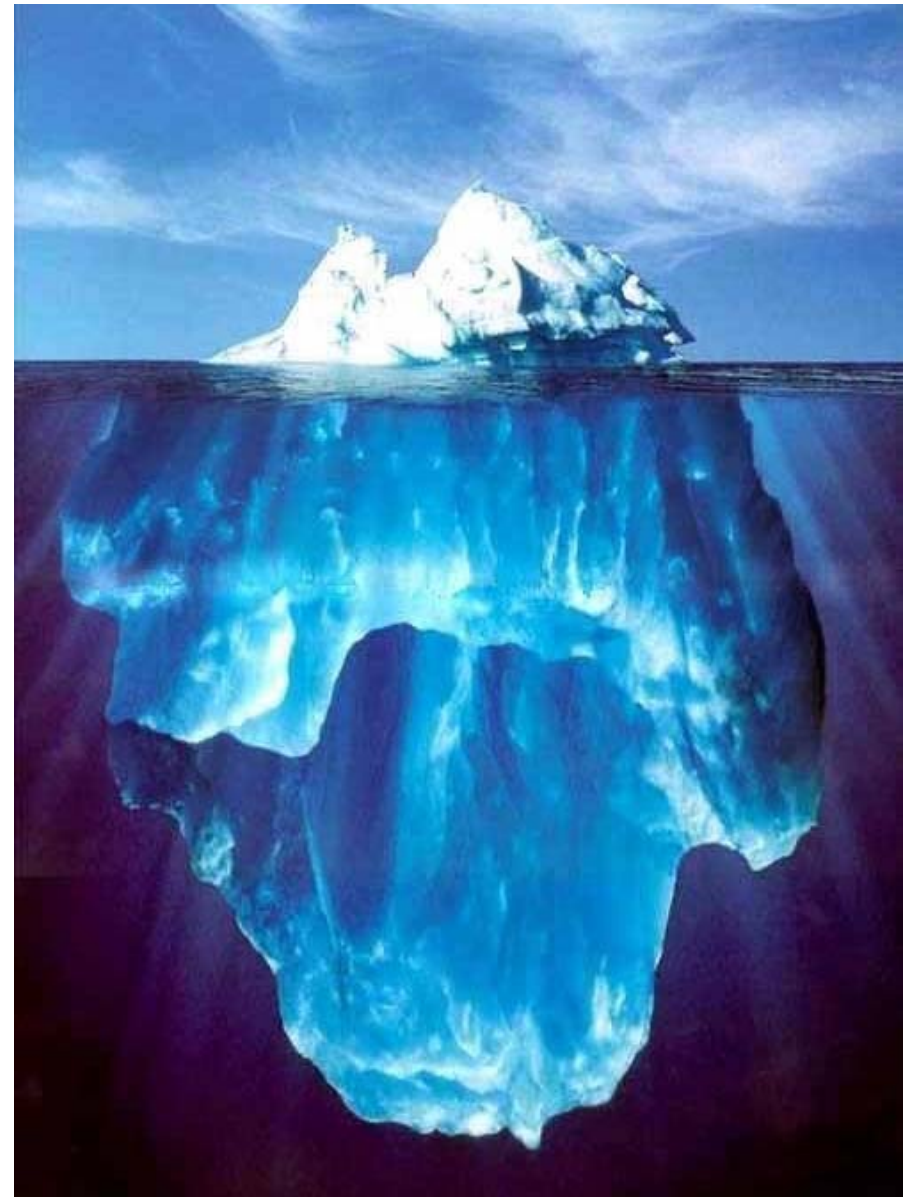


Formale Begriffsanalyse



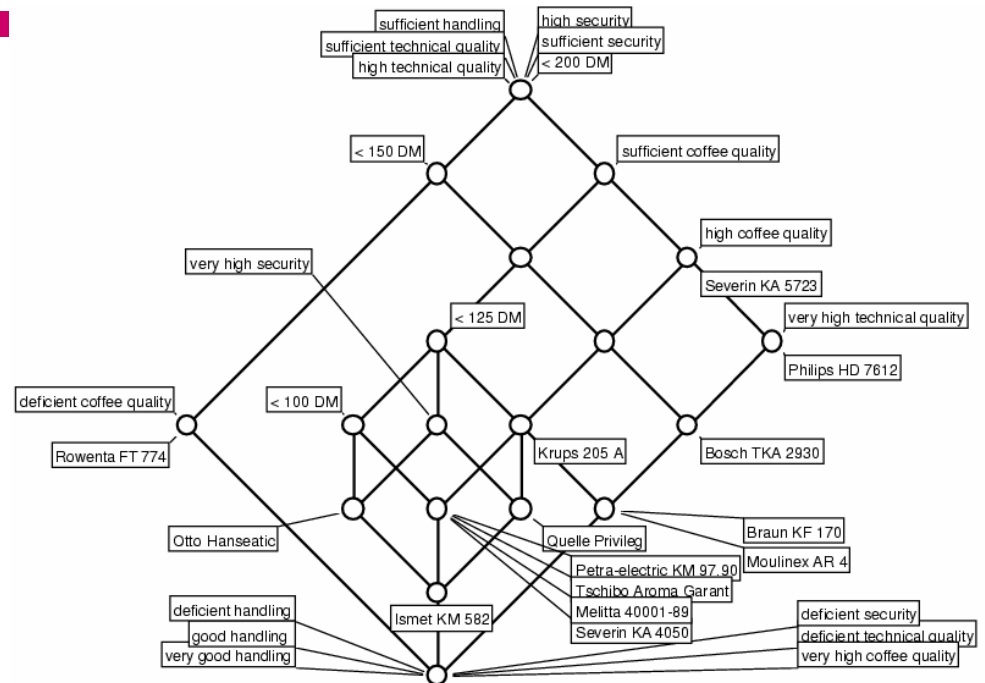
Formale Begriffsanalyse ist um 1980 in Darmstadt entstanden als mathematische Theorie, die eine Formalisierung des Begriffs vom „Begriff“ liefert.

FBA hat seitdem zunehmend Verbreitung in der Informatik gefunden, u.a. in

- der Datenanalyse,
- der Wissensentdeckung,
- dem Software Engineering.

Ausgehend von Datensätzen leitet FBA Begriffshierarchien ab.

FBA bietet die Erzeugung und Visualisierung der Begriffshierarchien auf einer mathematisch fundierten Basis.



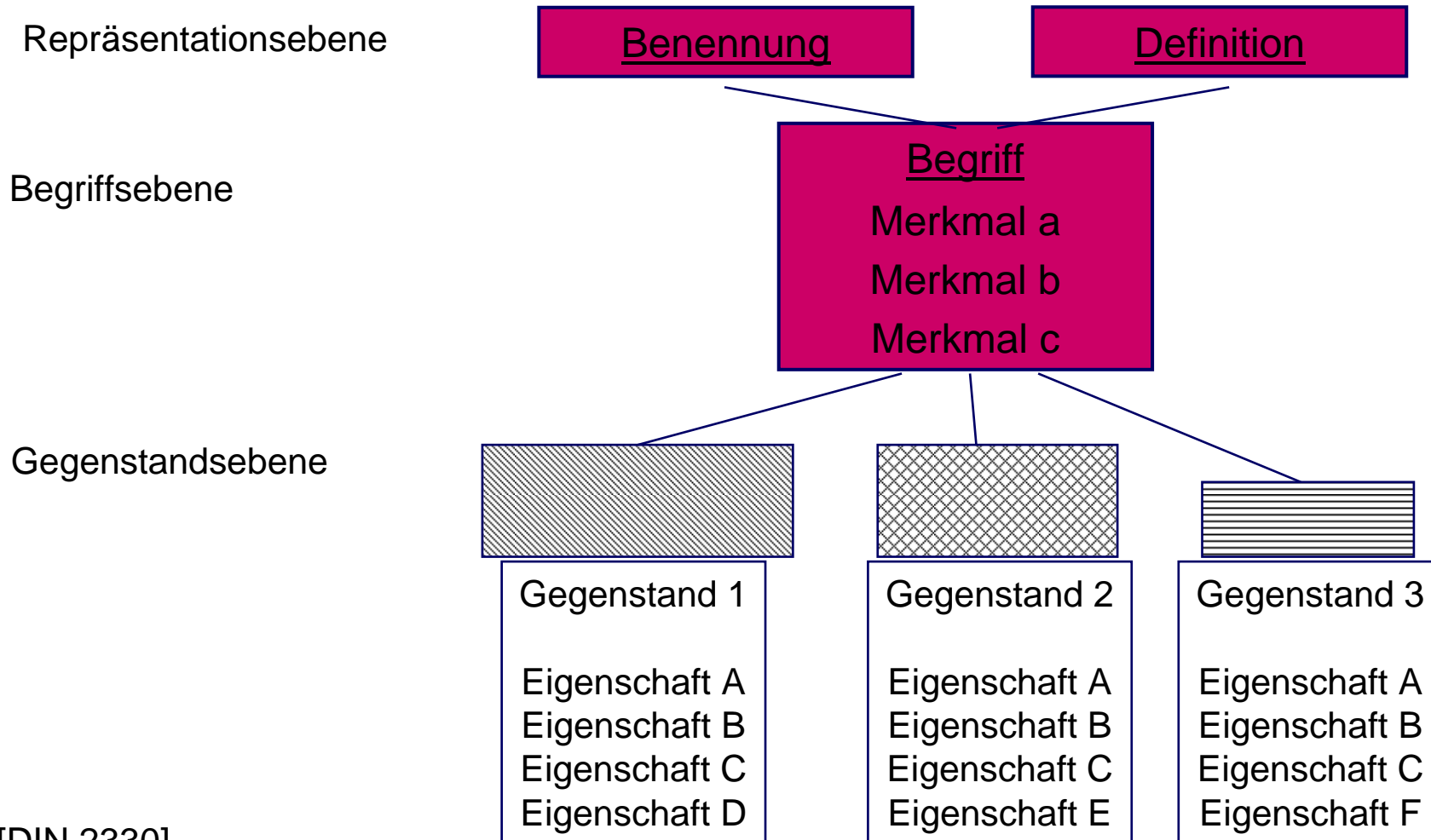
STIFTUNG WARENTEST test KOMPASS KAFFEEMASCHINEN MIT WARMHALTEKANNE (8 bis 10 Tassen) test Ausgabe 12/98

