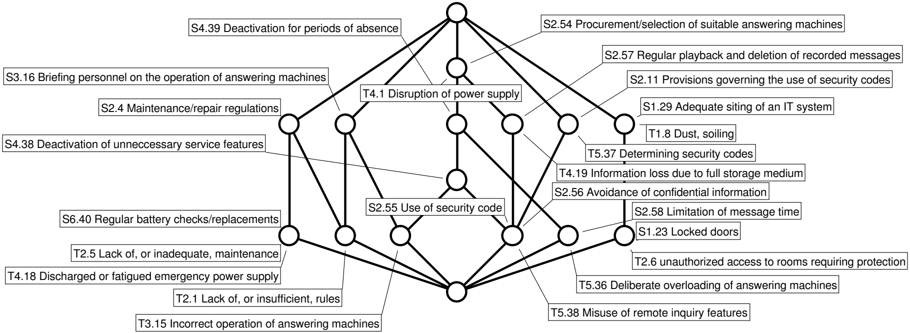
[1] J. Hereth, G. Stumme, R. Wille, and U. Wille. Conceptual knowledge discovery and data analysis. In *Proc. ICCS 2000*, volume 1867 of *LNCS*, pages 421–437. Springer, Berlin/Heidelberg, 2000.

Analysis of Plane Movements at Frankfurt Airport[2]



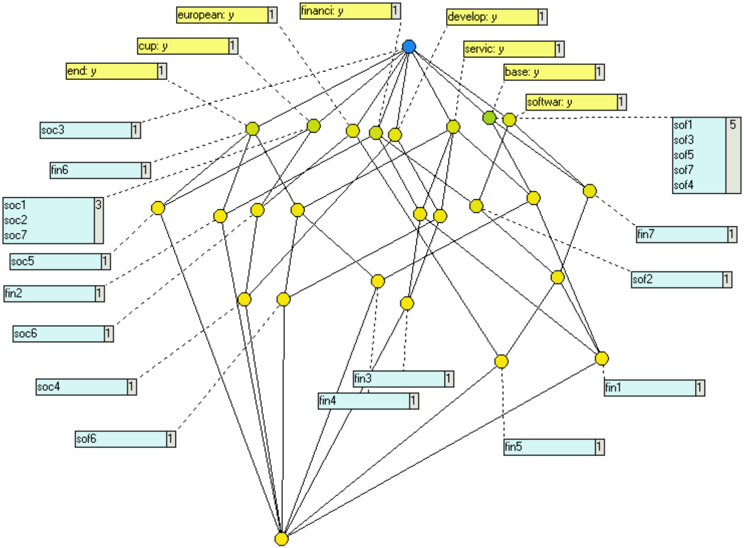
[2] G. Stumme, R. Wille, and U. Wille. Conceptual knowledge discovery in databases using formal concept analysis methods. In *Proc. PKDD*, volume 1510 of *LNCS*, pages 450–458. Springer, Berlin/Heidelberg, 1998.

IT Security Management[3]



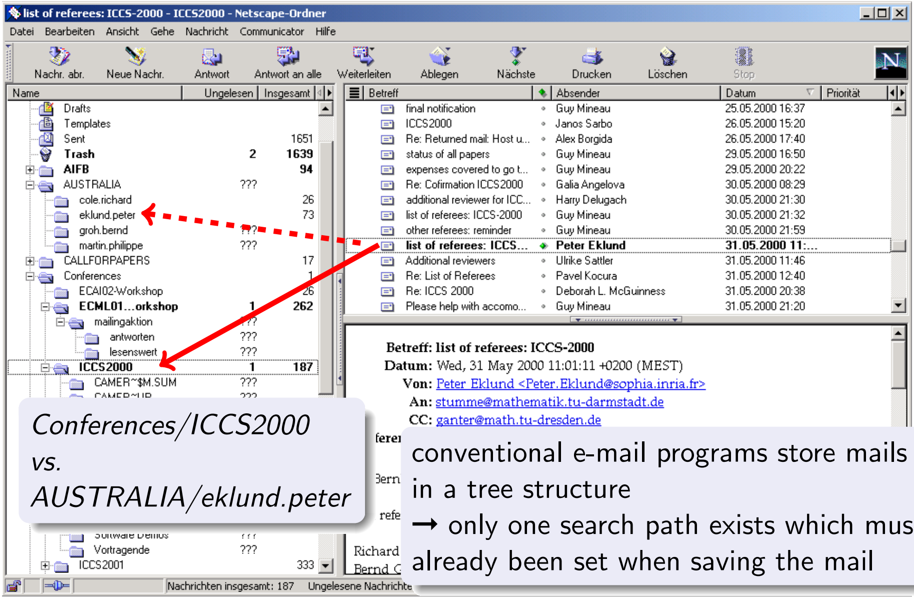
[3] K. Becker, G. Stumme, R. Wille, U. Wille, and M. Zickwolff. Conceptual information systems discussed through an it-security tool. In Rose Dieng and Olivier Corby, editors, *Knowledge Engineering and Knowledge Management Methods, Models, and Tools*, volume 1937 of *LNCS*, pages 352–365. Springer, Berlin/Heidelberg, 2000.

Text Clustering[4]



[4] A. Hotho, S. Staab, G. Stumme. Ontologies improve text document clustering. In *Proc. ICDM*, pages 541–544, 2003.

