



## VII.5 Formale Begriffsanalyse

© Institut AIFB, 2002.

Alle Rechte vorbehalten. Nachdruck oder photomechanische Wiedergabe nur mit Genehmigung des Verfassers.  
Zuwiderhandlungen unterliegen den strafrechtlichen Bedingungen des Urheberrechtsgesetzes.

## VII.5 Formale Begriffsanalyse

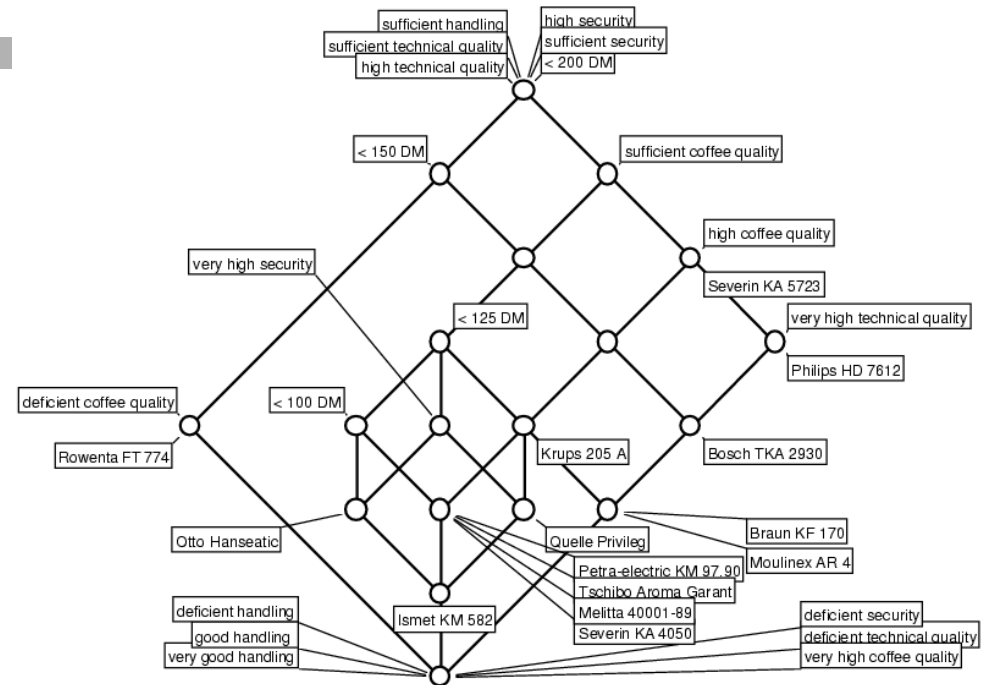
**Formale Begriffsanalyse** ist um 1980 als mathematische Theorie entstanden, die eine Formalisierung des Begriffs vom „Begriff“ liefert.

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 ermöglicht die Erzeugung und die Visualisierung der Begriffshierarchien.



| STIFTUNG WARENTEST<br>test         |                           | 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,- <sup>1)</sup> / □                               | baugl. mit Otto Hanseatic Best.-Nr. 4327357 |                    |            |            | <b>zufriedenst.</b>  |
| Otto Hanseatic Best.-Nr. 4327357   | 40,-                      | 30,- <sup>2)</sup> / □                               | ⊙                                           | +                  | ++         | ⊙          | <b>zufriedenst.</b>  |
| Quelle Privileg Best.-Nr. 7030720  | 40,-                      | 24,50 / 17,50                                        | baugl. mit Otto Hanseatic Best.-Nr. 4327357 |                    |            |            | <b>zufriedenst.</b>  |
| Severin KA 9660                    | 50,-                      | 35,- / 23,-                                          | baugl. mit Otto Hanseatic Best.-Nr. 4327357 |                    |            |            | <b>zufriedenst.</b>  |
| Severin KA 4050                    | 80,-                      | 50,- / □                                             | +                                           | +                  | +          | ⊙          | <b>gut</b>           |
| Tchibo Aroma Garant Art.-Nr. 48469 | 80,-                      | 27,50 / 19,50                                        | +                                           | +                  | +          | ⊙          | <b>gut</b>           |
| Ismet KM 582 starlight             | 84,-                      | 47,- / 14,-                                          | +                                           | +                  | ++         | ⊙          | <b>gut</b>           |

## VII.5 Formale Begriffsanalyse

### Formal Concept Analysis

[Wille 1982]

- FCA models concepts as units of thought, consisting of two parts:
  - The **extension** consists of all objects belonging to the concept.
  - The **intension** consists of all attributes common to all those objects.
- FCA is used for data analysis, information retrieval, and knowledge discovery.
- FCA can be understood as conceptual clustering method, which clusters simultaneously objects and their descriptions.
- FCA can be used for efficiently computing association rules.

## VII.5 Formale Begriffsanalyse

### VII.5.2 Basics of Formal Concept Analysis

- In its basic version, FCA handles object-attribute pairs.
- Object-attribute-value triples can also be handled, but this is not topic of this course.
- **Def.:** A **(formal) context** is a triple  $(G, M, I)$  where  $G$  and  $M$  are sets and  $I$  is a binary relation between  $G$  and  $M$ .
  - The elements of  $G$  are called objects, and the elements of  $M$  are called attributes.
  - $(g, m) \in I$  is read „object  $g$  has attribute  $m$ “.

Entspricht Items/Transaktionen bei Assoziationsregeln

## VII.5 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<br>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                         | ×                   | ×      | ×                | ×        | ×       | ×       | ×             | ×                   |

## VII.5 Formale Begriffsanalyse

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<br>in California    | A' |   |   |          |         |   |                                      |
|------------------------------------|----|---|---|----------|---------|---|--------------------------------------|
|                                    |    |   |   | Swimming | Boating |   | Bicycle Trail<br>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 |                                      |
|                                    |    |   |   |          |         |   | x                                    |
| Yosemite Natl. Park                | x  | x | x | x        | x       | x | x                                    |

## VII.5 Formale Begriffsanalyse

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

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

B

| National Parks<br>in California | B        |         |         |               |                     |
|---------------------------------|----------|---------|---------|---------------|---------------------|
|                                 | Swimming | Boating | Fishing | Bicycle Trail | Cross Country Trail |
| Cabrillo Natl. Mon.             |          |         |         | x             | x                   |
| Channel Islands Natl. Park      |          | x       |         | x             |                     |
| [Redacted]                      |          | x       |         |               | x                   |
| Fort Point Natl. Historic Site  |          | x       |         | x             |                     |
| [Redacted]                      |          | x       |         | x             | x                   |
| John Muir Natl. Historic Site   |          | x       |         |               |                     |
| [Redacted]                      |          |         |         | x             | x                   |
| [Redacted]                      |          | x       | x       | x             | x                   |
| Lava Beds Natl. Mon.            |          | x       | x       |               |                     |
| Muir Woods Natl. Mon.           |          |         | x       |               |                     |
| Pinnacles Natl. Mon.            |          |         | x       |               |                     |
| [Redacted]                      |          | x       |         | x             |                     |
| [Redacted]                      |          | x       |         | x             |                     |
| [Redacted]                      |          | x       | x       |               |                     |
| [Redacted]                      |          |         | x       |               | x                   |
| [Redacted]                      |          | x       | x       |               |                     |
| [Redacted]                      |          | x       | x       | x             | x                   |

B'

## VII.5 Formale Begriffsanalyse

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:

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

A

| National Parks<br>in California    | A' |   |   |   |          |         |   |               |                     |
|------------------------------------|----|---|---|---|----------|---------|---|---------------|---------------------|
|                                    |    |   |   |   | Swimming | Boating |   | 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 |               | 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       | x |               |                     |
|                                    |    |   |   |   |          |         |   |               | x                   |
| Yosemite Natl. Park                | x  | x | x | x | x        | x       | x | x             | x                   |



## VII.5 Formale Begriffsanalyse

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<br>in California | Inhalt |   |   |          |         |   |               |                     |
|---------------------------------|--------|---|---|----------|---------|---|---------------|---------------------|
|                                 |        |   |   | Swimming | Boating |   | Bicycle Trail | Cross Country Trail |
| Cabrillo Natl. Mon.             |        |   |   |          |         | × | ×             |                     |
| Channel Islands Natl. Park      |        | × |   | ×        |         | × |               |                     |
| Death Valley Natl. Mon.         | ×      | × | × | ×        |         |   | ×             |                     |
|                                 |        |   |   | ×        |         |   |               |                     |
| Fort Point Natl. Historic Site  | ×      |   |   |          |         | × |               |                     |
|                                 |        |   |   | ×        |         | × | ×             |                     |
| John Muir Natl. Historic Site   | ×      |   |   |          |         |   |               |                     |
| Joshua Tree Natl. Mon.          | ×      | × | × |          |         |   |               |                     |
|                                 |        |   |   |          |         |   |               | ×                   |
|                                 |        |   |   | ×        | ×       |   |               | ×                   |
| Lava Beds Natl. Mon.            | ×      | × |   |          |         |   |               |                     |
| Muir Woods Natl. Mon.           |        | × |   |          |         |   |               |                     |
| Pinnacles Natl. Mon.            |        | × |   |          |         |   |               |                     |
|                                 |        |   |   | ×        |         |   | ×             |                     |
|                                 |        |   |   | ×        |         |   |               |                     |
|                                 |        |   |   | ×        | ×       |   |               | ×                   |
|                                 |        |   |   | ×        | ×       |   | ×             | ×                   |

## VII.5 Formale Begriffsanalyse

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

der blaue Umfang ist im gelben Umfang enthalten.