	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⁾ / □					zufriedenst.
Otto Hanseatic Best.-Nr. 4327357	40,-	30,- ²⁾ / □	○	+	++	○	zufriedenst.
Quelle Privileg Best.-Nr. 7030720	40,-	24,50 / 17,50					zufriedenst.
Severin KA 9660	50,-	35,- / 23,-					zufriedenst.
Severin KA 4050	80,-	50,- / □	+	+	+	○	gut
Tchibo Aroma Garant Art.-Nr. 48469	80,-	27,50 / 19,50	+	+	+	○	gut
Ismet KM 582 starlight	84,-	47,- / 14,-	+	+	++	○	gut

-
- Die Formale Begriffsanalyse modelliert Begriffe als Einheiten des Denkens, die aus zwei Teilen bestehen:
 - Der **Begriffsumfang** besteht aus allen Gegenständen, die unter den Begriff fallen.
 - Der **Begriffsinhalt** enthält alle Merkmale, die all diesen Gegenständen zukommen
 - FBA wird unter anderem in der Datenanalyse, dem Information Retrieval, Data Mining, und Software Engineering angewandt.

DIN 2330: Begriffe und Ihre Benennungen

Die Begriffsanalyse bewegt sich auf der Begriffsebene, die Repräsentationsebene spielt keine (bzw. nur eine geringe) Rolle.



[DIN 2330]

Formale Begriffsanalyse

Def.: Ein **formaler Kontext** ist ein Tripel (G, M, I) , wobei

- G eine Menge von Gegenständen,
 - M eine Menge von Merkmalen
 - und I eine Relation zwischen G und M ist.
- $(g, m) \in I$ wird gelesen als „Gegenstand g hat Merkmal 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	×	×	×	×	×	×	×	×

Für $A \subseteq G$ definieren wir

$$A' := \{ m \in M \mid \forall g \in A: (g, m) \in I \}.$$

Für $B \subseteq M$ definieren wir dual

$$B' := \{ g \in G \mid \forall m \in B: (g, m) \in I \}.$$

A

National Parks in California	A'							
	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

Für $A, A_1, A_2 \subseteq G$ gilt:

- $A_1 \subseteq A_2 \Rightarrow A'_2 \subseteq A'_1$
- $A \subseteq A''$
- $A' = A''$

Für $B, B_1, B_2 \subseteq M$ gilt:

- $B_1 \subseteq B_2 \Rightarrow B'_2 \subseteq B'_1$
- $B \subseteq B''$
- $B' = B''$

National Parks in California	A'							
	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

A {

Inhalt

Def.: Ein formaler Begriff

ist ein Paar (A, B) mit

- $A \subseteq G$ und $B \subseteq M$,
- $A' = B$,
- $B' = A$.

A ist der **Umfang** und B der **Inhalt** des Begriffs.

Umfang

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

Der blaue Begriff ist ein **Unterbegriff** des gelben Begriffs, denn:

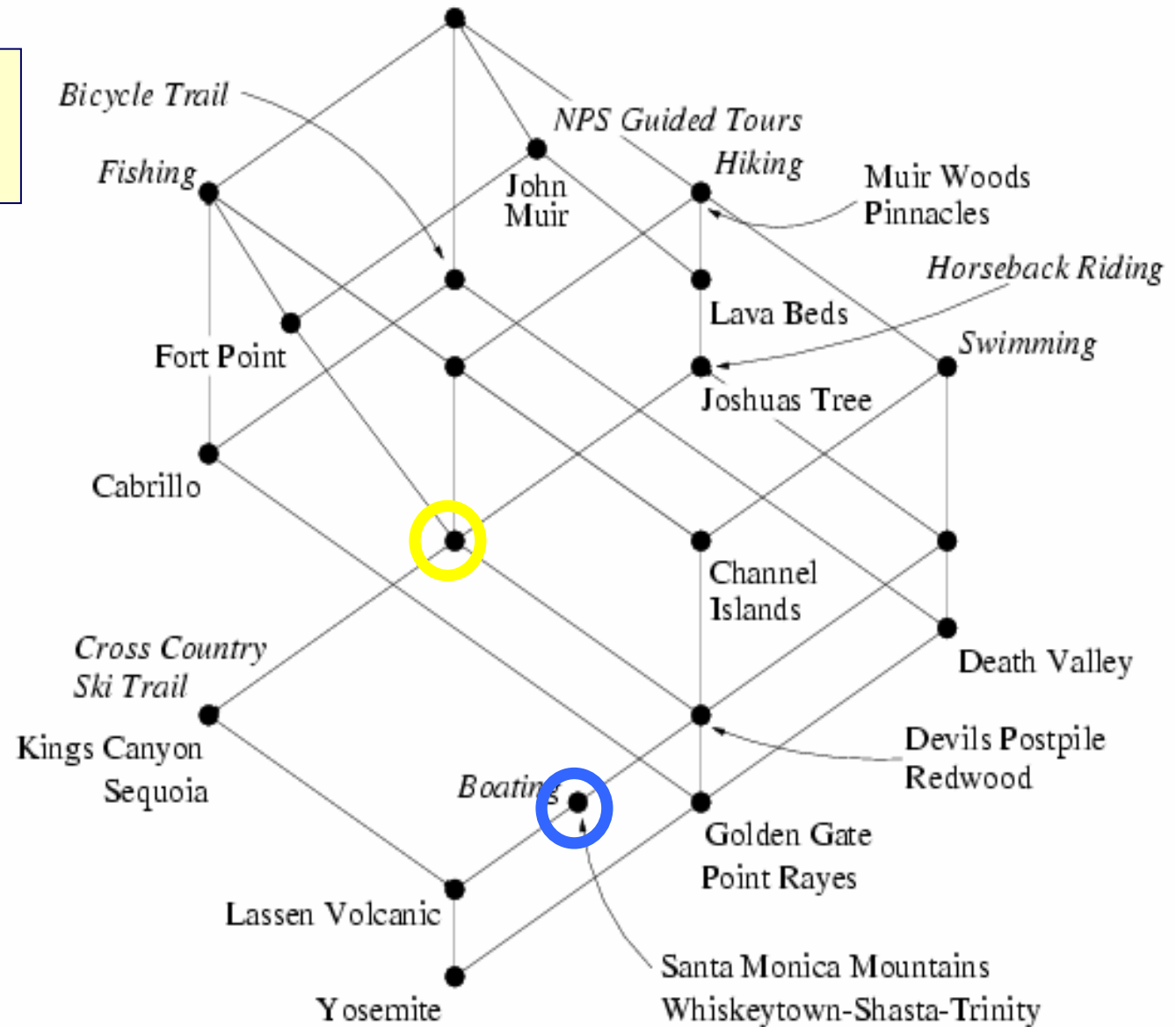
der blaue Umfang ist im gelben Umfang enthalten.

(\Leftrightarrow der gelbe Inhalt ist im blauen Inhalt enthalten.)