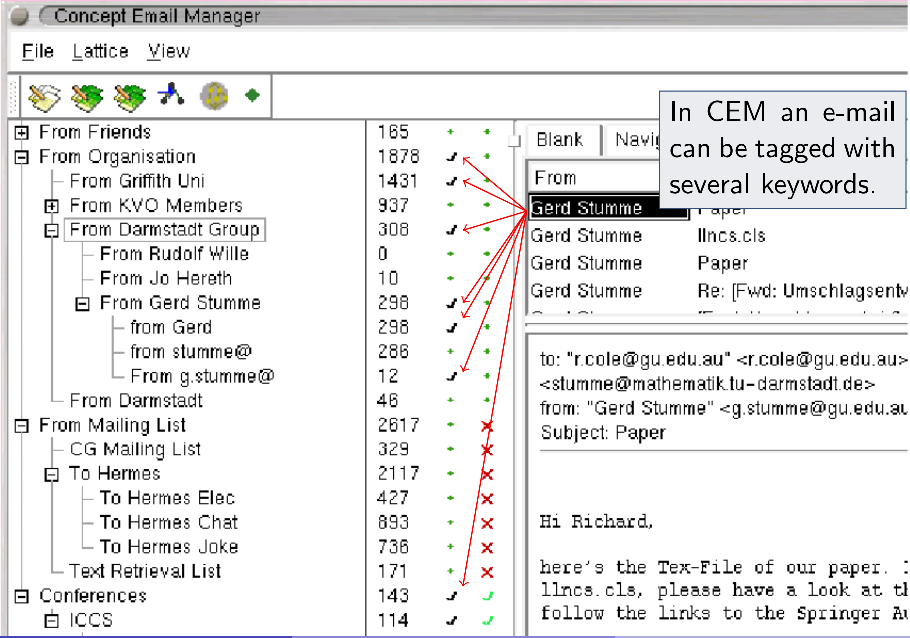
Concept Email Manager



Conferences/ICCS2000
vs.
AUSTRALIA/eklund.peter

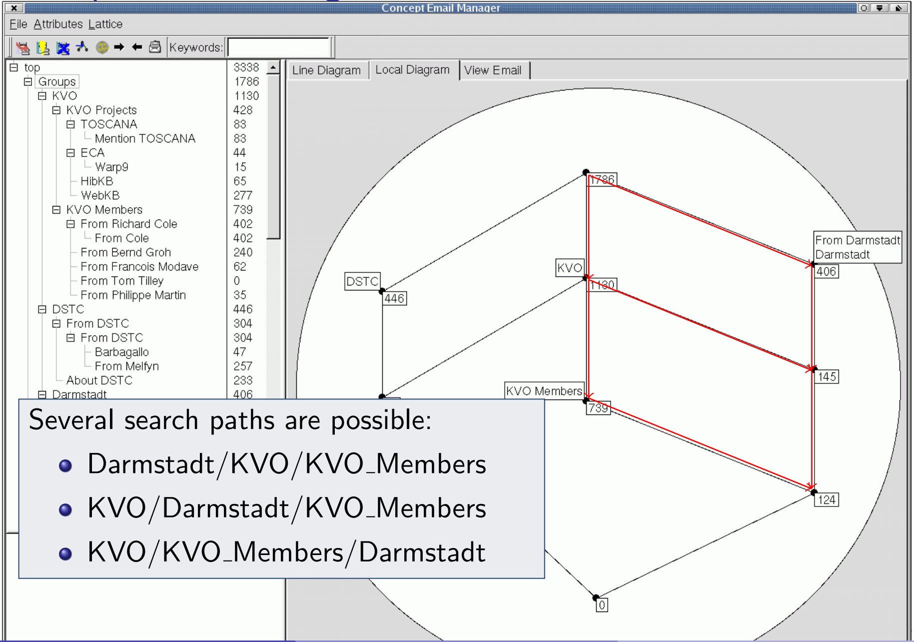
conventional e-mail programs store mails in a tree structure
→ only one search path exists which must already been set when saving the mail

Concept Email Manager



In CEM an e-mail can be tagged with several keywords.

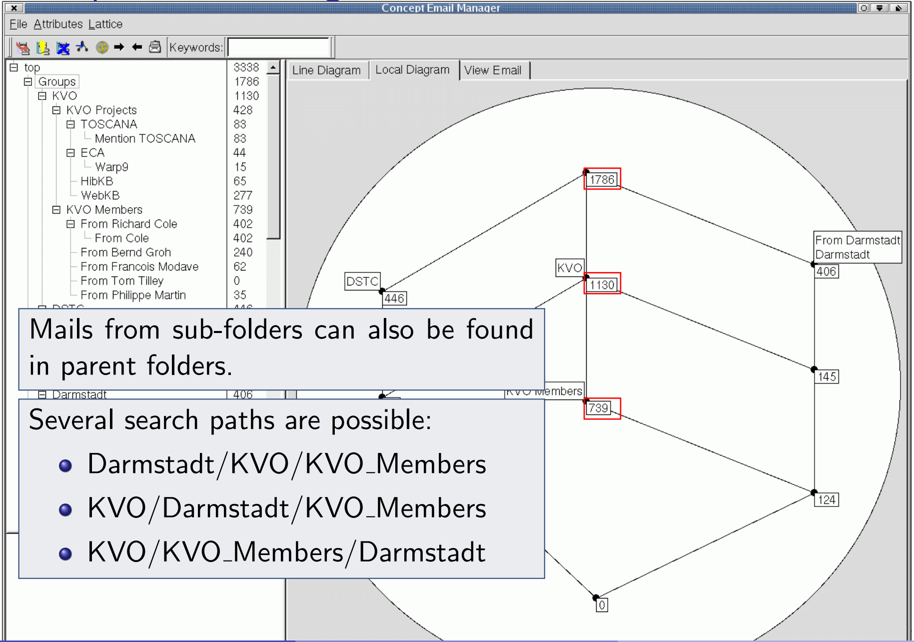
Concept Email Manager



Several search paths are possible:

- Darmstadt/KVO/KVO_Members
- KVO/Darmstadt/KVO_Members
- KVO/KVO_Members/Darmstadt

Concept Email Manager

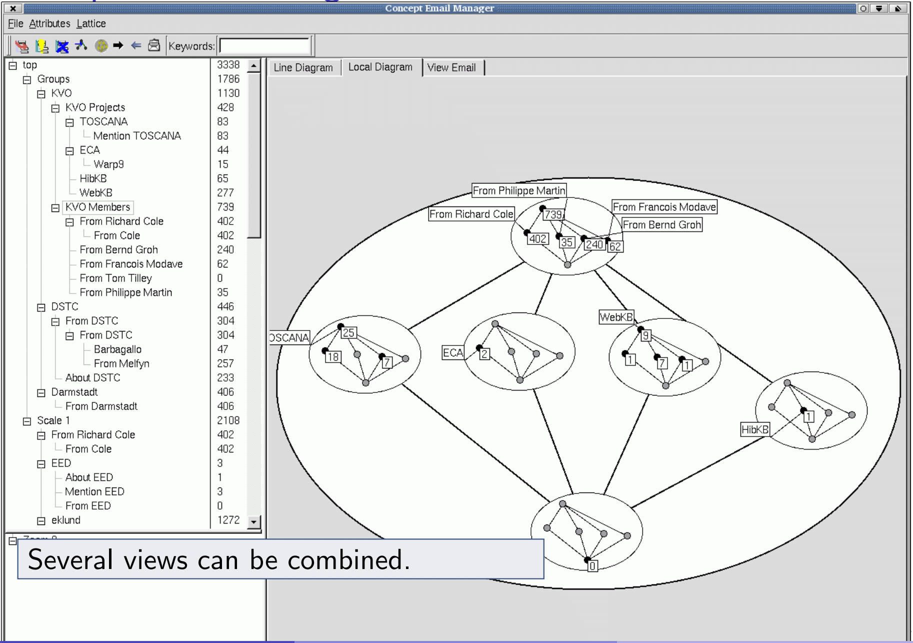


Mails from sub-folders can also be found in parent folders.

Several search paths are possible:

- Darmstadt/KVO/KVO_Members
- KVO/Darmstadt/KVO_Members
- KVO/KVO_Members/Darmstadt

Concept Email Manager



Several views can be combined.

Formal Concept Analysis

I Contexts, Concepts, and Concept Lattices

Sebastian Rudolph

Computational Logic Group
Technische Universität Dresden

slides based on a lecture by Prof. Gerd Stumme

Agenda

1. Concept Lattices
   - What is a concept?
   - Formal Context
   - Derivation Operators
   - Formal Concept
   - Concept Lattice
   - Computing All Concepts
   - Drawing Concept Lattices
   - Clarifying and Reducing a Formal Context
   - Interlude: ConExp
   - Additive Line Diagrams
   - Nested Line Diagrams

What is a concept?