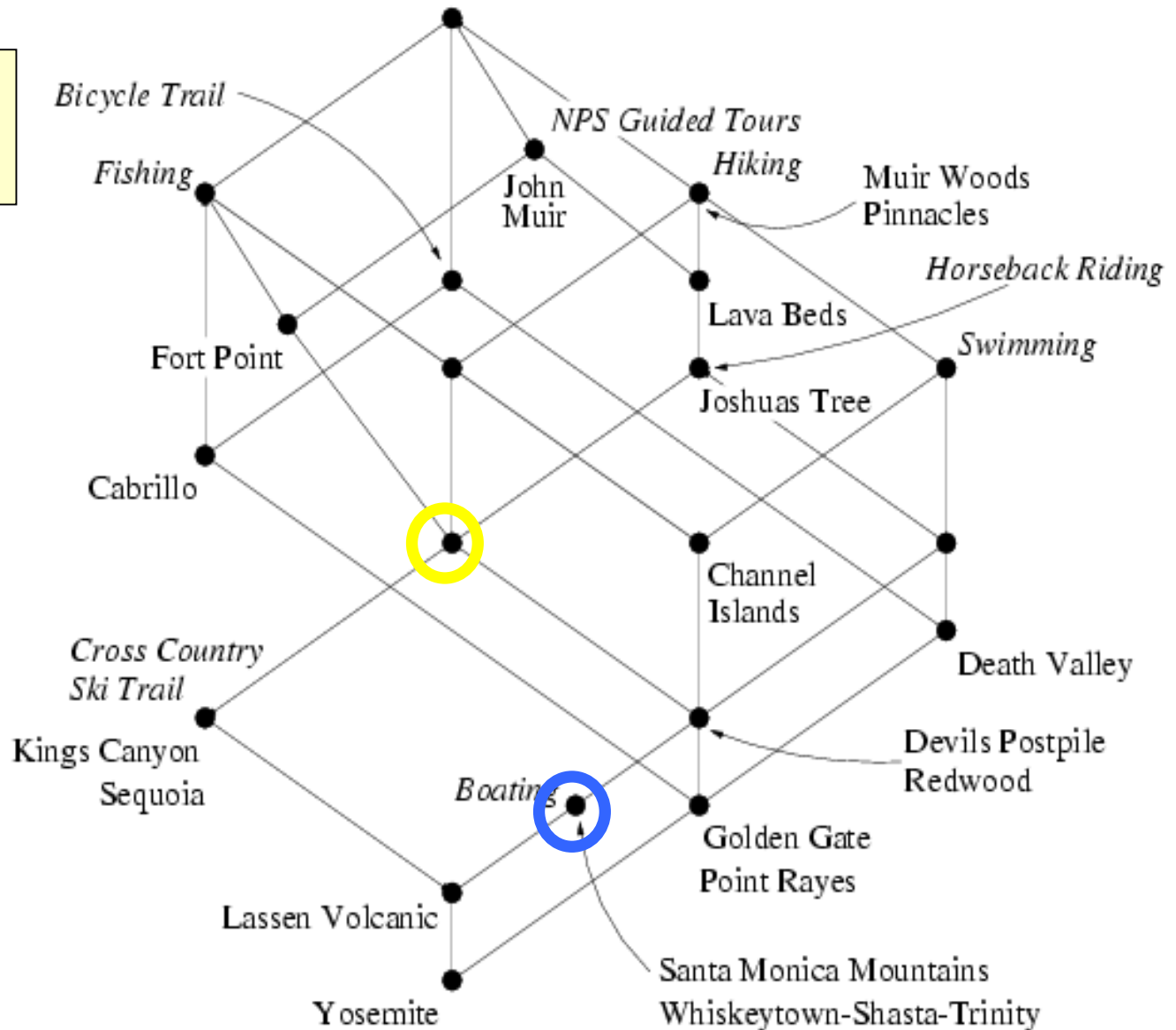
( $\Leftrightarrow$  der gelbe Inhalt ist im blauen Inhalt enthalten.)

$(A, B) \leq (C, D) : \Leftrightarrow A \subseteq C$   
 $(\Leftrightarrow B \supseteq D)$

| National Parks in California   |   |   |   |   |  | 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                   |
| Fort Point Natl. Historic Site | x |   |   |   |  | x             |                     |
| John Muir Natl. Historic Site  |   |   |   | x |  |               | x                   |
| Joshua Tree Natl. Mon.         | x | x | x |   |  |               |                     |
| Lava Beds Natl. Mon.           | x | x |   |   |  |               |                     |
| Muir Woods Natl. Mon.          |   | x |   |   |  |               |                     |
| Pinnacles Natl. Mon.           |   | x |   |   |  |               |                     |
|                                |   |   |   | x |  |               | x                   |
|                                |   |   |   | x |  |               |                     |
|                                |   |   |   |   |  |               | x                   |
|                                |   |   |   |   |  | x             | x                   |

# VII.5 Formale Begriffsanalyse

Der **Begriffsverband** zu dem Nationalpark-Kontext



| National Parks in California      | Bicycle Trail | Cross Country Trail |
|-----------------------------------|---------------|---------------------|
| Cabrillo Natl. Mon.               |               | x                   |
| Channel Islands Natl. Park        |               | x                   |
| Death Valley Natl. Mon.           | x             | x                   |
| Devils Postpile Natl. Mon.        | x             | x                   |
| Fort Point Natl. Historic Site    | x             | x                   |
| Golden Gate Natl. Recreation Area | x             | x                   |
| John Muir Natl. Historic Site     | x             |                     |
| Joshua Tree Natl. Mon.            | x             | x                   |
| Kings Canyon Natl. Park           | x             | x                   |
| Lava Beds Natl. Mon.              | x             |                     |
| Muir Woods Natl. Mon.             | x             |                     |
| Pinnacles Natl. Mon.              | x             |                     |
| Point Reyes Natl. Seashore        | x             | x                   |
| Redwood Natl. Park                | x             | x                   |
| Sequoia Natl. Park                | x             | x                   |

## VII.5 Formale Begriffsanalyse

- **Def.:** The **concept lattice [Begriffsverband]** of a formal context  $(G, M, I)$  is the set of all formal concepts of  $(G, M, I)$ , together with the partial order

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

The concept lattice is denoted by  $\underline{\mathcal{B}}(G, M, I)$  .

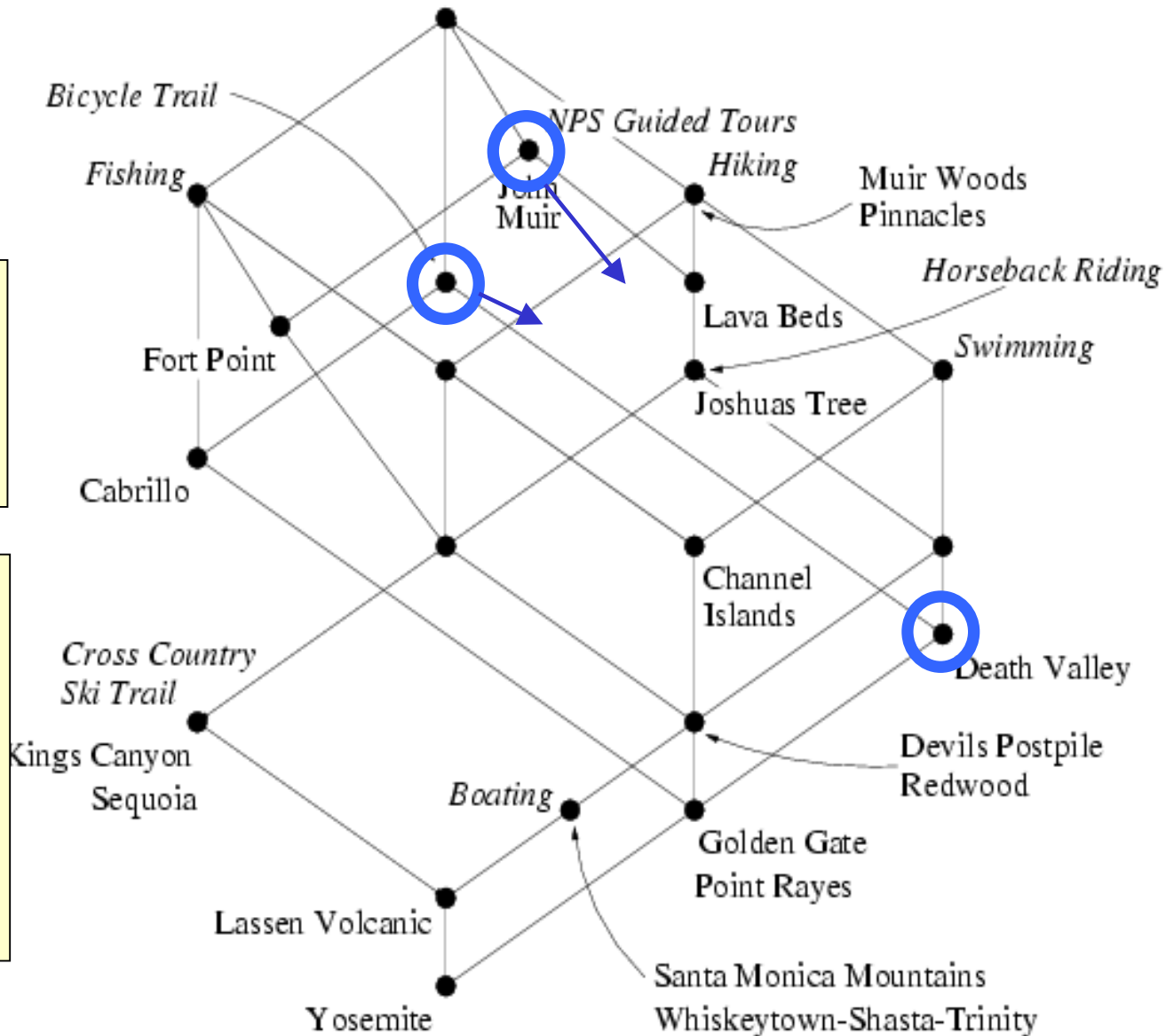
- **Theorem:** The concept lattice is a lattice, i.e. for two concepts  $(A_1, B_1)$  and  $(A_2, B_2)$ , there is always
  - a greatest common subconcept:  $(A_1 \cap A_2, (B_1 \cup B_2)'' )$
  - and a least common superconcept:  $((A_1 \cup A_2)'' , B_1 \cap B_2)$  .