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

Der **Begriffsverband** zu dem Nationalpark-Kontext

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		x						x
Death Valley Natl. Mon.		x	x	x				x
Devils Postpile Natl. Mon.		x	x	x				x
Fort Point Natl. Historic Site		x	x	x				
Golden Gate Natl. Recreation Area		x	x	x				x
John Muir Natl. Historic Site		x	x	x				
Joshua Tree Natl. Mon.		x	x	x				
Kings Canyon Natl. Park		x	x	x				x
Lassen Volcanic Natl. Park		x	x	x	x	x		x
Lava Beds Natl. Mon.		x	x	x				
Muir Woods Natl. Mon.		x						
Pinnacles Natl. Mon.		x						
Point Reyes Natl. Seashore		x	x	x				x
Redwood Natl. Park		x	x	x				
Santa Monica Mts. Natl. Recr. Area		x	x	x	x	x		x
Sequoia Natl. Park		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



Implikationen

Def.: Eine **Implikation**

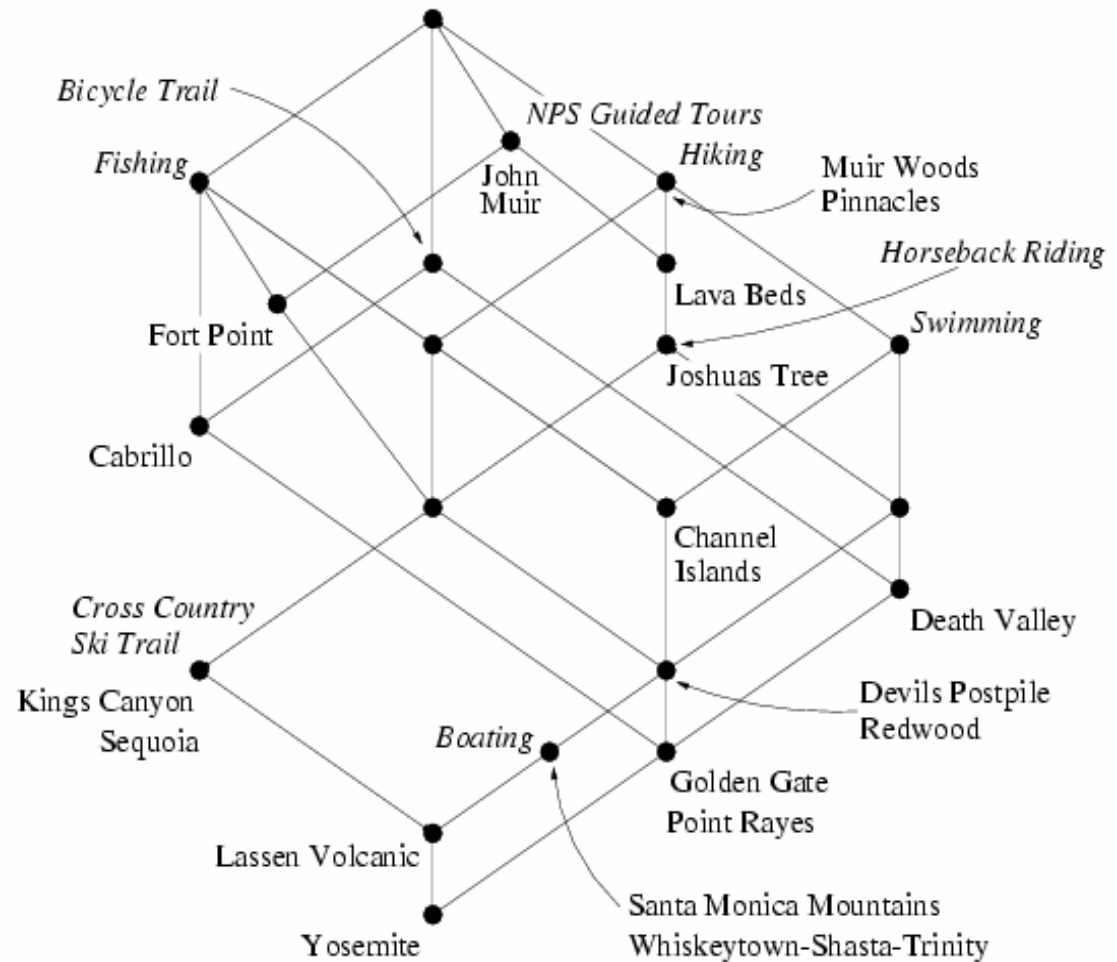
$X \rightarrow Y$ gilt in einem Kontext, wenn jeder Gegenstand, der alle Merkmale aus X hat, auch alle Merkmale aus Y hat.

• **Beispiele:**

{Swimming} \rightarrow {Hiking}

{Boating} \rightarrow {Swimming, Hiking, NPS Guided Tours, Fishing}

{Bicycle Trail, NPS Guided Tours} \rightarrow {Swimming, Hiking}



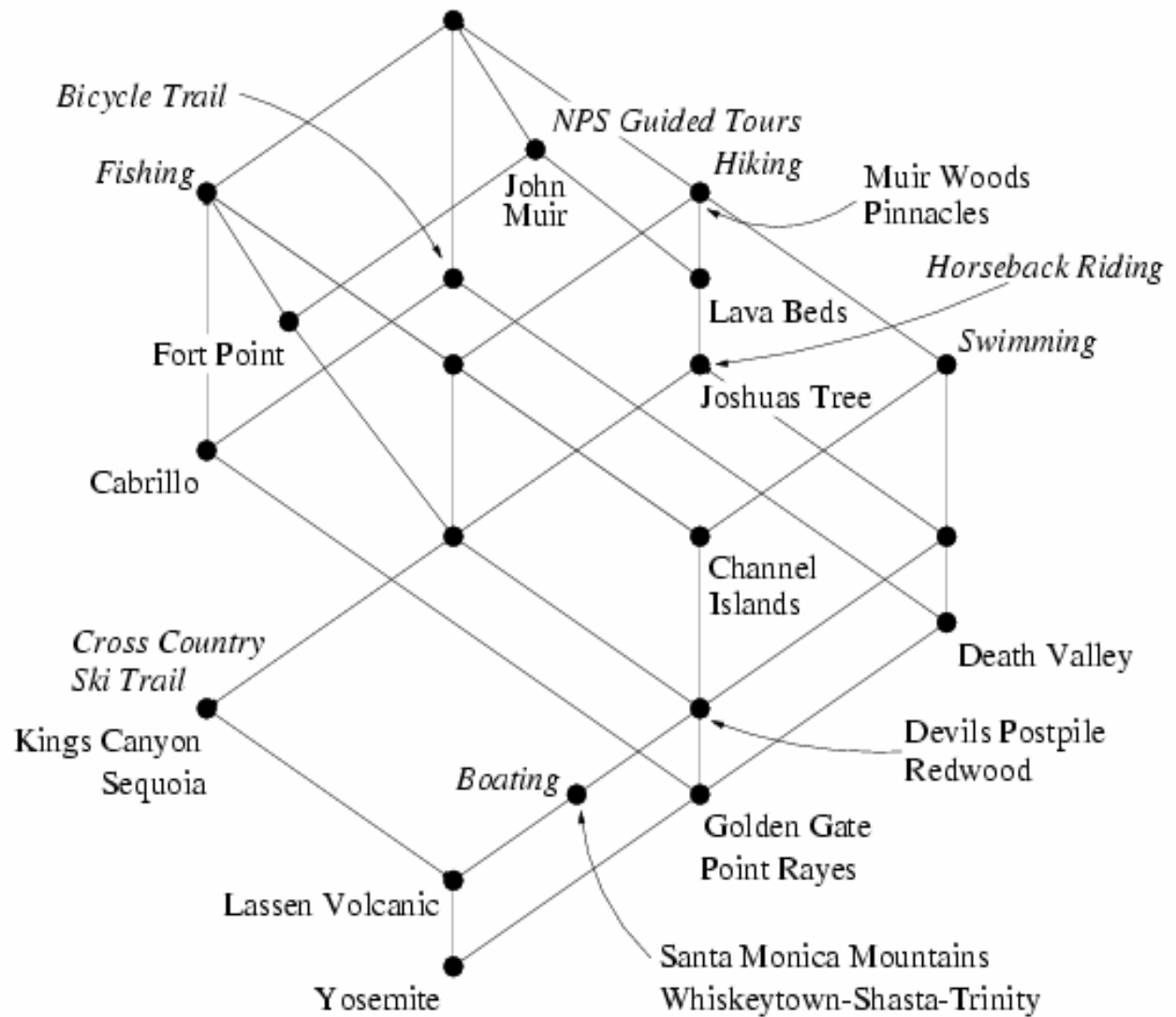
Unabhängigkeit

Def.: Sei $X \subseteq M$. Die Merkmale in X sind **unabhängig voneinander**, wenn es keine nichttrivialen Implikationen zwischen ihnen gibt.