Formal Concept Analysis models concepts as units of thought that consist of two parts:

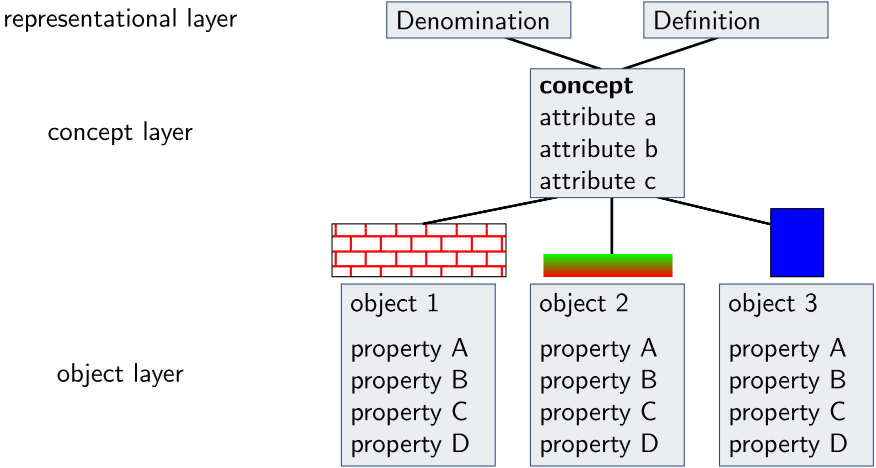
- The *concept extent* comprises all objects that belong to the concept.
- The *concept intent* contains all attributes that all of the objects have in common.

FCA is used, amongst others, data analysis, information retrieval, data mining and software engineering.

What is a concept?

DIN 2330/ISO 704: Concepts and their Denomination

FCA is working on the conceptual layer. The representational layer plays only a minor role.



Formal Context

Def.: A *formal context* is a triple (G, M, I), where

- G is a set of objects,
- M is a set of attributes, and
- I is a relation between G and M.

We read $(g, m) \in I$ as "object g has attribute m".

National Parks in California	NPS Guided Tours	Hiking	Horseback Riding	Swimming	Boating	Fishing	Bicycle Trail	Cross Country Trail
Cabrillo Natl. Mon.						×	×	
Channel Islands Natl. Park		×		×		×		
Death Valley Natl. Mon.	×	×	×	×			×	
Devils Postpile Natl. Mon.	×	×	×	×		×		
Fort Point Natl. Historic Site	×					×		
Golden Gate Natl. Recreation Area	×	×	×	×		×	×	
John Muir Natl. Historic Site	×							
Joshua Tree Natl. Mon.	×	×	×					
Kings Canyon Natl. Park	×	×	×			×		×
Lassen Volcanic Natl. Park	×	×	×	×	×	×		×
Lava Beds Natl. Mon.	×	×						
Muir Woods Natl. Mon.		×						
Pinnacles Natl. Mon.		×						
Point Reyes Natl. Seashore	×	×	×	×		×	×	
Redwood Natl. Park	×	×	×	×		×		
Santa Monica Mts. Natl. Recr. Area	×	×	×	×	×	×		
Sequoia Natl. Park	×	×	×			×		×
Whiskeytown-Shasta-Trinity Natl. Recr. Area	×	×	×	×	×	×		
Yosemite Natl. Park	×	×	×	×	×	×	×	×

Derivation Operators

For $A \subseteq G$ we define
$A' := \{m \in M \mid \forall g \in A : (g, m) \in I\}$.

For $B \subseteq M$ we define
$B' := \{g \in G \mid \forall m \in B : (g, m) \in I\}$.

(X' is spoken "X prime")

National Parks in California	NPS Guided Tours	Hiking	Horseback Riding	Swimming	Boating	Fishing	Bicycle Trail	Cross Country Trail
Cabrillo Natl. Mon.						×	×	
Channel Islands Natl. Park		×		×		×		
Death Valley Natl. Mon.	×	×	×	×			×	
Devils Postpile Natl. Mon.	×	×	×	×		×		
Fort Point Natl. Historic Site	×					×		
Golden Gate Natl. Recreation Area	×	×	×	×		×	×	
John Muir Natl. Historic Site	×							
Joshua Tree Natl. Mon.	×	×	×					
Kings Canyon Natl. Park	×	×	×			×		×
Lassen Volcanic Natl. Park	×	×	×	×	×	×		×
Lava Beds Natl. Mon.	×	×						
Muir Woods Natl. Mon.		×						
Pinnacles Natl. Mon.		×						
Point Reyes Natl. Seashore	×	×	×	×		×	×	
Redwood Natl. Park	×	×	×	×		×		
Santa Monica Mts. Natl. Recr. Area	×	×	×	×	×	×		
Sequoia Natl. Park	×	×	×			×		×
Whiskeytown-Shasta-Trinity Natl. Recr. Area	×	×	×	×	×	×		
Yosemite Natl. Park	×	×	×	×	×	×	×	×

A

Derivation Operators

For $A \subseteq G$ we define $A' := \{m \in M \mid \forall g \in A : (g, m) \in I\}$.

For $B \subseteq M$ we define $B' := \{g \in G \mid \forall m \in B : (g, m) \in I\}$.

(X' is spoken "X prime")

A'

National Parks in California	NPS Guided Tours	Hiking	Horseback Riding	Swimming	Boating	Fishing	Bicycle Trail	Cross Country Trail
Cabrillo Natl. Mon.						×	×	
Channel Islands Natl. Park		×		×		×		
Death Valley Natl. Mon.	×	×	×	×			×	
Devils Postpile Natl. Mon.	×	×	×	×		×		
Fort Point Natl. Historic Site	×					×		
Golden Gate Natl. Recreation Area	×	×	×	×		×	×	
John Muir Natl. Historic Site	×							
Joshua Tree Natl. Mon.	×	×	×					
Kings Canyon Natl. Park	×	×	×			×		×
Lassen Volcanic Natl. Park	×	×	×	×	×	×		×
Lava Beds Natl. Mon.	×	×						
Muir Woods Natl. Mon.		×						
Pinnacles Natl. Mon.		×						
Point Reyes Natl. Seashore	×	×	×	×		×	×	
Redwood Natl. Park	×	×	×	×		×		
Santa Monica Mts. Natl. Recr. Area	×	×	×	×	×	×		
Sequoia Natl. Park	×	×	×			×		×
Whiskeytown-Shasta-Trinity Natl. Recr. Area	×	×	×	×	×	×		
Yosemite Natl. Park	×	×	×	×	×	×	×	×

A

Derivation Operators: Properties

For $A, A_1, A_2 \subseteq G$

- $A_1 \subseteq A_2 \Rightarrow A_2' \subseteq A_1'$
- $A \subseteq A''$
- $A' = A'''$

holds.

For $B, B_1, B_2 \subseteq M$

- $B_1 \subseteq B_2 \Rightarrow B_2' \subseteq B_1'$
- $B \subseteq B''$
- $B' = B'''$

holds.