## VII.5 Formale Begriffsanalyse

„Welche Gegenstände haben sowohl das Merkmal ‚Bicycle Trail‘ als auch ‚NPS Guided Tours‘?“

„Welche Merkmale haben diese Gegenstände noch?“

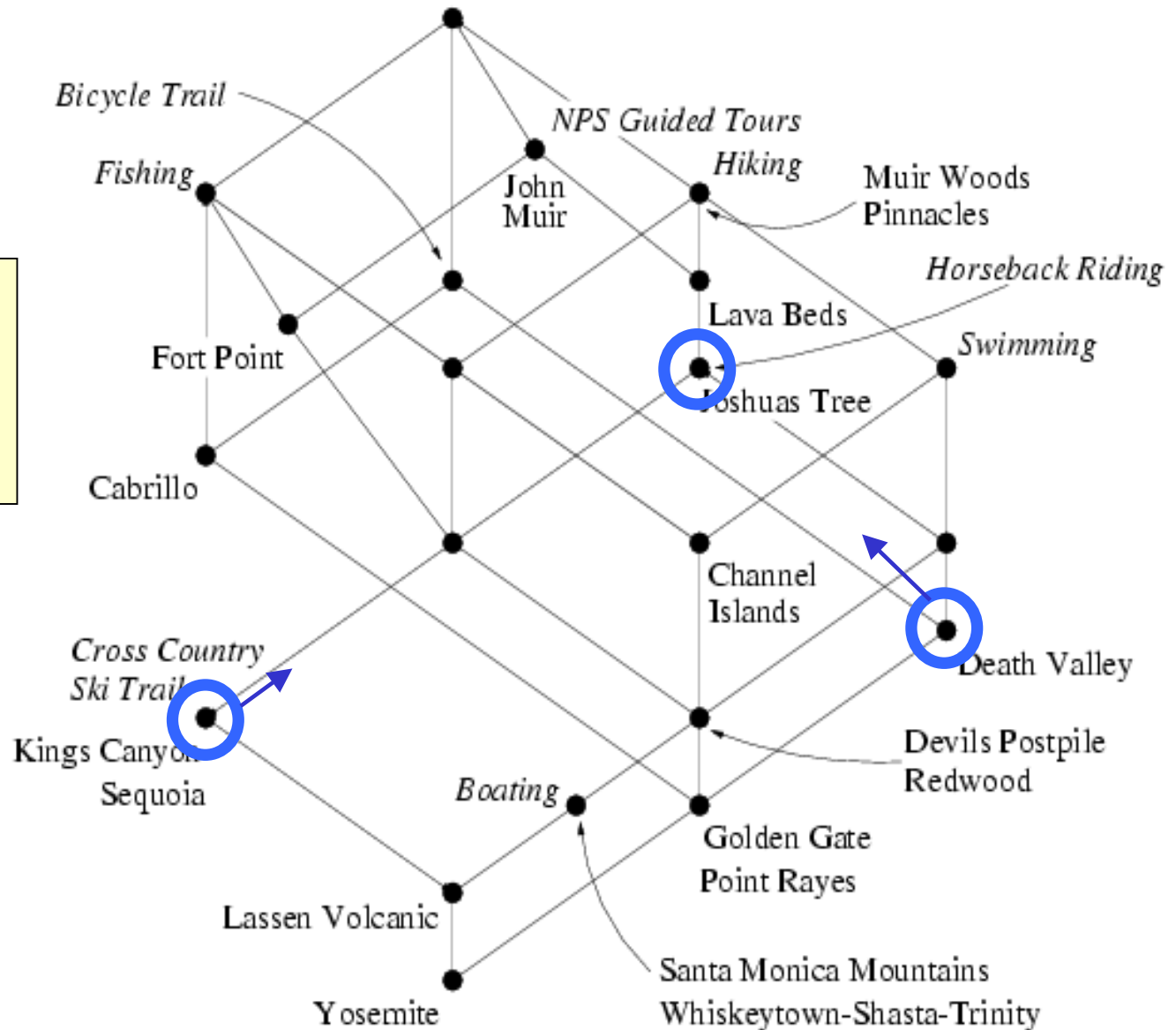
In anderen Worten:  
„Welche Merkmale folgen noch aus ‚Bicycle Trail‘ und ‚NPS Guided Tours‘?“



## VII.5 Formale Begriffsanalyse

„Welche Merkmale teilen sich die Gegenstände ‚Kings Canyon‘ und ‚Death Valley‘?“

„Welche Gegenstände haben noch diese Merkmale?“



## VII.5 Formale Begriffsanalyse

### VII.5.3 Formal Concept Analysis as Conceptual Clustering Method

- **Conceptual Clustering** methods are clustering methods which generate simultaneously descriptions of the clusters.
- Advantages of conceptual clustering against clustering as in Sect. VII.1:
  - A cluster is not only a set of objects, but there also exists an intensional description.
  - For FCA: The results do not depend on the order of the input
- Disadvantages:
  - The language used to describe the clusters restricts the type of clusters which can be built.
  - The computation has usually higher complexity.
- Other methods: Michalski & Stepp 1983; Lebowitz 1987; Fisher 1987; Gennari et al 1989

## VII.5 Formale Begriffsanalyse

**Iceberg concept lattices** only allow conjunctions of attributes as descriptions.

- In the notion of Formal Concept Analysis, the support of an itemset  $X \subseteq M$  can be written as

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

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

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

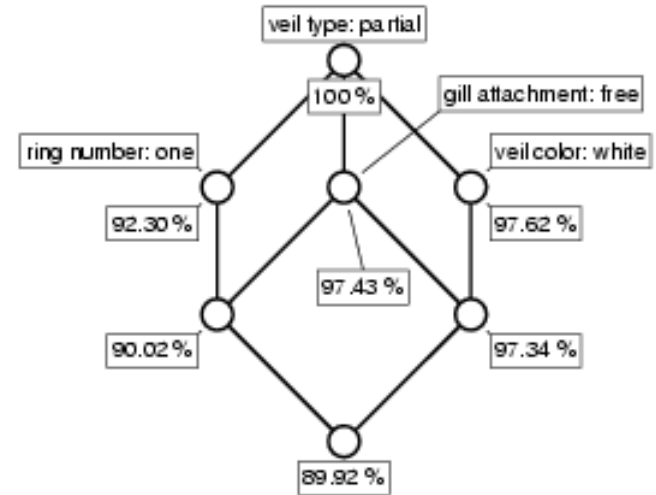
- It can be computed with the algorithm **TITANIC**. [Stumme et al 2001]



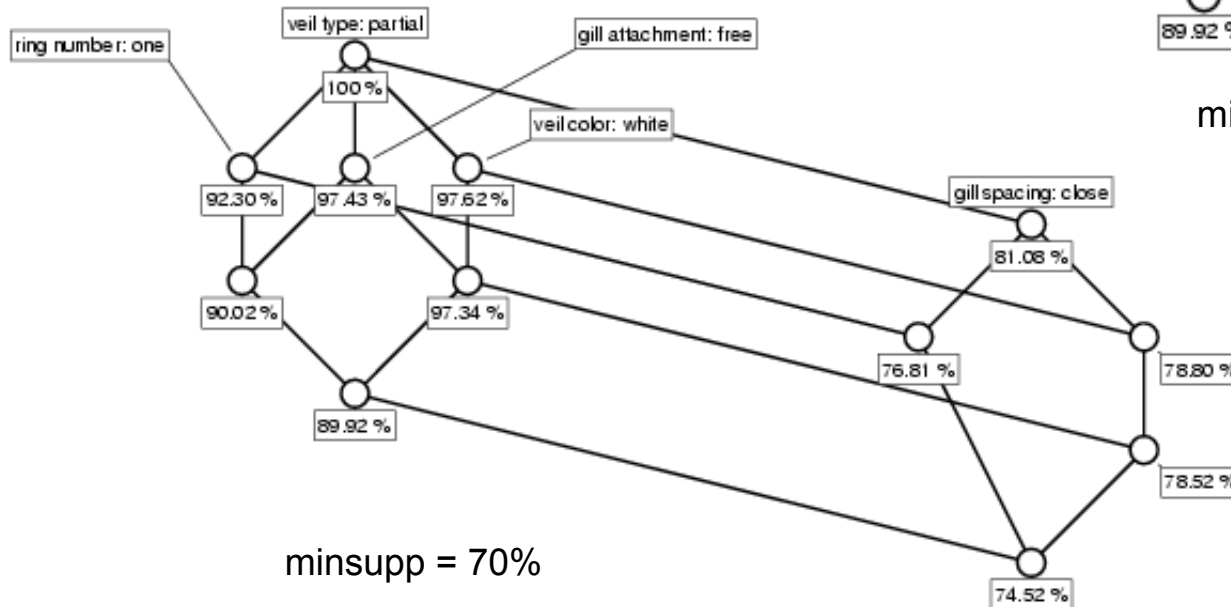
## VII.5 Formale Begriffsanalyse

### Example:

Iceberg Concept Lattices of the  
Mushrooms Database  
(<http://kdd.ics.uci.edu>).