Beispiel:

- Fishing
- Bicycle Trail
- Swimming

sind voneinander unabhängige Merkmale



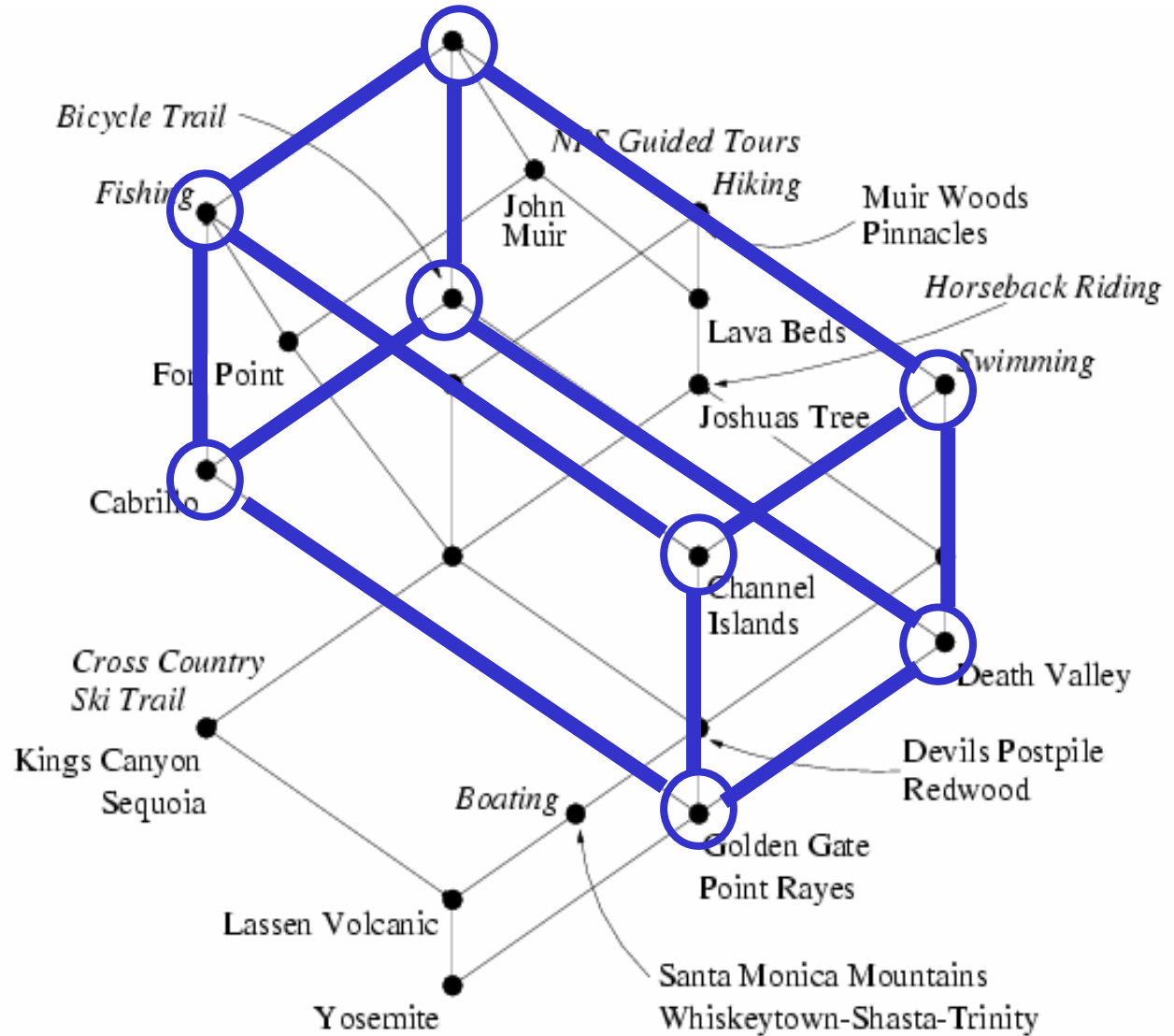
Unabhängigkeit

Lemma: Merkmale sind unabhängig, wenn sie einen (Hyper-) Würfel aufspannen.

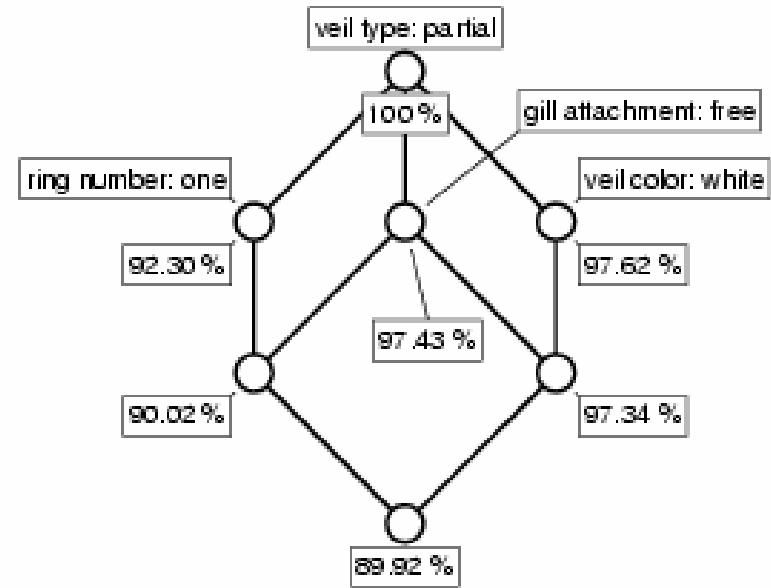
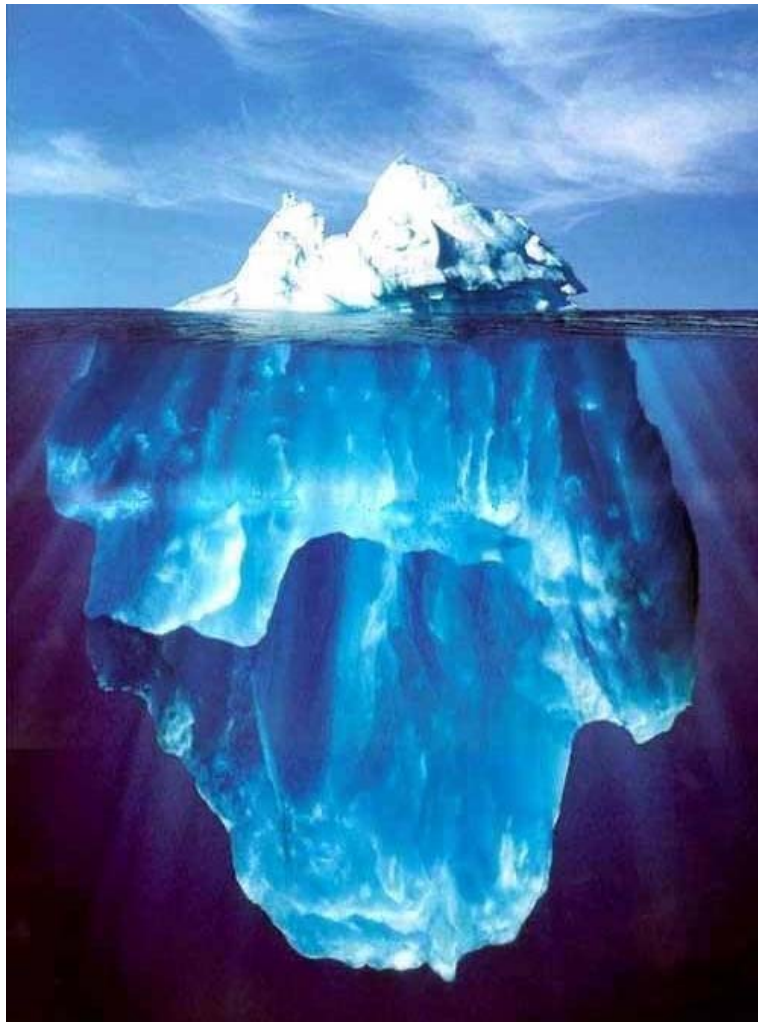
Beispiel:

- Fishing
- Bicycle Trail
- Swimming

sind voneinander unabhängige Merkmale



Iceberg Concept Lattices



minsupp = 85%

For minsupp = 85% the seven most general of the 32.086 concepts of the Mushrooms database <http://kdd.ics.uci.edu> are shown.