National Parks in California	NPS Guided Tours	Hiking	Horseback Riding	Swimming	Boating	Fishing	Bicycle Trail	Cross Country Trail
Cabrillo Natl. Mon.						×	×	
Channel Islands Natl. Park		×		×		×		
Death Valley Natl. Mon.	×	×	×	×			×	
Devils Postpile Natl. Mon.	×	×	×	×		×		
Fort Point Natl. Historic Site	×					×		
Golden Gate Natl. Recreation Area	×	×	×	×		×	×	
John Muir Natl. Historic Site	×							
Joshua Tree Natl. Mon.	×	×	×					
Kings Canyon Natl. Park	×	×	×			×		×
Lassen Volcanic Natl. Park	×	×	×	×	×	×		×
Lava Beds Natl. Mon.	×	×						
Muir Woods Natl. Mon.		×						
Pinnacles Natl. Mon.		×						
Point Reyes Natl. Seashore	×	×	×	×		×	×	
Redwood Natl. Park	×	×	×	×		×		
Santa Monica Mts. Natl. Recr. Area	×	×	×	×	×	×		
Sequoia Natl. Park	×	×	×			×		×
Whiskeytown-Shasta-Trinity Natl. Recr. Area	×	×	×	×	×	×		
Yosemite Natl. Park	×	×	×	×	×	×	×	×

Formal Concept

Def.: A *formal concept* is a pair (A, B) with

- $A \subseteq G$ and $B \subseteq M$
- $A' = B$
- $B' = A$

A is the *extent* and B the *intent* of the concept.

intent

extent

National Parks in California	NPS Guided Tours	Hiking	Horseback Riding	Swimming	Boating	Fishing	Bicycle Trail	Cross Country Trail
Cabrillo Natl. Mon.						×	×	
Channel Islands Natl. Park		×		×		×		
Death Valley Natl. Mon.	×	×	×	×			×	
Devils Postpile Natl. Mon.	×	×	×	×		×		
Fort Point Natl. Historic Site	×					×		
Golden Gate Natl. Recreation Area	×	×	×	×		×	×	
John Muir Natl. Historic Site	×							
Joshua Tree Natl. Mon.	×	×	×					
Kings Canyon Natl. Park	×	×	×			×		×
Lassen Volcanic Natl. Park	×	×	×	×	×	×		×
Lava Beds Natl. Mon.	×	×						
Muir Woods Natl. Mon.		×						
Pinnacles Natl. Mon.		×						
Point Reyes Natl. Seashore	×	×	×	×		×	×	
Redwood Natl. Park	×	×	×	×		×		
Santa Monica Mts. Natl. Recr. Area	×	×	×	×	×	×		
Sequoia Natl. Park	×	×	×			×		×
Whiskeytown-Shasta-Trinity Natl. Recr. Area	×	×	×	×	×	×		
Yosemite Natl. Park	×	×	×	×	×	×	×	×

Formal Concept

Lemma: (A, B) is a formal concept iff $A \subseteq G$, $B \subseteq M$ and A and B are both maximal with respect to $A \times B \subseteq I$.
I.e., every concept corresponds to a maximal rectangle in the relation I.

Def.: The set of all concepts of (G, M, I) is depicted as $\mathfrak{B}(G, M, I)$.

intent

National Parks in California	NPS Guided Tours	Hiking	Horseback Riding	Swimming	Boating	Fishing	Bicycle Trail	Cross Country Trail
Cabrillo Natl. Mon.						×	×	
Channel Islands Natl. Park		×		×		×		
Death Valley Natl. Mon.	×	×	×	×			×	
Devils Postpile Natl. Mon.	×	×	×	×		×		
Fort Point Natl. Historic Site	×					×		
Golden Gate Natl. Recreation Area	×	×	×	×		×	×	
John Muir Natl. Historic Site	×							
Joshua Tree Natl. Mon.	×	×	×					
Kings Canyon Natl. Park	×	×	×			×		×
Lassen Volcanic Natl. Park	×	×	×	×	×	×		×
Lava Beds Natl. Mon.	×	×						
Muir Woods Natl. Mon.		×						
Pinnacles Natl. Mon.		×						
Point Reyes Natl. Seashore	×	×	×	×		×	×	
Redwood Natl. Park	×	×	×	×		×		
Santa Monica Mts. Natl. Recr. Area	×	×	×	×	×	×		
Sequoia Natl. Park	×	×	×			×		×
Whiskeytown-Shasta-Trinity Natl. Recr. Area	×	×	×	×	×	×		
Yosemite Natl. Park	×	×	×	×	×	×	×	×

extent

Formal Concept: Subconcept and Superconcept

The blue concept is a *subconcept* of the yellow concept because

- the blue extent is contained in the yellow extent
- ($\Leftrightarrow$ the yellow intent is contained in the blue intent)

Def.:

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

National Parks in California	*NPS Guided Tours*	*Hiking*	*Horseback Riding*	*Swimming*	*Boating*	*Fishing*	*Bicycle Trail*	*Cross Country Trail*
Cabrillo Natl. Mon.						×	×	
Channel Islands Natl. Park		×		×		×		
Death Valley Natl. Mon.	×	×	×	×			×	
Devils Postpile Natl. Mon.	×	×	×	×		×		
Fort Point Natl. Historic Site	×					×		
Golden Gate Natl. Recreation Area	×	×	×	×		×	×	
John Muir Natl. Historic Site	×							
Joshua Tree Natl. Mon.	×	×	×					
Kings Canyon Natl. Park	×	×	×			×		×
Lassen Volcanic Natl. Park	×	×	×	×	×	×		×
Lava Beds Natl. Mon.	×	×						
Muir Woods Natl. Mon.		×						
Pinnacles Natl. Mon.		×						
Point Reyes Natl. Seashore	×	×	×	×		×	×	
Redwood Natl. Park	×	×	×	×		×		
Santa Monica Mts. Natl. Recr. Area	×	×	×	×	×	×		
Sequoia Natl. Park	×	×	×			×		×
Whiskeytown-Shasta-Trinity Natl. Recr. Area	×	×	×	×	×	×		
Yosemite Natl. Park	×	×	×	×	×	×	×	×

Concept Lattice

(Recapitulation: Partial Order)

Def. (recap.): $(A_1, B_1) \leqslant (A_2, B_2) :\Leftrightarrow A_1 \subseteq A_2 (\Leftrightarrow B_1 \supseteq B_2)$

Def.: The set of all concepts $\mathfrak{B}(G, M, I)$ together with the partial order $\leqslant$ is the *concept lattice* of the context (G, M, I) and is depicted with $\underline{\mathfrak{B}}(G, M, I)$.

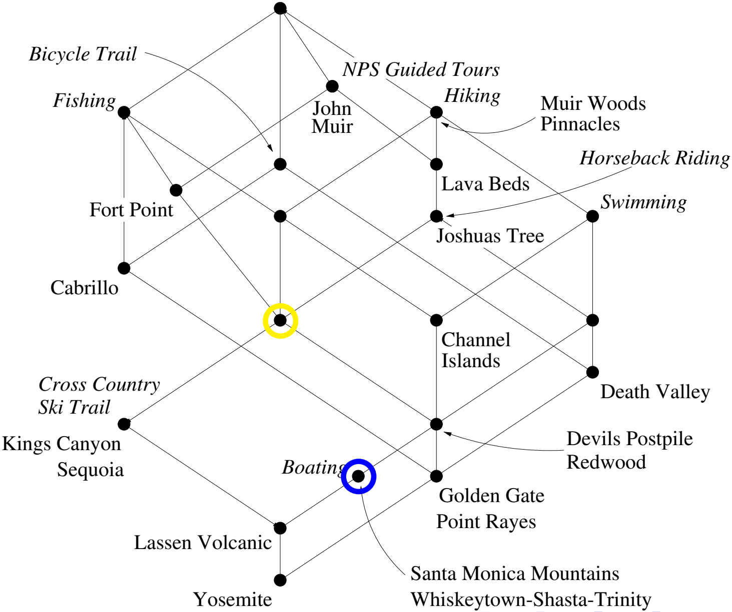
On the blackboard:

- definition of partial order
- definition of total order
- examples

Concept Lattice: as Line Diagram

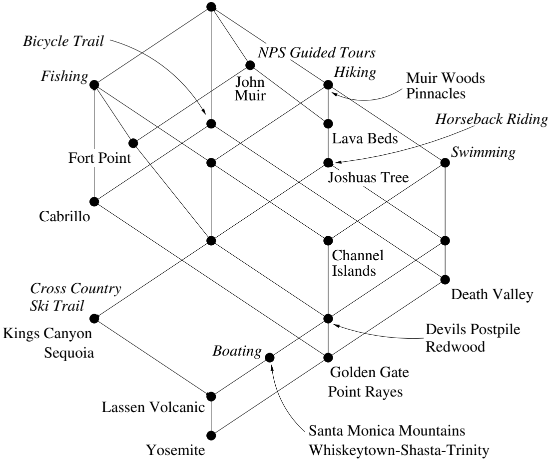
The *concept lattice* for the national park context.