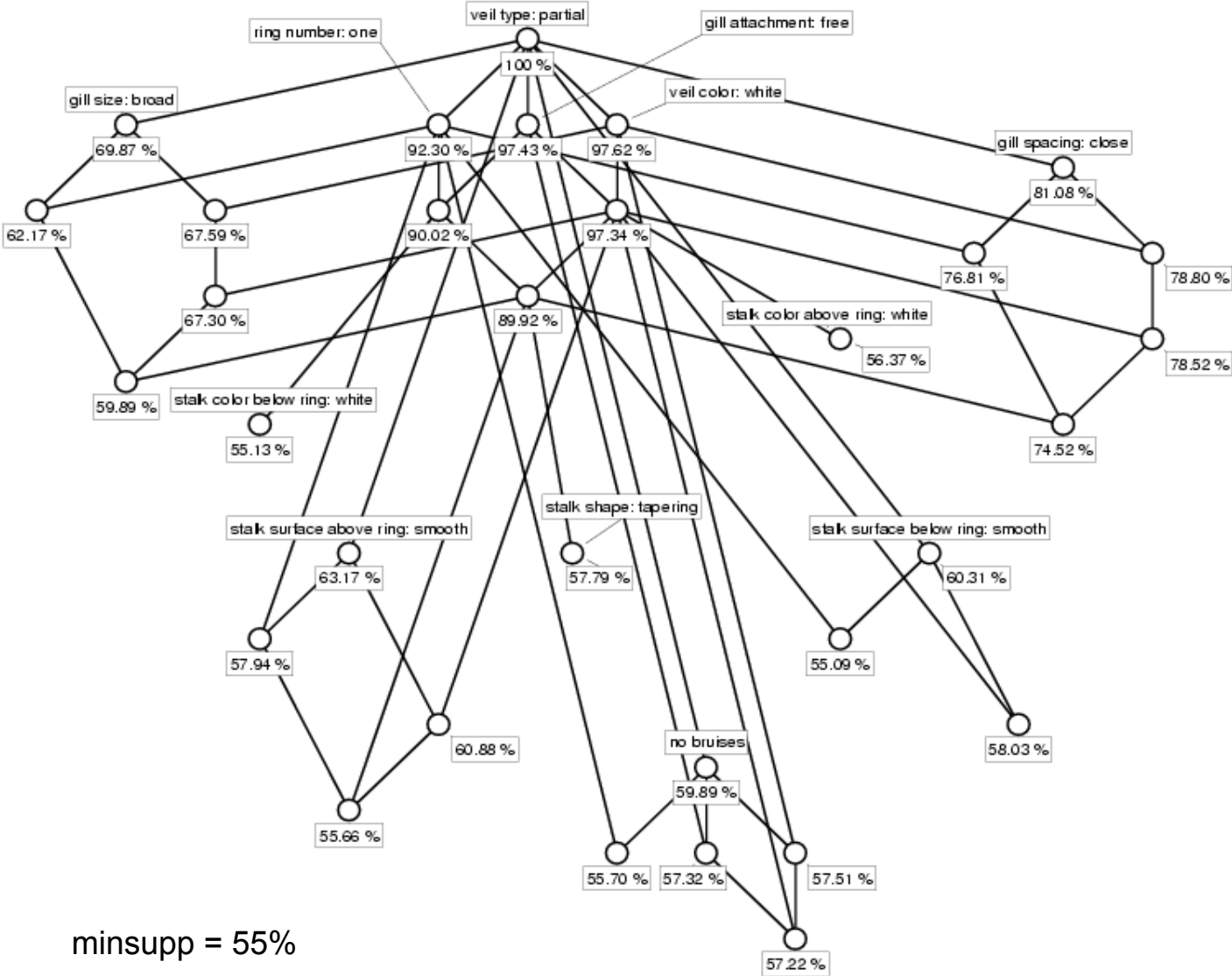


minsupp = 85%



minsupp = 70%

# VII.5 Formale Begriffsanalyse



## VII.5 Formale Begriffsanalyse

### VII.5.4 Computation of (Iceberg) Concept Lattices

- There exist a number of algorithms for computing concept lattices
  - Next-Closure [Ganter 1984]
  - Titanic [Stumme et al 2001]
  - and some incremental algorithms
- The following method is also suitable for manual computation. [Wille 1982]
- It provides the best worst-case time complexity. [Nourine, Raynoud 1999]

↳ Example „Faces“ on the Blackboard

## VII.5 Formale Begriffsanalyse

### How to compute/draw a concept lattice:

- From left to right, consider all intersections of each column extent with every column extent to the left of it. If the resulting extent is not already a column, add it as column at the right end of the context. Repeat this until the last (added) column is reached.
- Add a full column, unless there is already one. (Now each column stands for one concept.)
- Draw a circle for the full column.
- Draw for each column, starting for the ones with a maximal number of crosses, a circle, and link it with a line to the circles where the column comprises the current column.
- Attach every attribute label to the circle of the corresponding column.
- Attach every object label to the circle laying exactly below the circles of the attributes in its intent.

## VII.5 Formale Begriffsanalyse

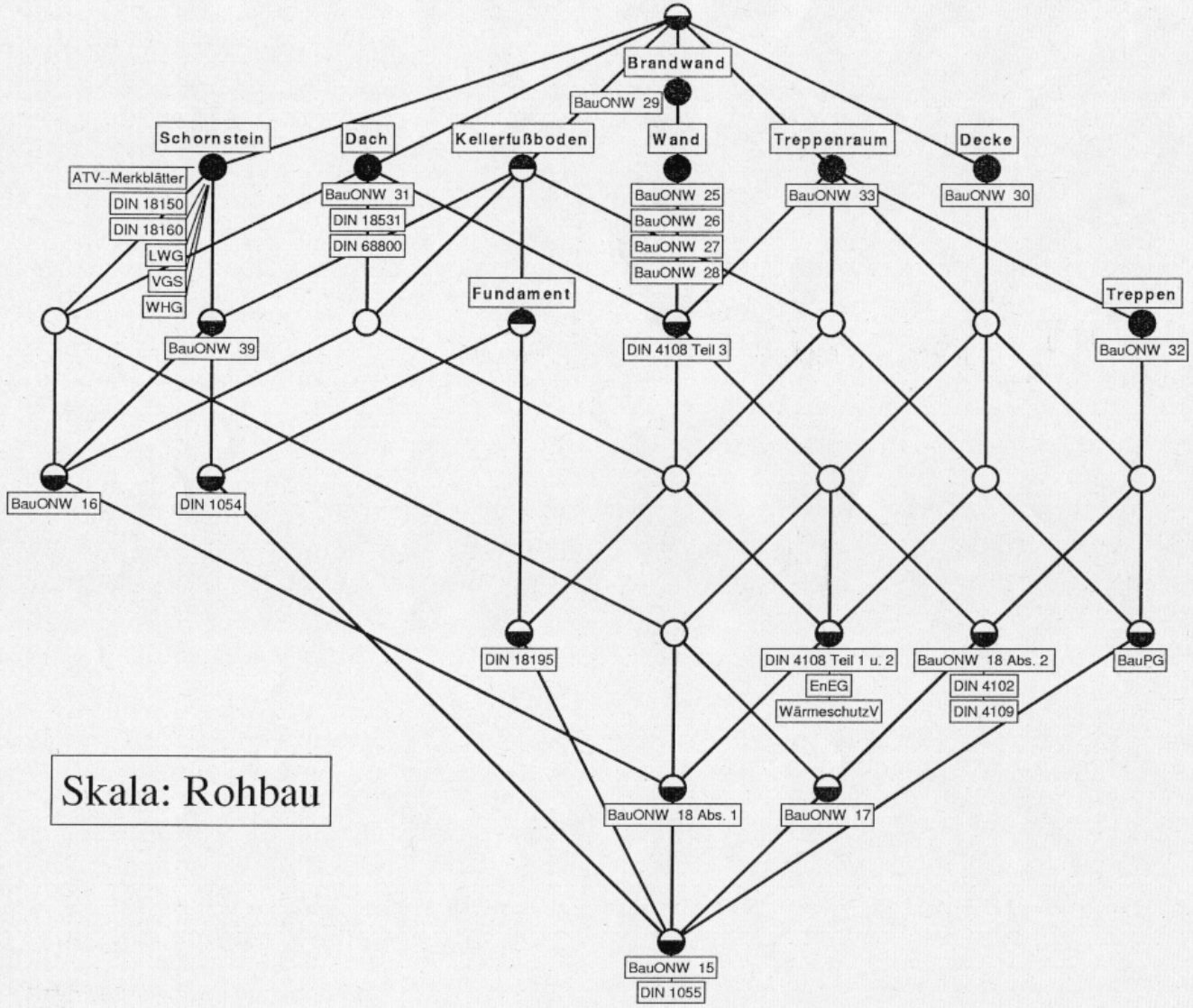
### How to check the drawing of a concept lattice:

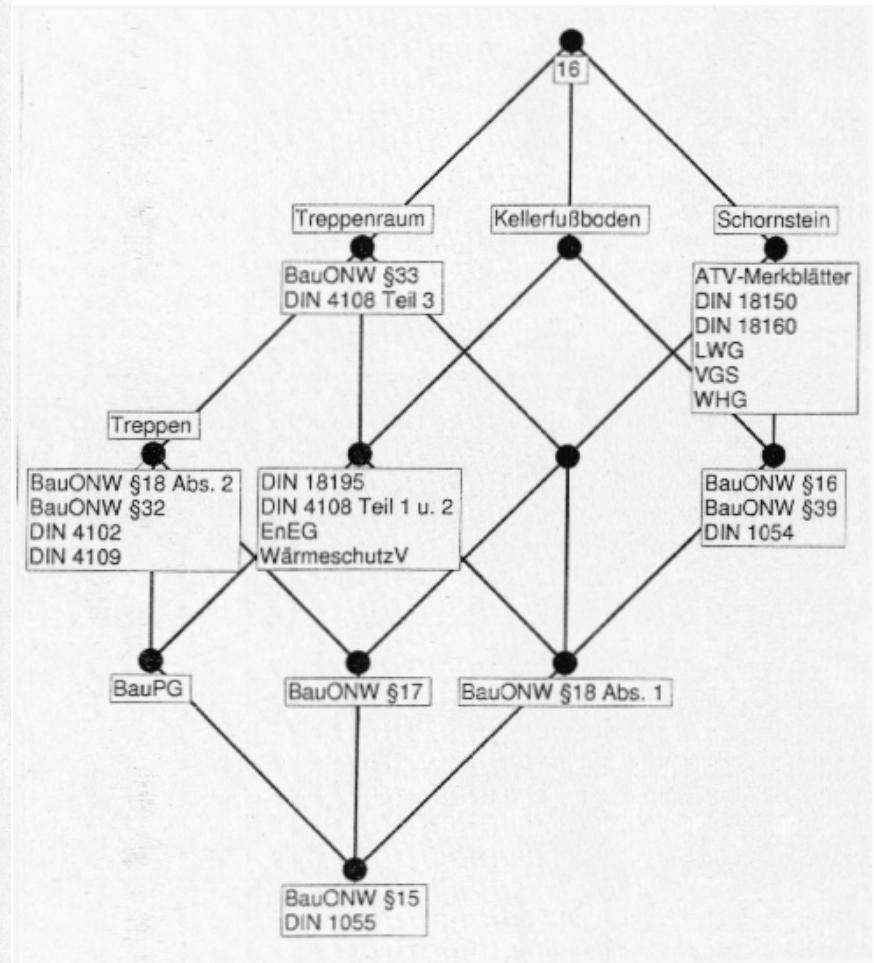
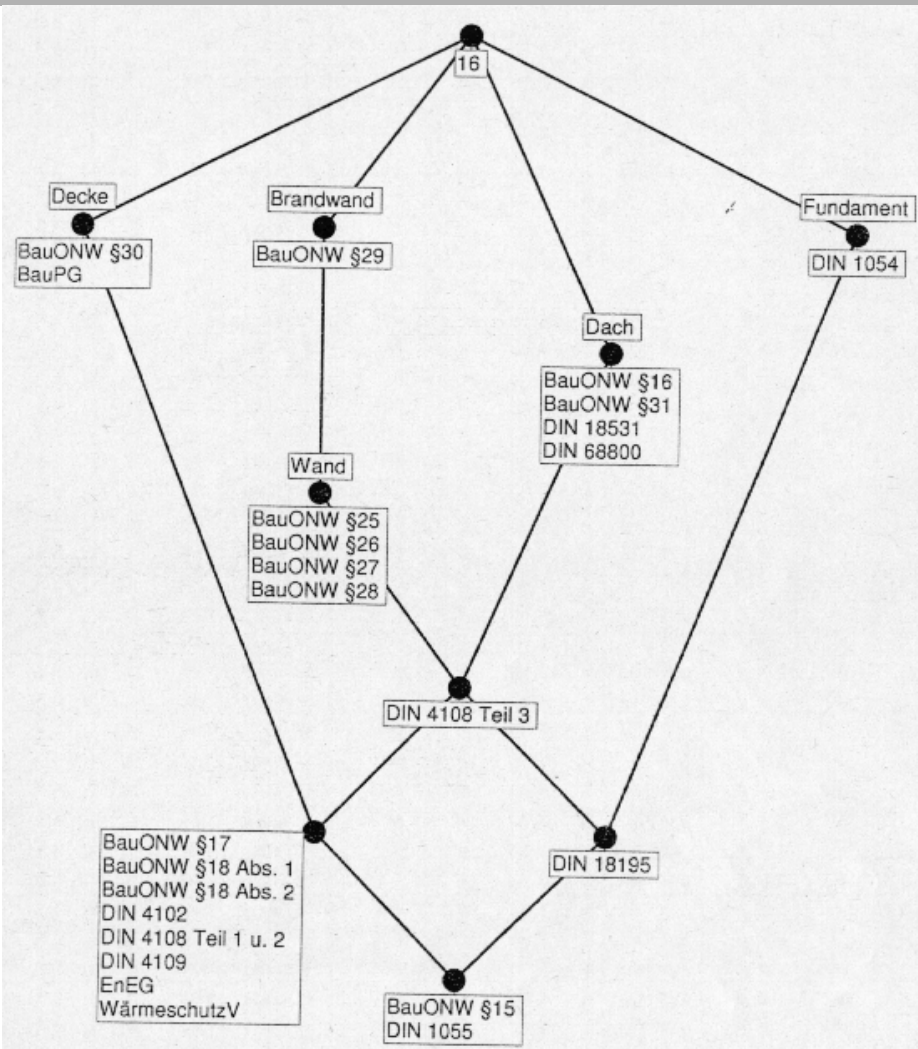
- Is it really a lattice? (This test is usually skipped.)
- Is every concept with exactly one upper neighbor labeled by at least one attribute?
- Is every concept with exactly one lower neighbor labeled by at least one object?
- Is, for all  $g \in G$  and all  $m \in M$ , the label of object  $g$  below the label of attribute  $m$  iff  $(g,m) \in I$  ?

- Another method for reducing the complexity of the diagram is conceptual scaling.
- In our case (only Boolean attributes), this means selecting only subsets of the attribute set.
- The resulting concept lattices are smaller.
- If combinations are of interest, they can be put together again.
- Conceptual scaling can be combined with Iceberg concept lattices.

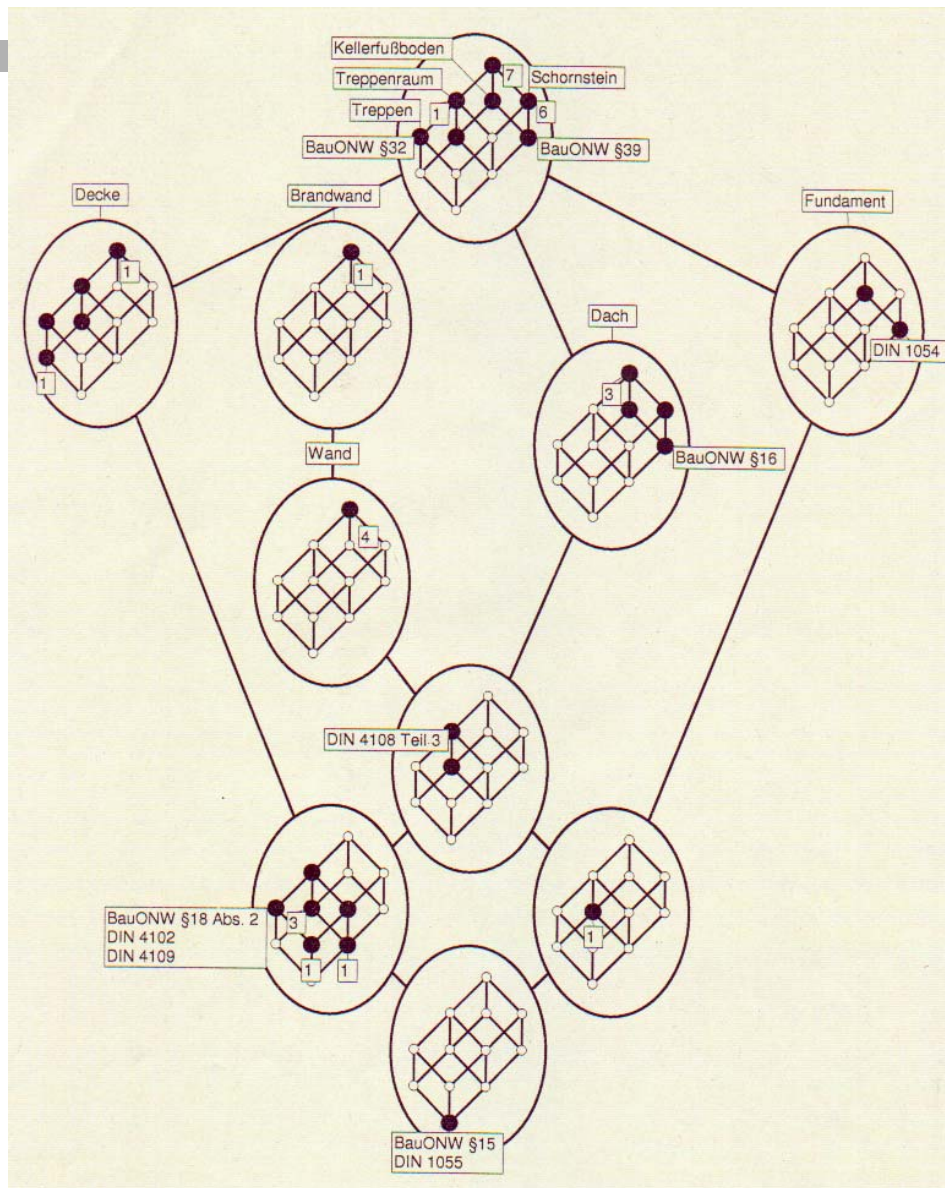
**Beispiel:** Baurecht in Nordrhein-Westfalen

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









## VII.5 Formale Begriffsanalyse

- Some **typical applications** of FCA:
  - analysis of children suffering from diabetes
  - IT security management system
  - database marketing in a Swiss department store
  - **Conceptual Email Management system**
  - developing qualitative theories in music esthetics
  - analysis of flight movements at Frankfurt airport

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: Cofirmation 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         |           |
| <b>list of referees: ICCS...</b> | <b>Peter Eklund</b>   | <b>31.05.2000 11:...</b> |           |
| 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](mailto:stumme@mathematik.tu-darmstadt.de)  
**CC:** [ganter@math.tu-dresden.de](mailto:ganter@math.tu-dresden.de)

**Referenzen:** 1

Hi Bernhard/Gerd. .