The support of a set $X \subseteq M$ of attributes is given by

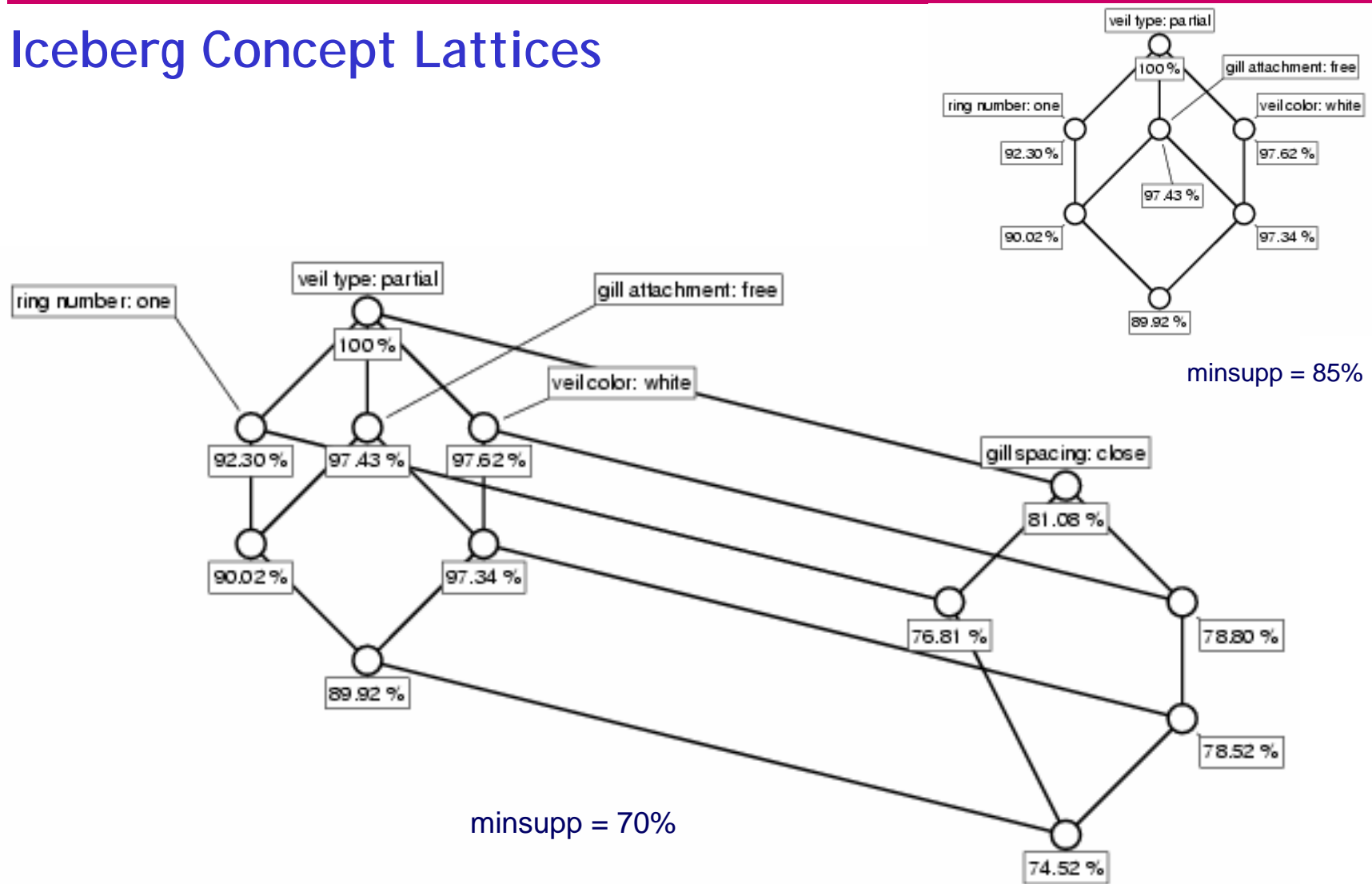
$$\text{supp}(X) = \frac{|X|}{|G|}$$

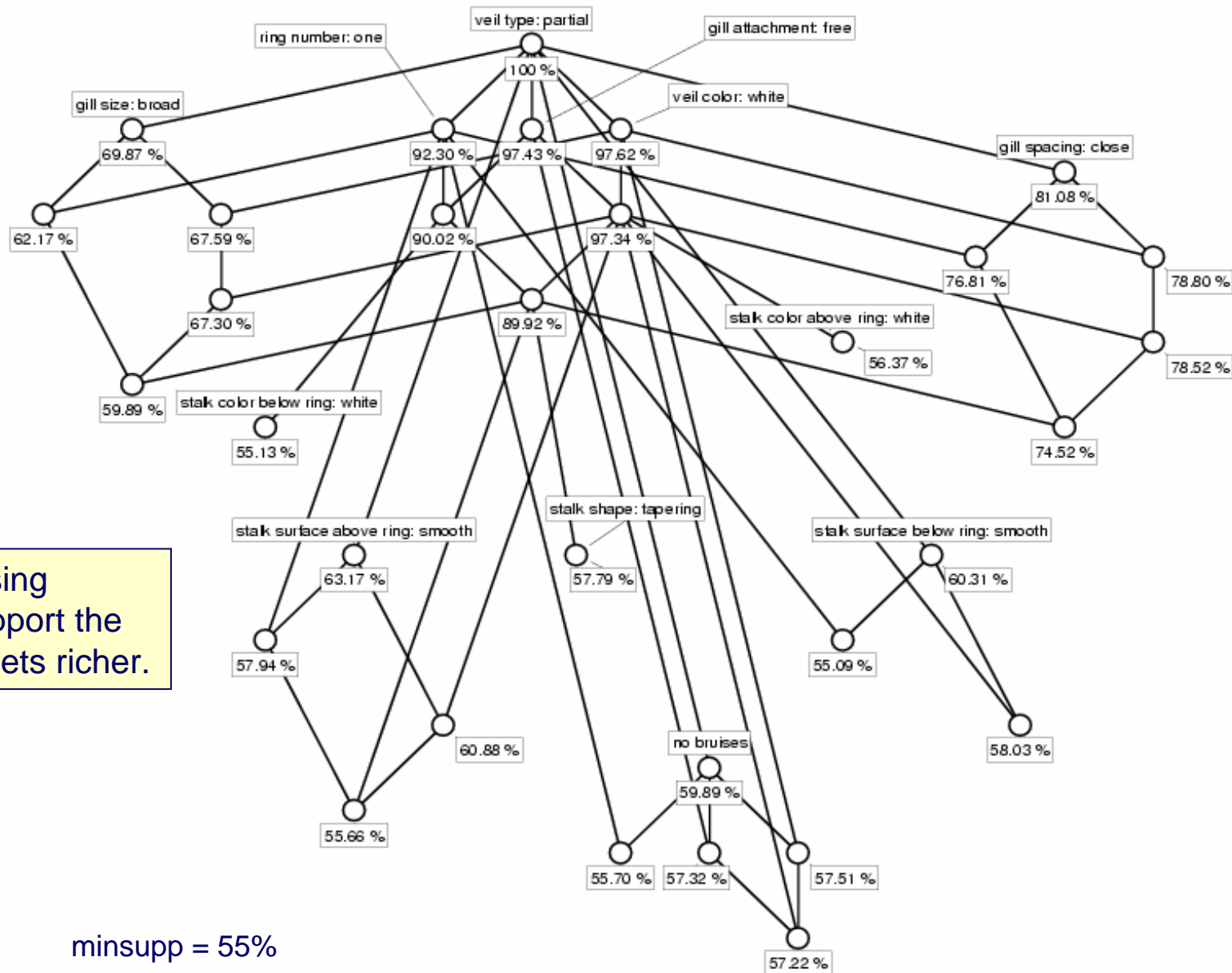
- Def.: The **iceberg concept lattice** of a formal context (G, M, I) for a given minimal support minsupp is the set

$$\{ (A, B) \in \underline{\mathbf{B}}(G, M, I) \mid \text{supp}(B) \geq \text{minsupp} \}$$

- It can be computed with **TITANIC**. [Stumme et al 2001]
(Will be discussed after the Association Rule Section.)