National Parks in California	*NPS Guided Tours*	*Hiking*	*Horseback Riding*	*Swimming*	*Boating*	*Fishing*	*Bicycle Trail*	*Cross Country Trail*
Cabrillo Natl. Mon.						x	x	
Channel Islands Natl. Park		x		x		x		
Death Valley Natl. Mon.	x	x	x	x			x	
Devils Postpile Natl. Mon.	x	x	x	x		x		
Fort Point Natl. Historic Site	x					x		
Golden Gate Natl. Recreation Area	x	x	x	x		x	x	
John Muir Natl. Historic Site	x							
Joshua Tree Natl. Mon.	x	x	x					
Kings Canyon Natl. Park	x	x	x			x		x
Lassen Volcanic Natl. Park	x	x	x	x	x	x		x
Lava Beds Natl. Mon.	x	x						
Muir Woods Natl. Mon.		x						
Pinnacles Natl. Mon.		x						
Point Reyes Natl. Seashore	x	x	x	x		x	x	
Redwood Natl. Park	x	x	x	x		x		
Santa Monica Mts. Natl. Recr. Area	x	x	x	x	x	x		
Sequoia Natl. Park	x	x	x			x		x
Whiskeytown-Shasta-Trinity Natl. Recr. Area	x	x	x	x	x	x		
Yosemite Natl. Park	x	x	x	x	x	x	x	x



Concept Lattice: Implications (Preview)

Def.: An *implication* $X \to Y$ *holds* in a context, if every object that has all attributes from X also has all attributes from Y.



Examples:

- $\{$*Swimming*$\} \to \{$*Hiking*$\}$
- $\{$*Boating*$\} \to \{$*Swimming, Hiking, NPS Guided Tours, Fishing, Horseback Riding*$\}$
- $\{$*Bicycle Trail, NPS Guided Tours*$\} \to \{$*Swimming, Hiking, Horseback Riding*$\}$

Concept Lattice: Dual Context

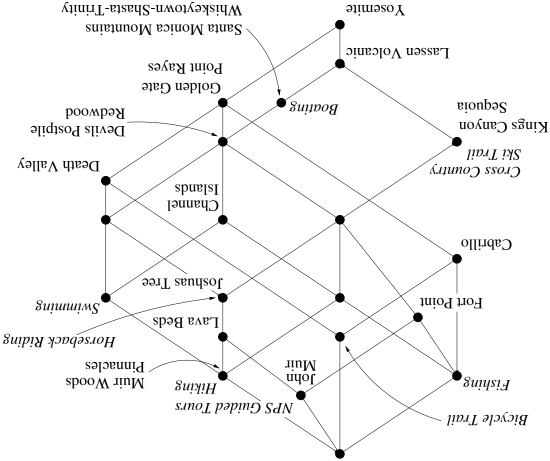
Def.: Let (G, M, I) be a context. Then (M, G, I^{-1}) with

$$(m, g) \in I^{-1} \iff (g, m) \in I$$

is the *dual context* of (G, M, I).

Theorem: Its concept lattice is isomorphic to $(\mathfrak{B}(G, M, I), \geqslant)$.

Remark: In general, G and M need not be disjunct, they can even be identical.



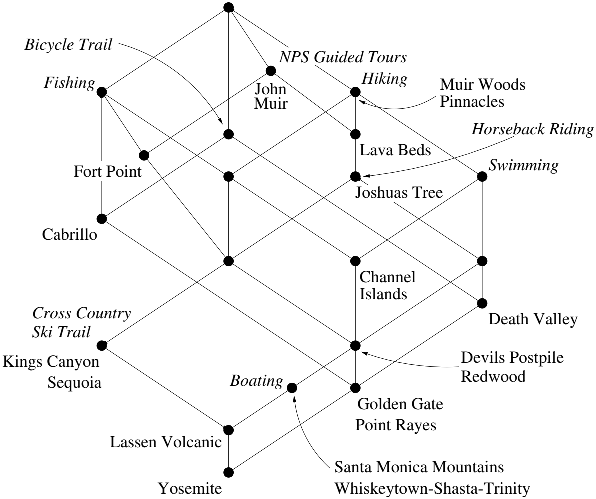
Concept Lattice

Recapitulation: Lattices

On the blackboard:

- lower bound, upper bound
- infimum (join, $\bigwedge$), supremum (meet, $\bigvee$)
- **Lemma:** For two formal concepts (A_1, B_1), (A_2, B_2) we get
  - the infimum $(A_1, B_1) \wedge (A_2, B_2)$ as $(A_1 \cap A_2, (B_1 \cup B_2)'')$
  - the supremum $(A_1, B_1) \vee (A_2, B_2)$ as $((A_1 \cup A_2)'', B_1 \cap B_2)$
- Def. (complete) lattice $(V, \leqslant)$
- $\mathbf{0}_V$, $\mathbf{1}_V$

Concept Lattice: The Basic Theorem on Concept Lattices



On the blackboard:

- supremum/infimum reducible, irreducible, dense
- isomorphisms of lattices
- Basic Theorem

Concept Lattice: The Duality Principle

- Let $(V, \leqslant)$ be a (complete) lattice. Then $(V, \geqslant)$ is also a (complete) lattice.
- (cf. with the definition of the dual context)
- If a theorem holds for (complete) lattices, then the 'dual theorem' also holds, i.e., the theorem where all occurrences of $\leqslant, \cap, \cup, \wedge, \vee, \mathbf{1}_V, \mathbf{0}_V$, etc. have been replaced by $\geqslant, \cup, \cap, \vee, \wedge, \mathbf{0}_V, \mathbf{1}_V$, etc.

Computing All Concepts

There are several algorithms to compute all concepts:

- naive approach
- intersection method
- Next-Closure (Ganter 1984) → Chapter 3
- TITANIC (Stumme et al. 2001) → Chapter 3
- and several incremental algorithms

Computing All Concepts: Naive Approach

Theorem

Each concept of a context (G, M, I) *has the form* (X'', X') *for some subset* $X \subseteq G$ *and* (Y', Y'') *for some subset* $Y \subseteq M$*.*
Conversely, all such pairs are concepts.

Algorithm

Determine for every subset Y of M the pair (Y', Y'').

Computing All Concepts: Naive Approach

Theorem

Each concept of a context (G, M, I) has the form (X'', X') for some subset $X \subseteq G$ and (Y', Y'') for some subset $Y \subseteq M$.
Conversely, all such pairs are concepts.