The referees I u

Richard Cole  
 Bernd Groh

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

Im CEM kann eine Email mehreren Schlagworten zugeordnet werden.

| From                   | 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         | 893   | + | ✗ |
| 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... |

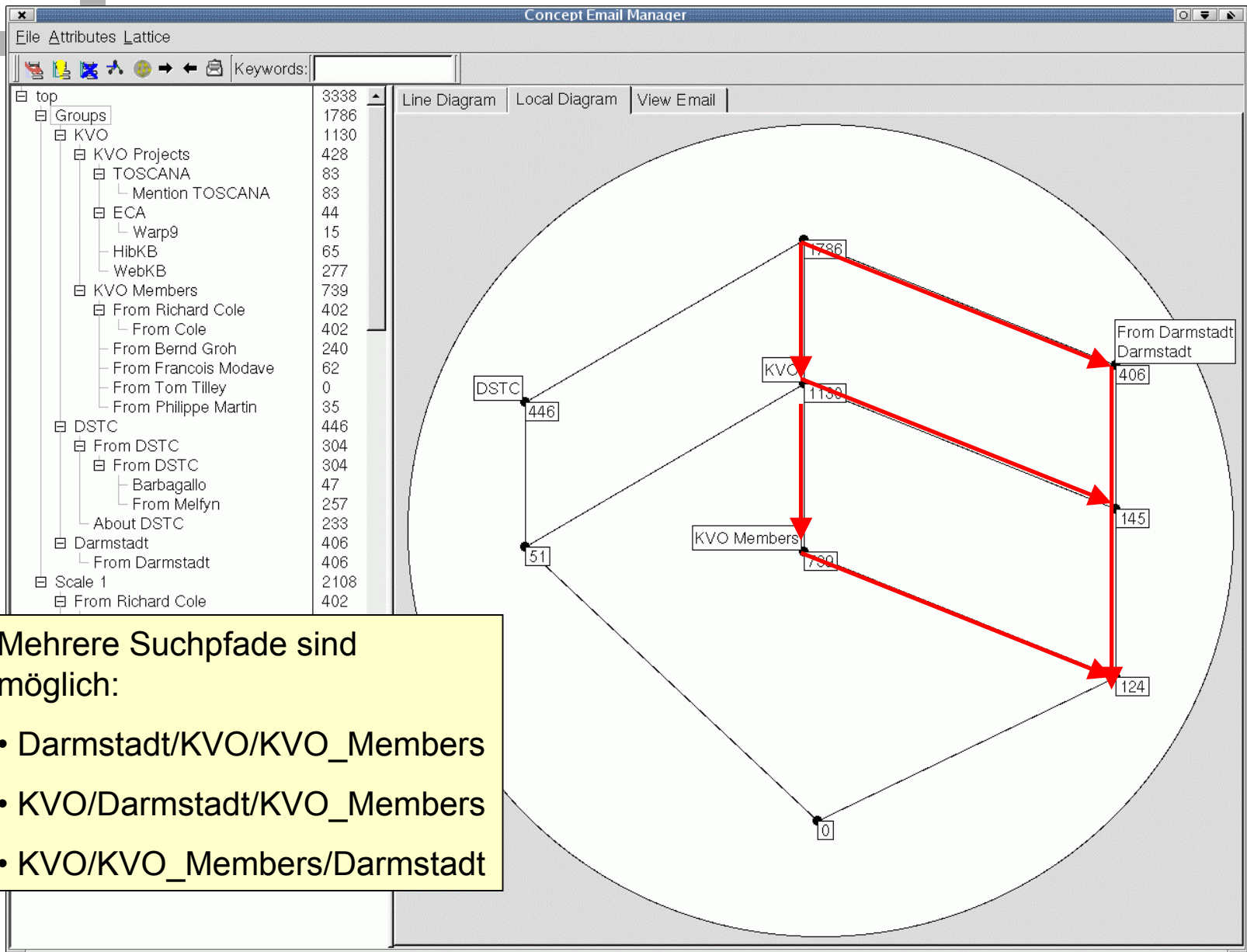
to: "r.cole@gu.edu.au" <r.cole@gu.edu.au>  
 <stumme@mathematik.tu-darmstadt.de>  
 from: "Gerd Stumme" <g.stumme@gu.edu.au>  
 Subject: Paper

Hi Richard,

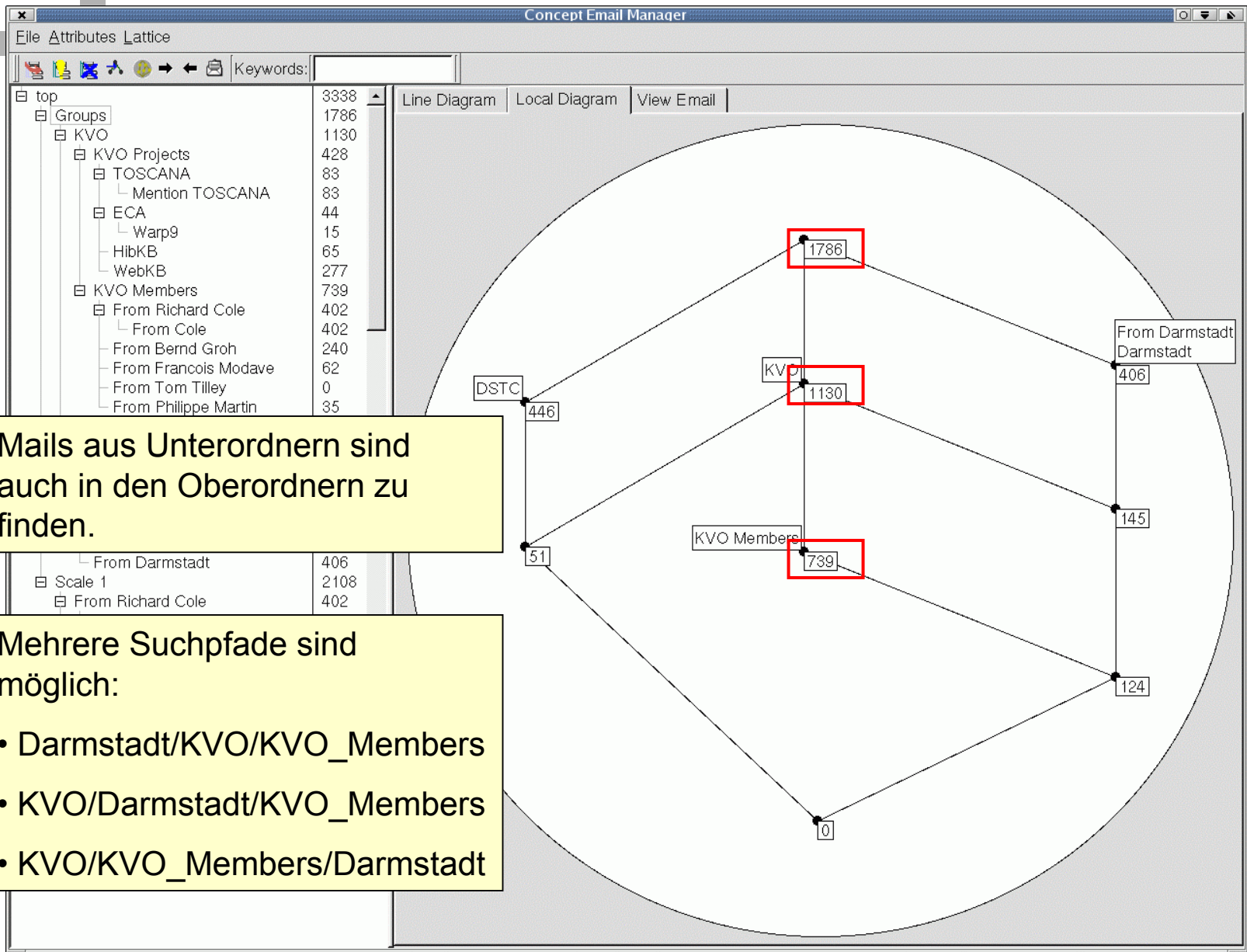
here's the Tex-File of our paper. :  
 llncs.cls, please have a look at tl  
 follow the links to the Springer A

See you at the Sushi place  
 Gerd

# Browsing basierend auf Formaler Begriffsanalyse



# Browsing basierend auf Formaler Begriffsanalyse



Concept Email Manager

File Attributes Lattice

Keywords:

Line Diagram Local Diagram View Email

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

Zoom 0

- KVO
- KVO

Verschiedene Sichten können kombiniert werden.

## VII.5 Formale Begriffsanalyse

### VII.5.5 A reduced representation of association rules based on Formal Concept Analysis

- The input data of association rules algorithms can be written as a formal context  $(G, M, I)$ :  $M$  is the set of items,  $G$  consists of the transaction IDs, and the relation  $I$  is the list of transactions.
  - We will distinguish between exact and approximate association rules:
  - **Def.:** An **association rule**  $X \rightarrow Y$  (with  $X, Y \subseteq M$ ) is called **exact** if  $\text{conf}(X \rightarrow Y) = 1$  and **approximate** else.
- An exact association rule is also called an **implication**.