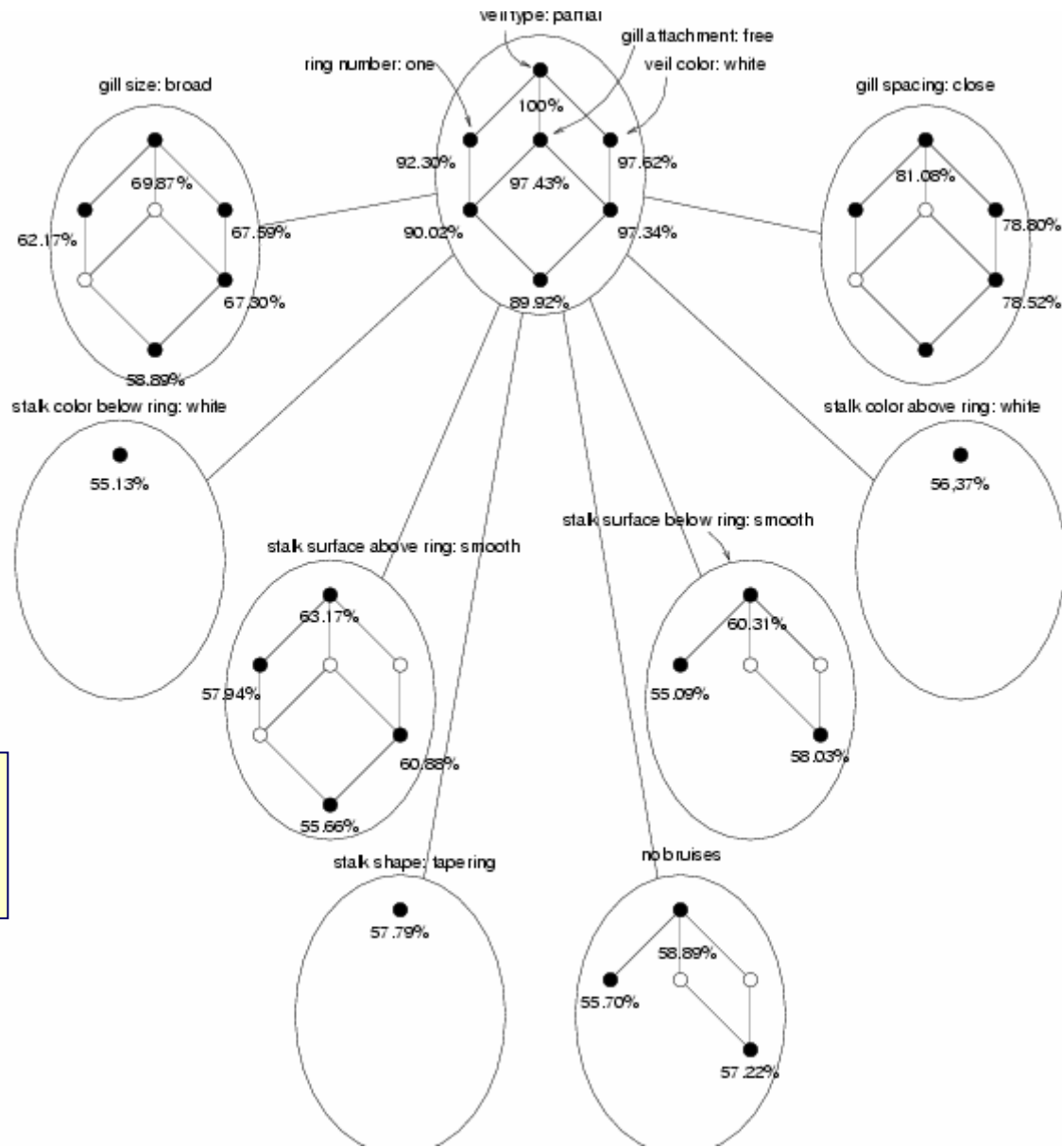
Iceberg Concept Lattices





With decreasing minimum support the information gets richer.

minsupp = 55%



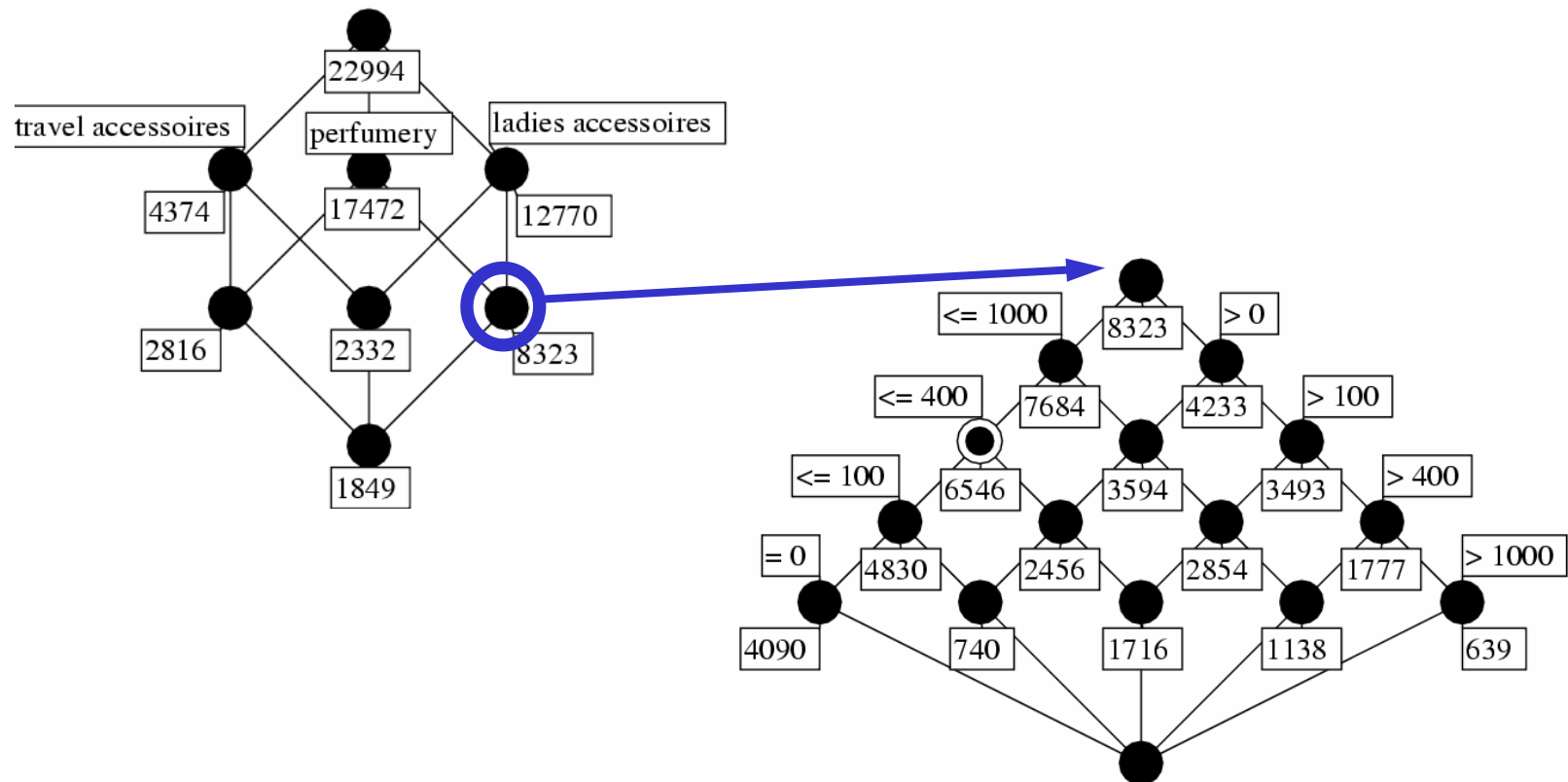
The visualization as a nested line diagram indicates implications.

Einige Anwendungen der Formalen Begriffsanalyse

- Analyse von Diabetes-kranken Kindern
- Entwicklung qualitativer Theorien in Musik-Ästhetik
- Database Marketing in einem Schweizer Warenhaus
- Baurecht in NRW
- Analyse der Flugbewegungen am Frankfurter Flughafen
- IT Sicherheitsmanagement
- Begrifflicher Email Manager

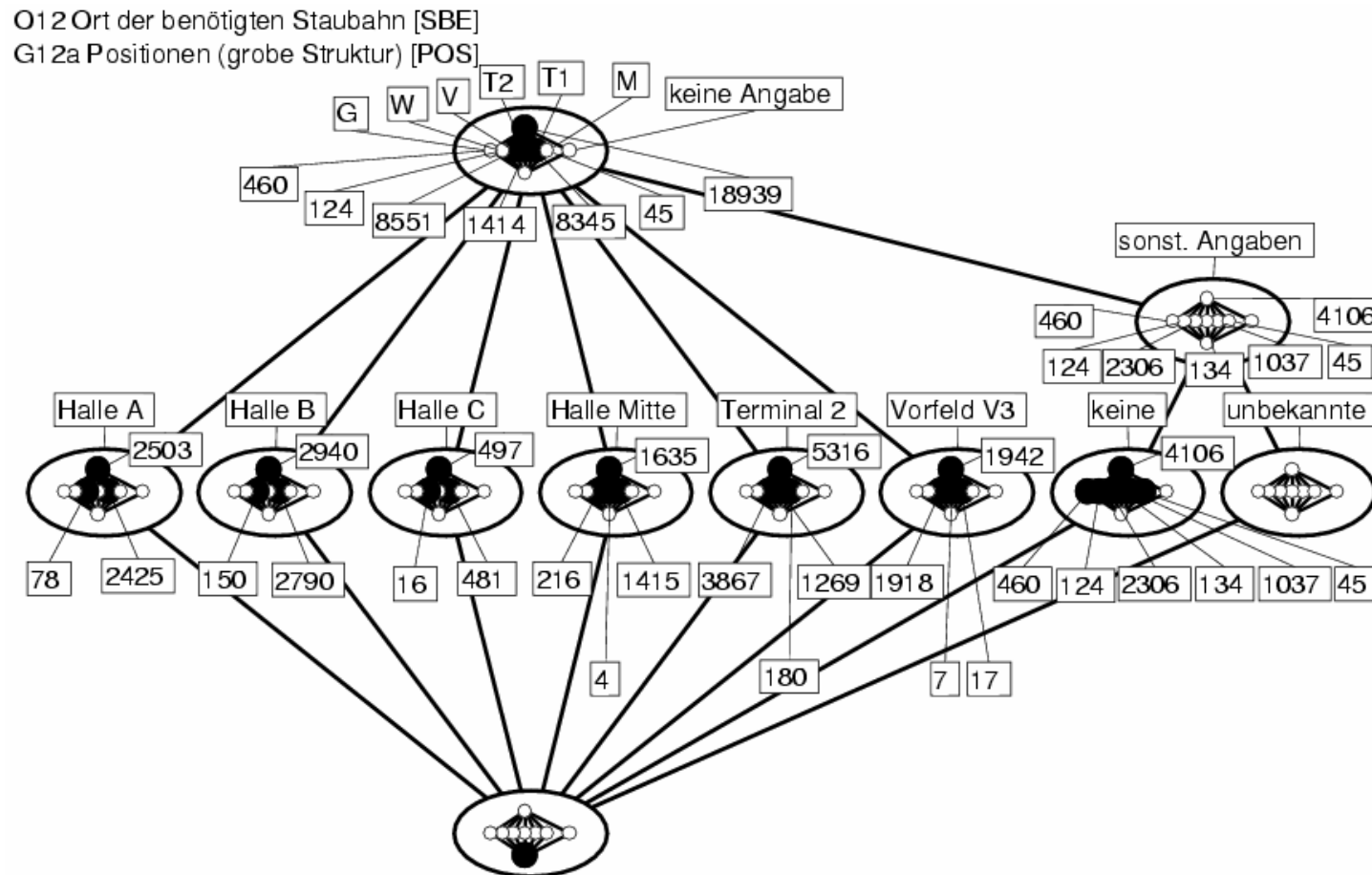
2.5 Einige Anwendungen der Formalen Begriffsanalyse