Algorithm

Determine for every subset Y of M the pair (Y', Y'').

Inefficient! (Too) many concepts are generated multiple times.

Computing All Concepts: Intersection Method

- Suitable for manual computation (Wille 1982)
- Best worst-case time complexity (Nourine, Raynoud 1999)
- Based on the following

Theorem

Every extent is the intersection of attribute extents. (I.e., the closure system of all extents is generated by the attribute extents.)

Which intersections of attribute extents should we take?

On the blackboard: attribute extent $\{m\}'$, closure system

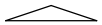
Computing All Concepts: Intersection Method

How to determine all formal concepts of a formal context:

1. For each attribute $m \in M$ compute the attribute extent $\{m\}'$.
2. For any two sets in this list, compute their intersection. If it is not yet contained in the list, add it.
3. Repeat until no new extents are generated.
4. If G is not yet contained in the list, add it.
5. For every extent A in the list compute the corresponding intent A'.

Computing All Concepts: Intersection Method

On the blackboard: "triangle" example

Triangles				
abbreviation	coordinates			diagram
T1	(0,0)	(6,0)	(3,1)	
T2	(0,0)	(1,0)	(1,1)	
T3	(0,0)	(4,0)	(1,2)	
T4	(0,0)	(2,0)	$(1,\sqrt{3})$	
T5	(0,0)	(2,0)	(5,1)	
T6	(0,0)	(2,0)	(1,3)	
T7	(0,0)	(2,0)	(0,1)	

Attributes	
symbol	property
a	equilateral
b	isoceles
c	acute angled
d	obtuse angled
e	right angled

	a	b	c	d	e
$T1$		×		×	
$T2$		×			×
$T3$			×		
$T4$	×	×	×		
$T5$				×	
$T6$		×	×		
$T7$					×

Drawing Concept Lattices

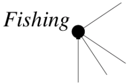
How to draw a concept lattice by hand:

1. Draw a small circle for the extent G at the top.
2. For every extent (starting with the one's with the most elements) draw a small circle and connect it with the lowest circle(s) whose extent contains the current extent.
3. Every attribute is written slightly above the circle of its attribute extent.
4. Every object is written slightly below the circle that is exactly below the circles that are labeled with the attributes of the object.

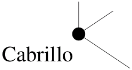
Drawing Concept Lattices

How you can check the drawn diagram:

1. Is it really a lattice? (that's often skipped)
2. Is every concept with exactly one upper neighbor labeled with at least one attribute?



3. Is every concept with exactly one lower neighbor labeled with at least one object?



4. Is for every $g \in G$ and $m \in M$ the label of the object g below the label of the attribute m iff $(g, m) \in I$ holds?

Clarifying and Reducing a Formal Context

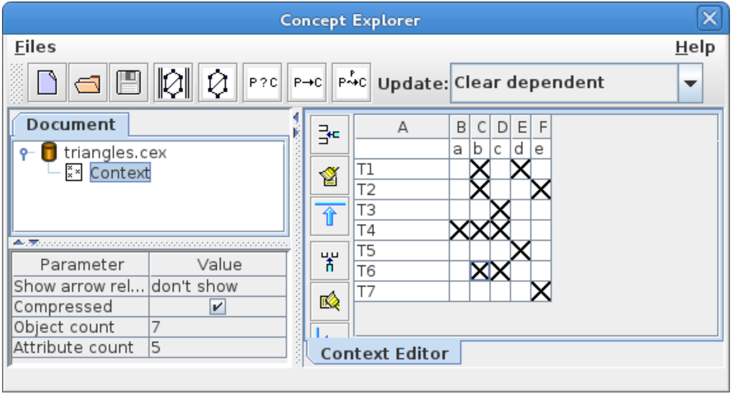
On the blackboard:

- proper subconcept ($<$)
- lower neighbor ($\prec$)
- reducible objects/attributes
- clarifying and reducing
- reduced context, standard context

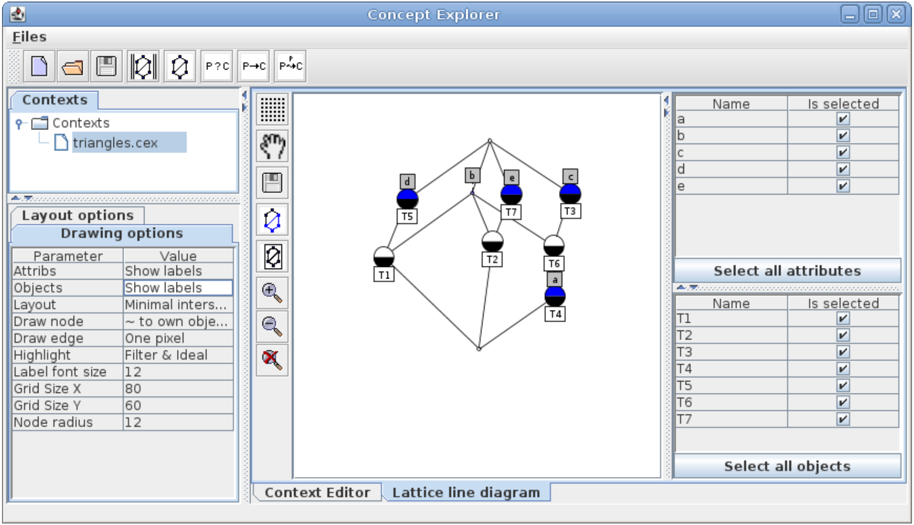
Theorem

A finite context and its reduced context have isomorphic concept lattices. For every finite lattice L there is (up to isomorphism) exactly one reduced context, the concept lattice of which is isomorphic to L, namely its standard context.

Interlude: ConExp



Interlude: ConExp



Additive Line Diagrams

Def.: An attribute $m \in M$ is called *irreducible*, if there is no set X of attributes with $m \notin X$ such that $\{m\}' = \bigcap_{x \in X} \{x\}'$.
The set of irreducible attributes is depicted as M_{irr}.

We define the map $\mathrm{irr} : \underline{\mathfrak{B}}(G, M, I) \to \mathfrak{P}(M_{irr})$ as

$$\mathrm{irr}(A, B) := \{m \in B \mid m \text{ irreducible}\}.$$

Let $\mathrm{vec} : M_{irr} \to \mathbb{R} \times \mathbb{R}_{<0}$. Then

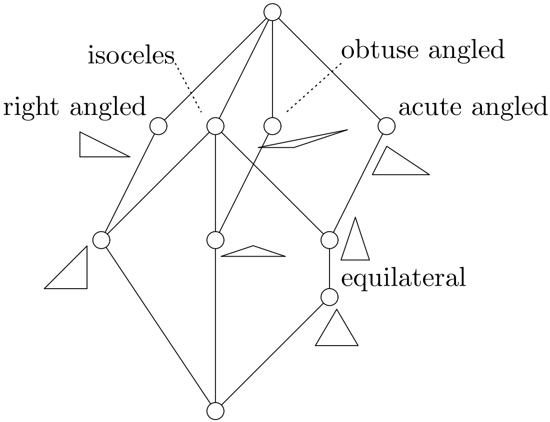
$$\mathrm{pos} : \underline{\mathfrak{B}}(G, M, I) \to \mathbb{R}^2 \text{ with } \mathrm{pos}(A, B) := \sum_{m \in \mathrm{irr}(A,B)} \mathrm{vec}(m)$$

is an *additive line diagram* of the concept lattice $\underline{\mathfrak{B}}(G, M, I)$.