## VII.5 Formale Begriffsanalyse

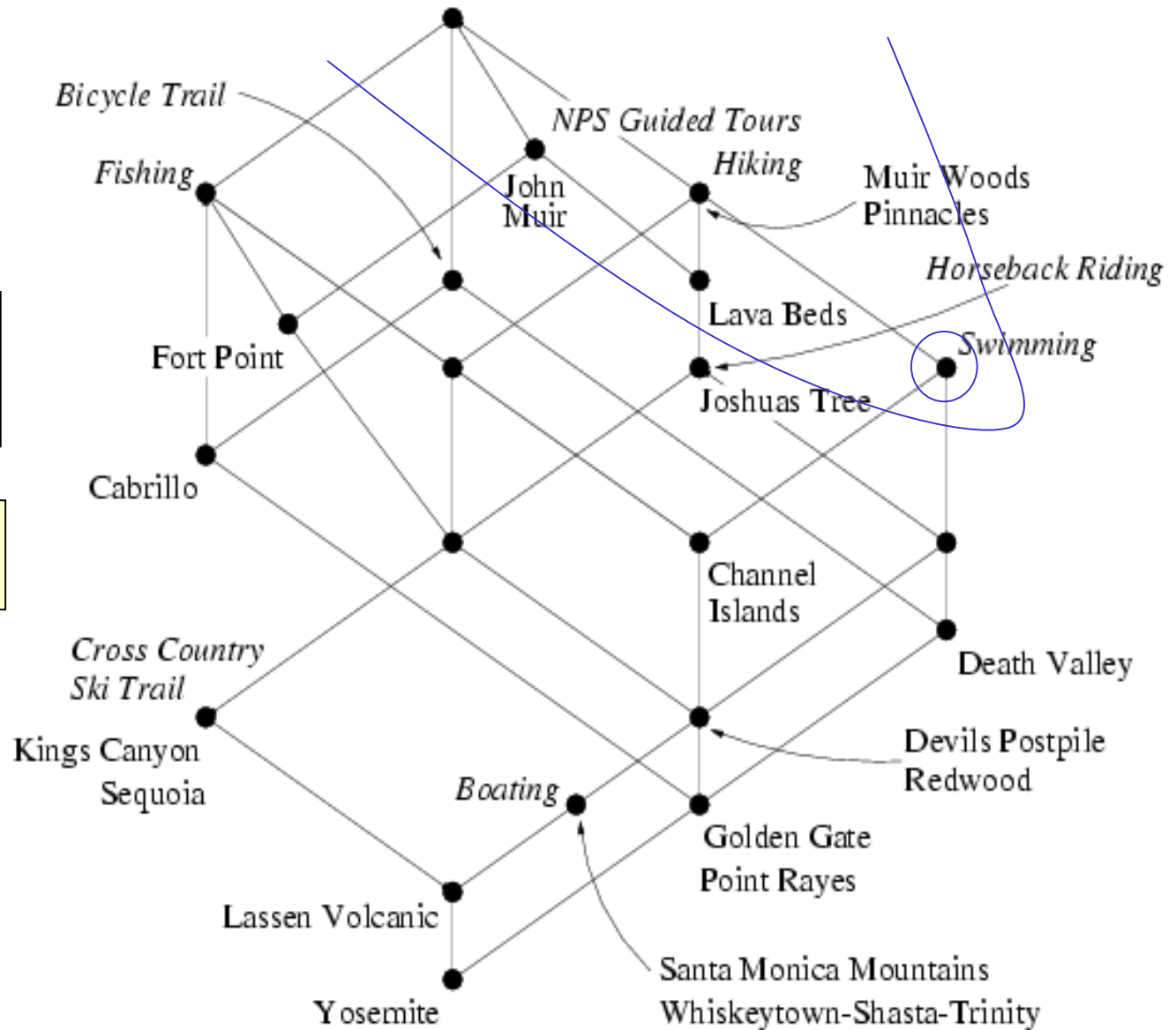
- In concept lattices, **exact association rules** can be directly read from the diagram:
- **Lemma:** An implication  $X \rightarrow Y$  holds iff the largest concept which is below all concepts generated by the attributes in  $X$  is below all concepts generated by attributes in  $Y$ .
- **Examples:**
  - {Swimming}  $\rightarrow$  {Hiking}  
(supp=10/19  $\approx$  52.6%, conf = 100%)

-

## VII.5 Formale Begriffsanalyse

„Welche Merkmale  
folgen noch aus  
,Swimming‘?“

„Hiking, (NPS Guided  
Tours)“



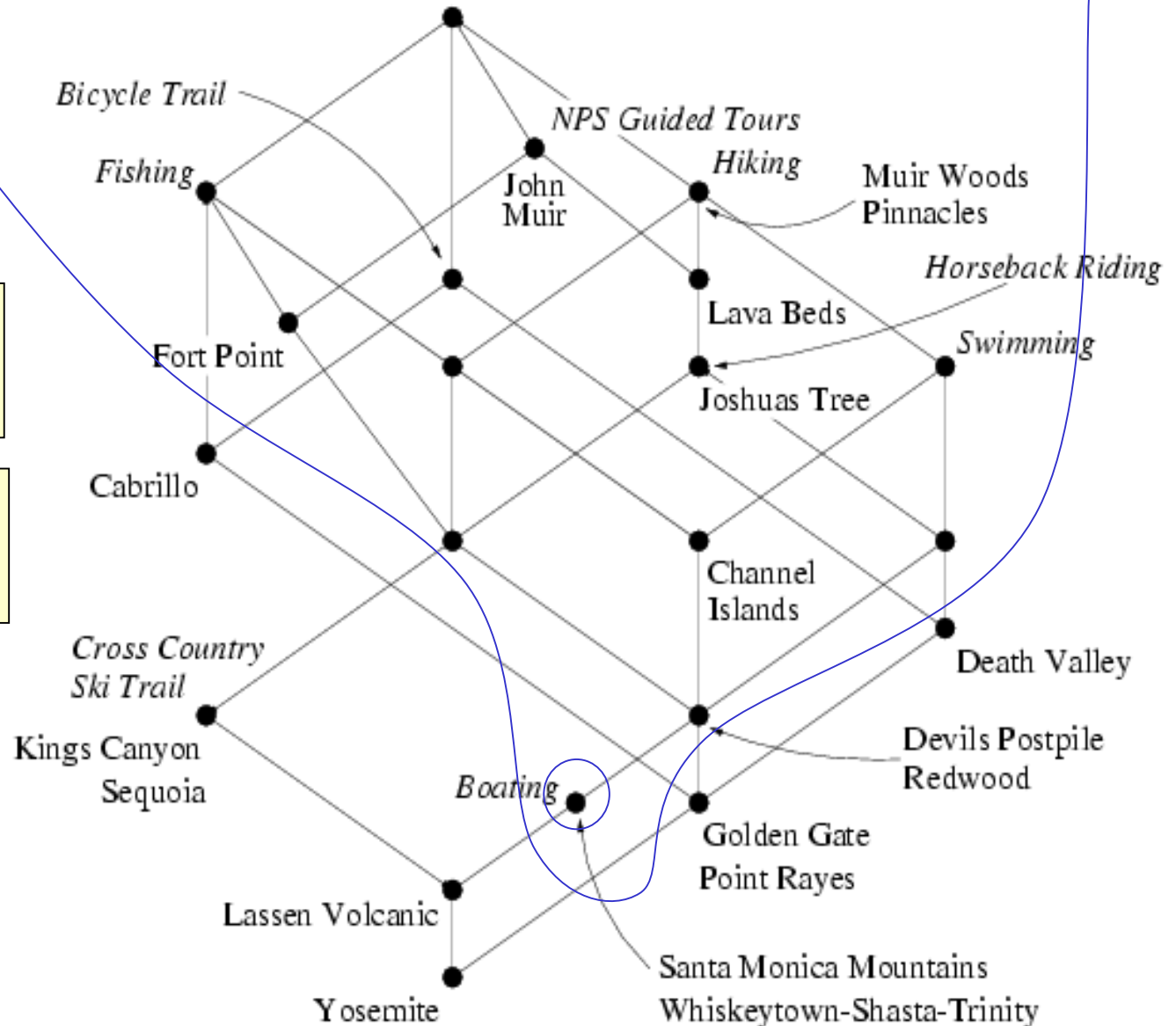
## VII.5 Formale Begriffsanalyse

- In concept lattices, **exact association rules** can be directly read from the diagram:
- **Lemma:** An implication  $X \rightarrow Y$  holds iff the largest concept which is below all concepts generated by the attributes in  $X$  is below all concepts generated by attributes in  $Y$ .
- **Examples:**
  - $\{\text{Swimming}\} \rightarrow \{\text{Hiking}\}$   
(supp=10/19  $\approx$  52.6%, conf = 100%)
  - $\{\text{Boating}\} \rightarrow \{\text{Swimming, Hiking, NPS Guided Tours, Fishing}\}$   
(supp=4/19  $\approx$  21.0%, conf = 100%)
  -