- Database Marketing bei Jelmoli AG, Zürich



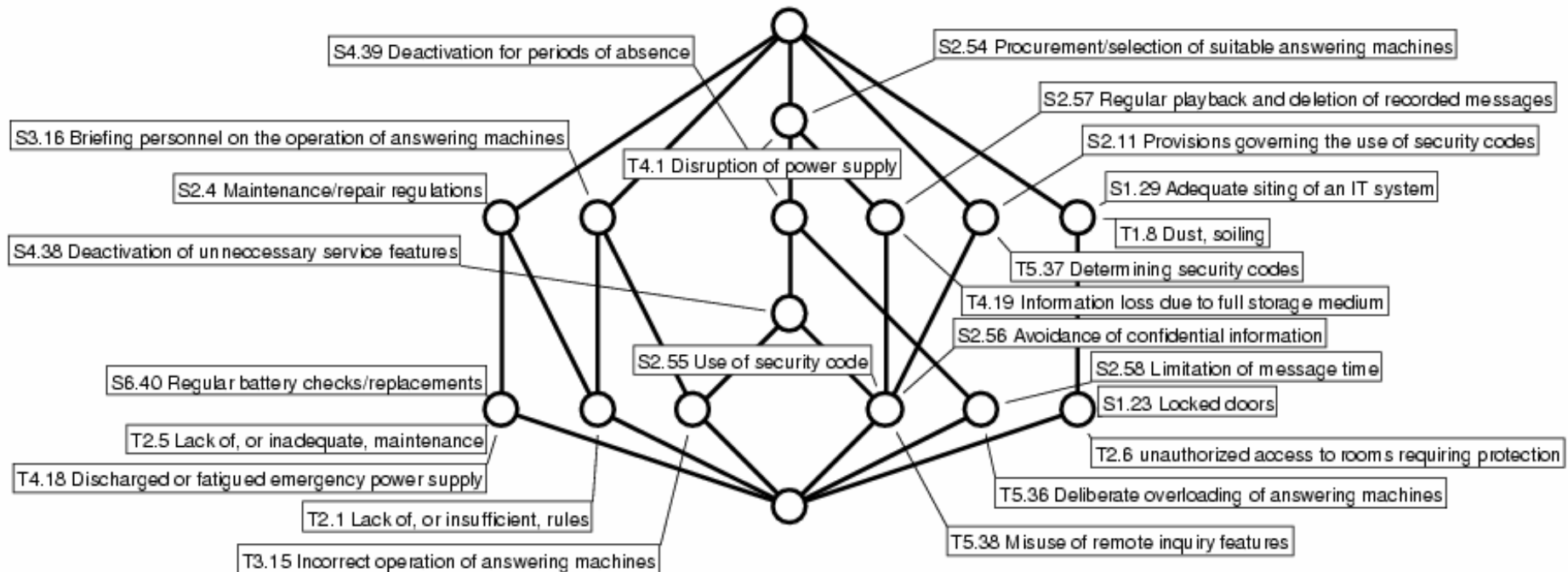
2.5 Einige Anwendungen der Formalen Begriffsanalyse

- Analyse der Flugbewegungen am Flughafen Frankfurt



2.5 Einige Anwendungen der Formalen Begriffsanalyse

- IT-Grundschutzhandbuch



list of referees: ICCS-2000 - ICCS2000 - Netscape-Ordner

Datei Bearbeiten Ansicht Gehe Nachricht Communicator Hilfe

Nachr. abr. Neue Nachr. Antwort Antwort an alle Weiterleiten Ablegen Nächste Drucken Löschen Stop

Name	Ungelesen	Insgesamt
Drafts		
Templates		
Sent		1651
Trash	2	1639
AIFB		94
AUSTRALIA	???	
cole.richard		26
eklund.peter		73
groh.bernd	???	
martin.philippe	???	
CALLFORPAPERS		17
Conferences	1	
ECAI02\Workshop		26
ECML01...orkshop	1	262
mailingaktion	???	
antworten	???	
lesenswert	???	
ICCS2000	1	187
CAMER~\$M.SUM	???	
CAMER~UR	???	
PositionPapers	???	
Software Demos	???	
Vortragende	???	
ICCS2001		333

Betreff	Absender	Datum	Priorität
final notification	Guy Mineau	25.05.2000 16:37	
ICCS2000	Janos Sarbo	26.05.2000 15:20	
Re: Returned mail: Host u...	Alex Borgida	26.05.2000 17:40	
status of all papers	Guy Mineau	29.05.2000 16:50	
expenses covered to go t...	Guy Mineau	29.05.2000 20:22	
Re: Confirmation ICCS2000	Galia Angelova	30.05.2000 08:29	
additional reviewer for ICC...	Harry Delugach	30.05.2000 21:30	
list of referees: ICCS-2000	Guy Mineau	30.05.2000 21:32	
other referees: reminder	Guy Mineau	30.05.2000 21:59	
list of referees: ICCS...	Peter Eklund	31.05.2000 11:...	
Additional reviewers	Ulrike Sattler	31.05.2000 11:46	
Re: List of Referees	Pavel Kocura	31.05.2000 12:40	
Re: ICCS 2000	Deborah L. McGuinness	31.05.2000 20:38	
Please help with accomo...	Guy Mineau	31.05.2000 21:20	

Betreff: list of referees: ICCS-2000
Datum: Wed, 31 May 2000 11:01:11 +0200 (MEST)
Von: [Peter Eklund <Peter.Eklund@sophia.inria.fr>](mailto:Peter.Eklund@sophia.inria.fr)
An: stumme@mathematik.tu-darmstadt.de
CC: ganter@math.tu-dresden.de
Referenzen: 1

Hi Bernhard/Gerd. .

The referees I u

Richard Cole
 Bernd Groh

Nachrichten insgesamt: 187 Ungelesene Nachrichten: 1

Conferences/ICCS2000
 vs.
 AUSTRALIA/eklund.peter

In konventionellen Email-Managern erfolgt
 Abspeicherung der Mails in Baumstruktur
 → nur ein möglicher Suchpfad, der bereits bei
 Abspeicherung festgelegt werden muss

Concept Email Manager

File Lattice View

Folder	Count	✓	✗
From Friends	165	+	+
From Organisation	1878	✓	+
From Griffith Uni	1431	✓	+
From KVO Members	937	+	+
From Darmstadt Group	308	✓	+
From Rudolf Wille	0	+	+
From Jo Hereth	10	+	+
From Gerd Stumme	298	✓	+
from Gerd	298	✓	+
from stumme@	286	+	+
From g.stumme@	12	✓	+
From Darmstadt	46	+	+
From Mailing List	2617	+	✗
CG Mailing List	329	+	✗
To Hermes	2117	+	✗
To Hermes Elec	427	+	✗
To Hermes Chat	693	+	✗
To Hermes Joke	736	+	✗
Text Retrieval List	171	+	✗
Conferences	143	✓	✓
ICCS	114	✓	✓
ICCS 00	26	✓	✓
ICCS Paper with Stumme	1	✓	✓
ICCS 99	7	+	+

From	Subject
Gerd Stumme	Paper
Gerd Stumme	llncs.cls
Gerd Stumme	Paper
Gerd Stumme	Re: [Fwd: Umschlagsentw...

to: "r.cole@gu.edu.au" <r.cole@gu.edu.au>
 <stumme@mathematik.tu-darmstadt.de>
 from: "Gerd Stumme" <g.stumme@gu.edu.au>
 Subject: Paper