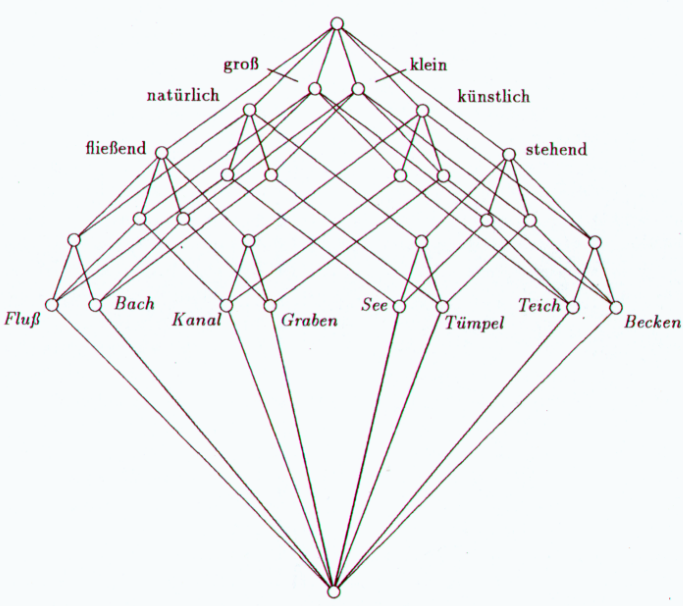
Additive Line Diagrams

An additive line diagram of the triangles context.

The position of the attribute concepts defines the position of all remaining concepts. If we consider the distance between $\mathbf{1}_{\underline{\mathfrak{B}}}$ and the attribute extents as vectors, then the position of any concept is equal to the sum of the vectors that belong to its concept intent.

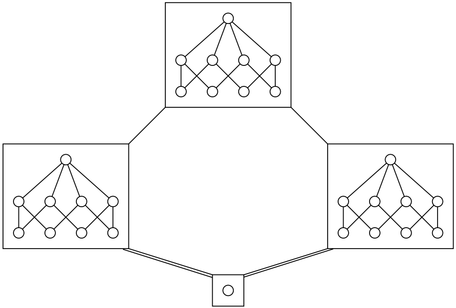


Additive Line Diagrams

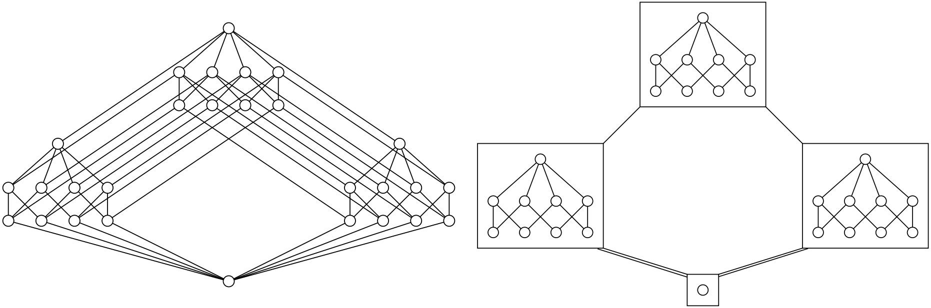


Nested Line Diagrams: Motivation and Idea

- readability of line diagrams often lost for many concepts ($\gtrapprox 50$)
- *nested line diagrams* allow us to go further
- and: support the visualization of changes caused by the addition of further attributes
- *basic idea:* cluster parts of an ordinary diagram and replace bundles of parallel lines between these parts by one line each



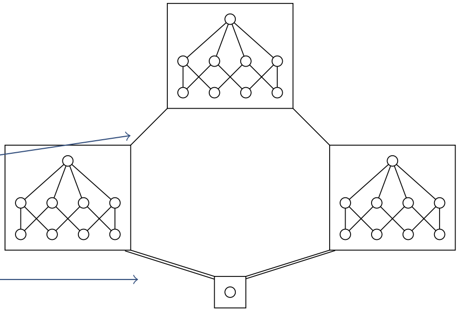
Nested Line Diagrams: Example



The previous concept lattice as ordinary line diagram and as nested line diagram. (For simplification, object and attribute labels have been omitted.)

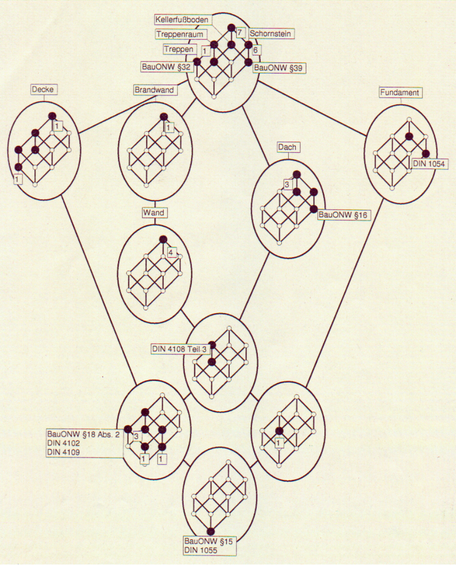
Nested Line Diagrams

- a nested line diagram consists of *boxes* which contain parts of the ordinary diagram and which are *connected by lines*
- simplest case: two boxes that are connected by a line are congruent → corresponding circles are direct neighbors
- double lines between two boxes: every element of the upper box is larger than every element of the lower box



Nested Line Diagrams

- two boxes connected by a single line need not be congruent but contain a part of two congruent figures
- the two congruent figures are drawn as "background structure" into the boxes
- elements are drawn as bold circles if they are part of the respective substructure
- the line connecting both boxes indicates that the respective pairs of elements of the background shall be connected with each other



Nested Line Diagrams: Drawing Example

Die Ducks. Psychogramm einer Sippe.

	generation			sex		financial status		
	older	middle	younger	♂	♀	rich	carefree	indebted
Tick			×	×			×	
Trick			×	×			×	
Track			×	×			×	
Donald		×		×				×
Daisy		×			×		×	
Gustav		×		×			×	
Dagobert	×			×		×		
Annette	×				×		×	
Primus v. Quack	×			×			×	

Taken from: Grobian Gans: *Die Ducks. Psychogramm einer Sippe.* Rowohlt, Reinbek bei Hamburg 1972, ISBN 3-499-11481-X

Nested Line Diagrams: Construction

1. split the attribute set: $M = M_1 \cup M_2$
   (needs not be disjoint, more important: both sets bear meaning)
2. draw the line diagrams of the subcontexts

$$\mathbb{K}_i := (G, M_i, I \cap G \times M_i), i \in \{1, 2\}$$

   and label them with with objects and attributes, as usual

   On the blackboard: **Theorem 2** (script, p. 35)

3. sketch a nested diagram of the product of the concept lattices $\underline{\mathfrak{B}}(\mathbb{K}_i)$
   1. draw a large diagram of $\underline{\mathfrak{B}}(\mathbb{K}_1)$ where the concepts are large boxes
   2. draw a copy of $\underline{\mathfrak{B}}(\mathbb{K}_2)$ into each box

Nested Line Diagrams: Labeling

- by Theorem 2, $\underline{\mathfrak{B}}(G, M, I)$ is embedded in this product as $\bigvee$-semilattice
- if a list of elements of $\underline{\mathfrak{B}}(G, M, I)$ exists, enter them according to their intents
- otherwise, enter the object concepts (whose intents can be read off directly from the context) and form all suprema

This gives us another method for determining a concept lattice by hand:

- split up the attribute set as appropriate
- determine the (small) concept lattices of the subcontexts
- draw their product as nested line diagram
- enter the object concepts and close against suprema

This is particularly advisable in order to arrive at a useful diagram quickly.

Nested Line Diagrams: Example

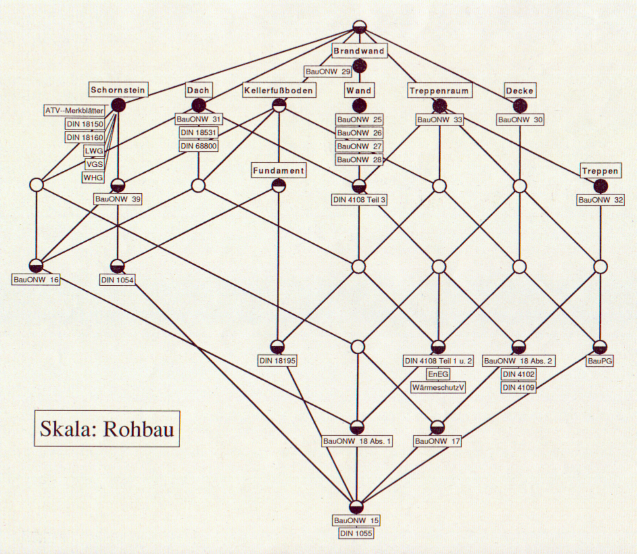
Baurecht in Nordrhein-Westfalen

Taken from: D. Eschenfelder, W. Kollewe, M. Skorsky, R. Wille: *Ein Erkundungssystem zum Baurecht: Methoden der Entwicklung eines TOSCANA-Systems*. In: G. Stumme, R. Wille (Eds.): Begriffliche Wissensverarbeitung – Methoden und Anwendungen. Springer 2000

	Dach	Decke	Wand	Brandwand	Treppen	Treppenraum	Fundament	Kellerfußboden	Schornstein
BauONW 15	×	×	×	×	×	×	×	×	×
BauONW 16	×							×	×
BauONW 17	×	×	×	×	×	×			×
BauONW 18 Abs. 1	×	×	×	×		×		×	×
BauONW 18 Abs. 2	×	×	×	×	×	×			
BauONW 25			×	×					
BauONW 26			×	×					
BauONW 27			×	×					
BauONW 28			×	×					
BauONW 29				×					
BauONW 30		×							
BauONW 31	×								
BauONW 32					×	×			
BauONW 33						×			
BauONW 36									
BauONW 39								×	×
BauONW 40									
BimSchG									
BauPG		×			×	×		×	
EnEG	×	×	×	×		×		×	
WHG									×
LWG									×
WärmeschutzV	×	×	×	×		×		×	
HeizAnlV									
BImSchV									
VGS									×
DIN 1054							×	×	×
DIN 1055	×	×	×	×	×	×	×	×	×
DIN 4102	×	×	×	×	×	×			
DIN 4108 Teil 1 u. 2	×	×	×	×		×		×	
DIN 4108 Teil 3	×		×	×		×			
DIN 4109	×	×	×	×	×	×			
DIN 18150									×
DIN 18160									×
DIN 18195	×		×	×		×	×	×	
DIN 18531	×								
DIN 68800	×								
DIN-Normen für Feuerungsanlagen									
DIN-Normen für Entwässerung									
ATV-Merkblätter									×

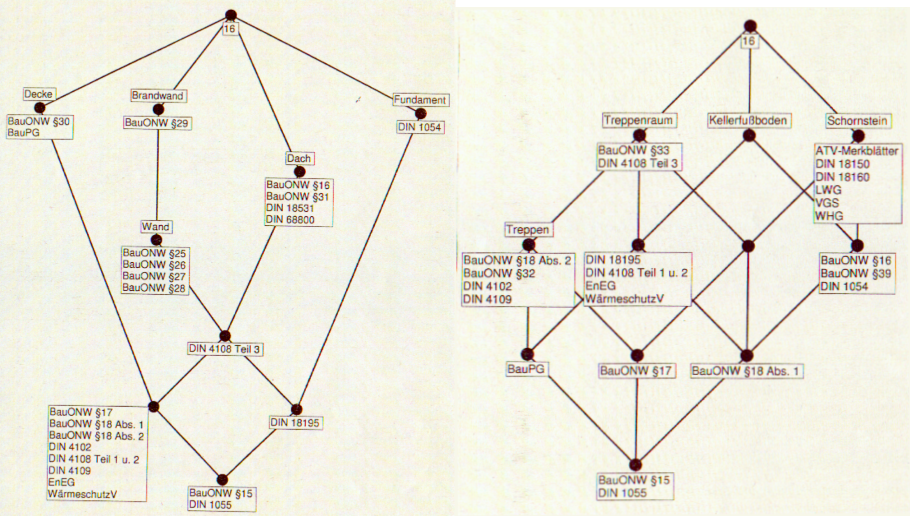
Nested Line Diagrams: Example

Baurecht in Nordrhein-Westfalen



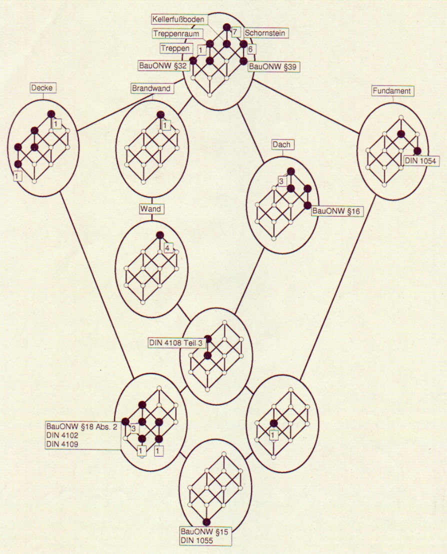
Nested Line Diagrams: Example

Baurecht in Nordrhein-Westfalen



Nested Line Diagrams: Example

Baurecht in Nordrhein-Westfalen



Nested Line Diagrams: Reading off Implications

- implications *within the inner scale* can be read off at the top concept:

$$\{\text{Treppen}\} \Rightarrow \{\text{Treppenraum}\}$$

- implications *within the outer scale* can be read off at it:

$$\{\text{Wand}\} \Rightarrow \{\text{Brandwand}\}$$
$$\{\text{Decke, Brandwand}\} \Rightarrow \{\text{Wand, Brandwand}\}$$
$$\{\text{Decke, Fundament}\} \Rightarrow \{?\}$$

- implications *between the inner and the outer scale* are shown by "not realized" concepts: premise = intent of the not-realized concept, conclusion = intent of the largest realized subconcept:

$$\{\text{Decke, Kellerfußboden}\} \Rightarrow \{\text{Treppenraum}\}$$
$$\{\text{Treppenraum, Schornstein}\} \Rightarrow \{\text{Decke, Wand, Brandwand, Dach}\}$$
$$\{\text{Fundament}\} \Rightarrow \{?\}$$
$$\{\text{Wand, Dach, Schornstein}\} \Rightarrow \{?\}$$