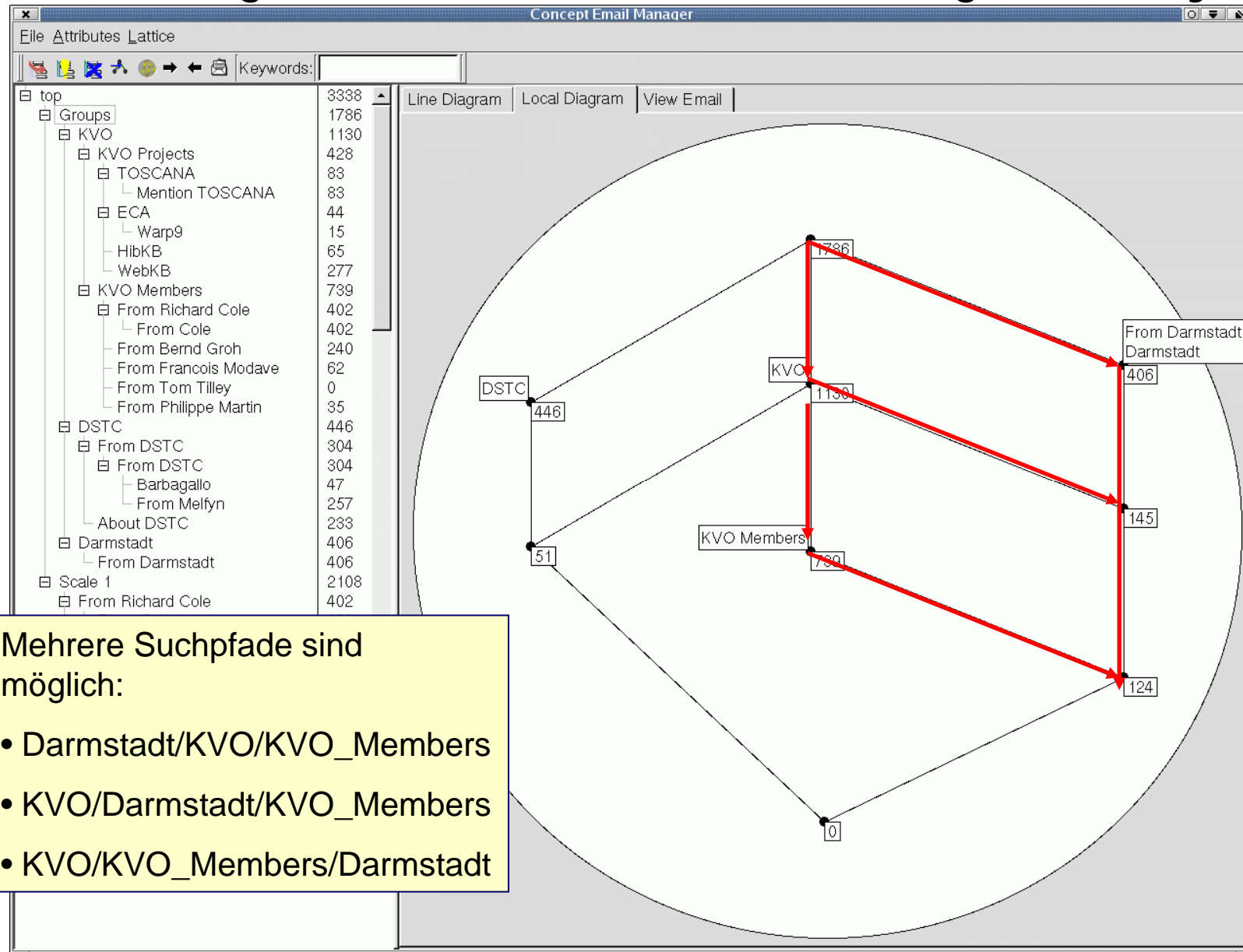
Hi Richard,

here's the Tex-File of our paper. :
 llncs.cls, please have a look at tl
 follow the links to the Springer A

See you at the Sushi place
 Gerd

Im CEM kann eine Email mehreren Schlagworten zugeordnet werden.

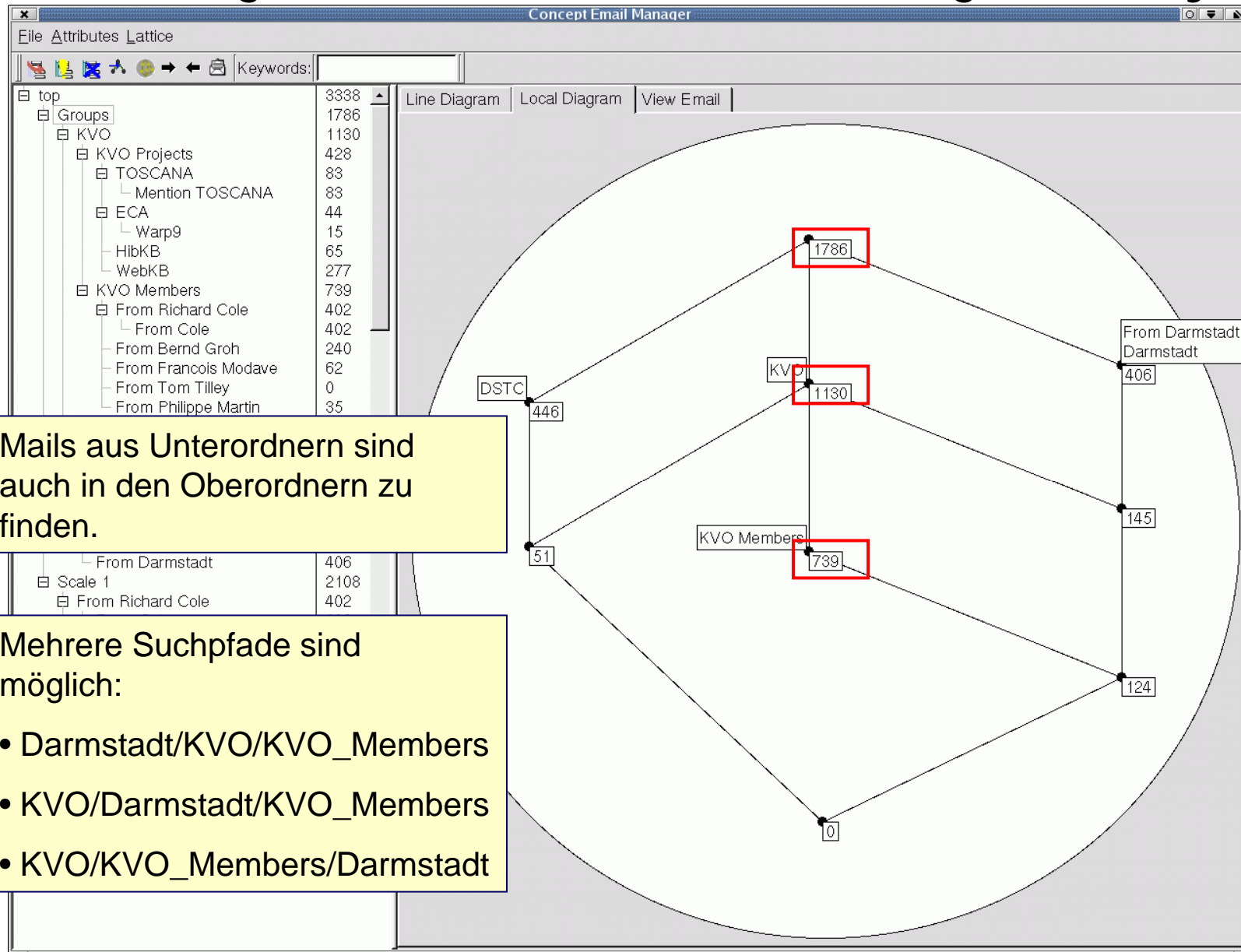
Browsing basierend auf Formaler Begriffsanalyse



Mehrere Suchpfade sind möglich:

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

Browsing basierend auf Formaler Begriffsanalyse



Mails aus Unterordnern sind auch in den Oberordnern zu finden.

- Mehrere Suchpfade sind möglich:
- Darmstadt/KVO/KVO_Members
 - KVO/Darmstadt/KVO_Members
 - KVO/KVO_Members/Darmstadt

Concept Email Manager

File Attributes Lattice

Keywords:

top 3338

- Groups 1786
 - KVO 1130
 - KVO Projects 428
 - TOSCANA 83
 - Mention TOSCANA 83
 - ECA 44
 - Warp9 15
 - HibKB 65
 - WebKB 277
 - KVO Members 739
 - From Richard Cole 402
 - From Cole 402
 - From Bernd Groh 240
 - From Francois Modave 62
 - From Tom Tilley 0
 - From Philippe Martin 35
 - DSTC 446
 - From DSTC 304
 - From DSTC 304
 - Barbagallo 47
 - From Melfyn 257
 - About DSTC 233
 - Darmstadt 406
 - From Darmstadt 406
 - Scale 1 2108
 - From Richard Cole 402
 - From Cole 402
 - EED 3
 - About EED 1
 - Mention EED 3
 - From EED 0
 - eklund 1272

Line Diagram Local Diagram View Email

Verschiedene Sichten können kombiniert werden.