## VII.5 Formale Begriffsanalyse

„Welche Merkmale folgen noch aus ‚Boating‘?“

„Swimming, Hiking, NPS Guided Tours, Fishing“



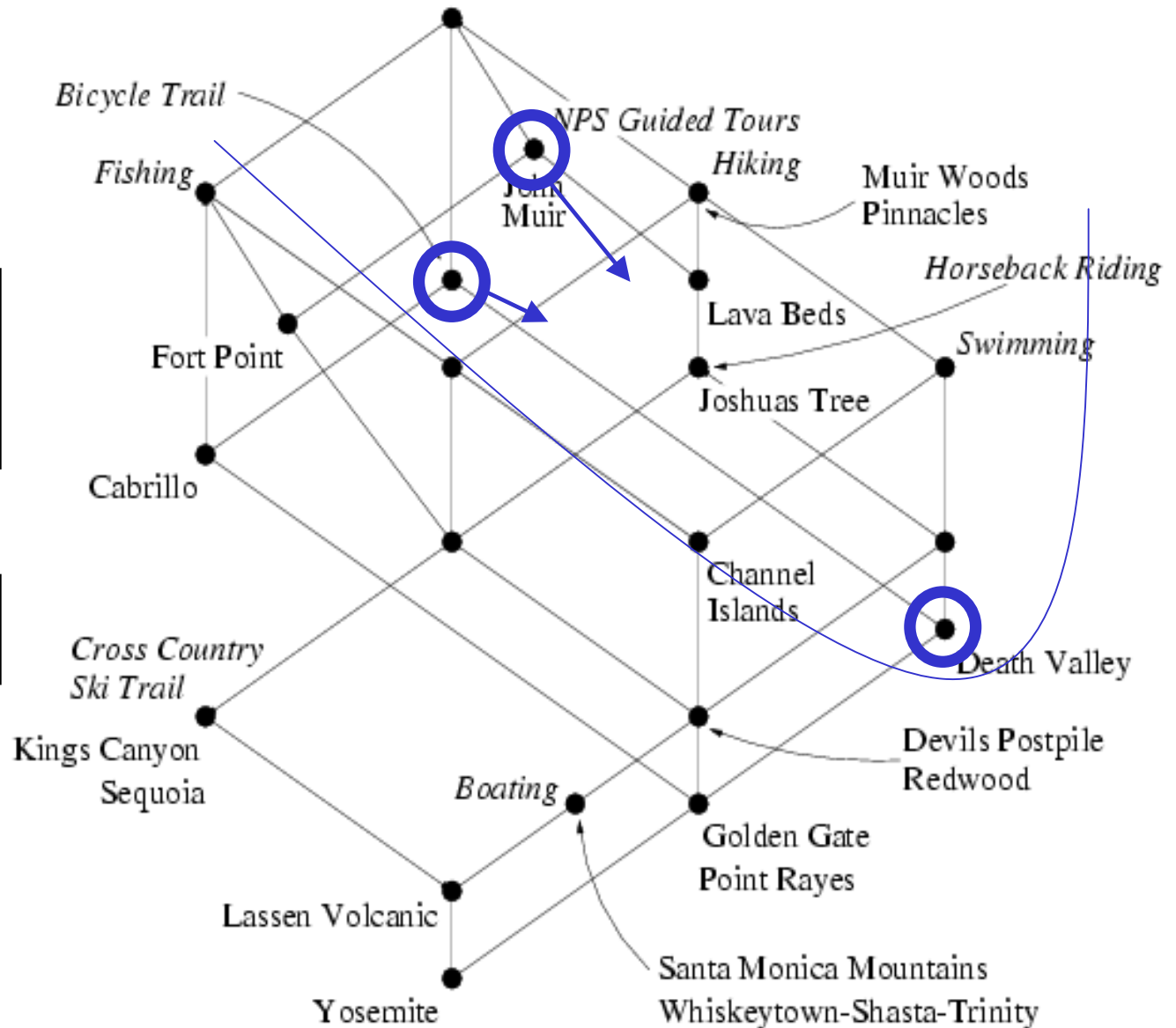
## VII.5 Formale Begriffsanalyse

- In concept lattices, **exact association rules** can be directly read from the diagram:
- **Lemma:** An implication  $X \rightarrow Y$  holds iff the largest concept which is below all concepts generated by the attributes in  $X$  is below all concepts generated by attributes in  $Y$ .
- **Examples:**
  - {Swimming}  $\rightarrow$  {Hiking}  
(supp=10/19  $\approx$  52.6%, conf = 100%)
  - {Boating}  $\rightarrow$  {Swimming, Hiking, NPS Guided Tours, Fishing}  
(supp=4/19  $\approx$  21.0%, conf = 100%)
  - {Bicycle Trail, NPS Guided Tours}  $\rightarrow$  {Swimming, Hiking}  
(supp=4/19  $\approx$  21.0%, conf = 100%)

## VII.5 Formale Begriffsanalyse

„Welche Merkmale folgen noch aus ‚Bicycle Trail‘ und ‚NPS Guided Tours‘?“

„Swimming, Hiking, Horseback Riding“



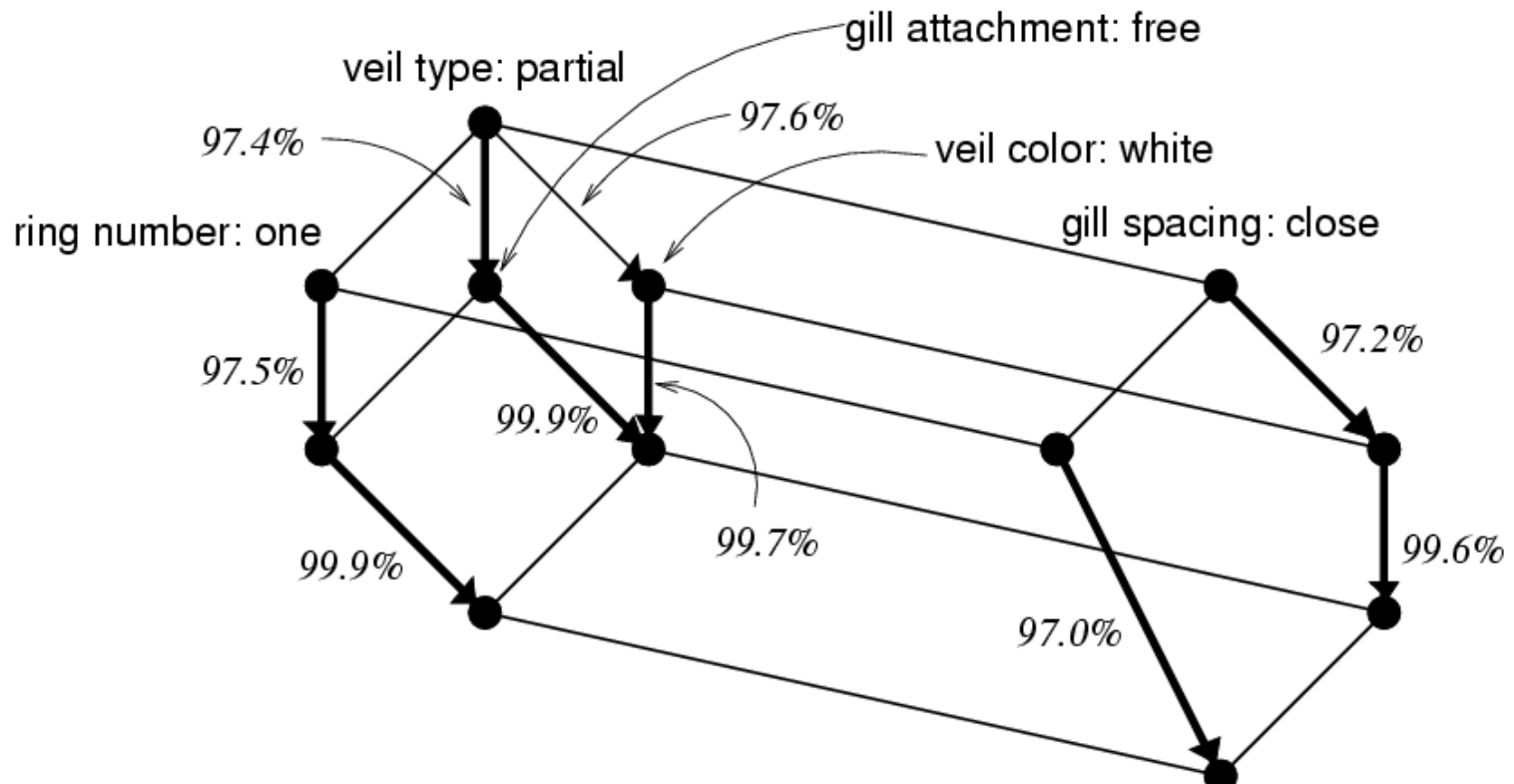
## VII.5 Formale Begriffsanalyse

- **Lemma:** For  $A \subseteq G$  and  $B \subseteq M$ , the following holds:
  - $A_1 \subseteq A_2 \Rightarrow A'_2 \subseteq A'_1$
  - $A \subseteq A''$
  - $A' = A'''$
  - $B_1 \subseteq B_2 \Rightarrow B'_2 \subseteq B'_1$
  - $B \subseteq B''$
  - $B' = B'''$
  - $A \subseteq B' \Leftrightarrow B \subseteq A' \Leftrightarrow A \times B \subseteq I$

From  $B' = B'''$  follows  $\text{supp}(B) = \frac{|B'|}{|G|} = \frac{|B'''}{|G|} = \text{supp}(B')$

Hence for computing association rules, it is sufficient to compute the supports of all frequent sets with  $B = B''$  (i.e., the intents [Inhalte] of the iceberg concept lattice).

## VII.5 Formale Begriffsanalyse



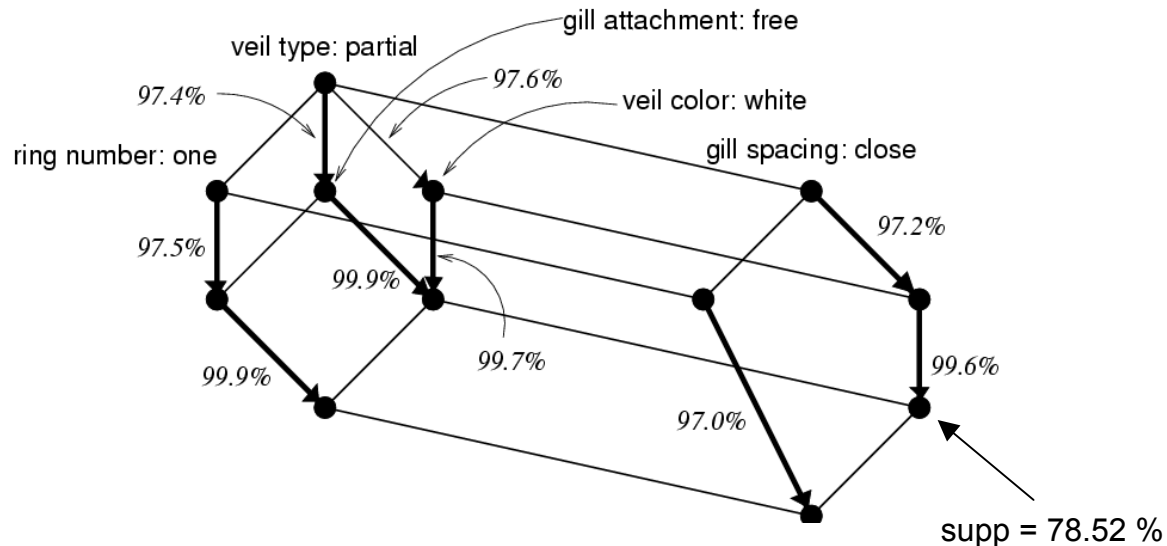
Association Rules can be visualized in the line diagram:

- exact rules as described before
- **approximate rules** as above ⑤ Luxenburger basis



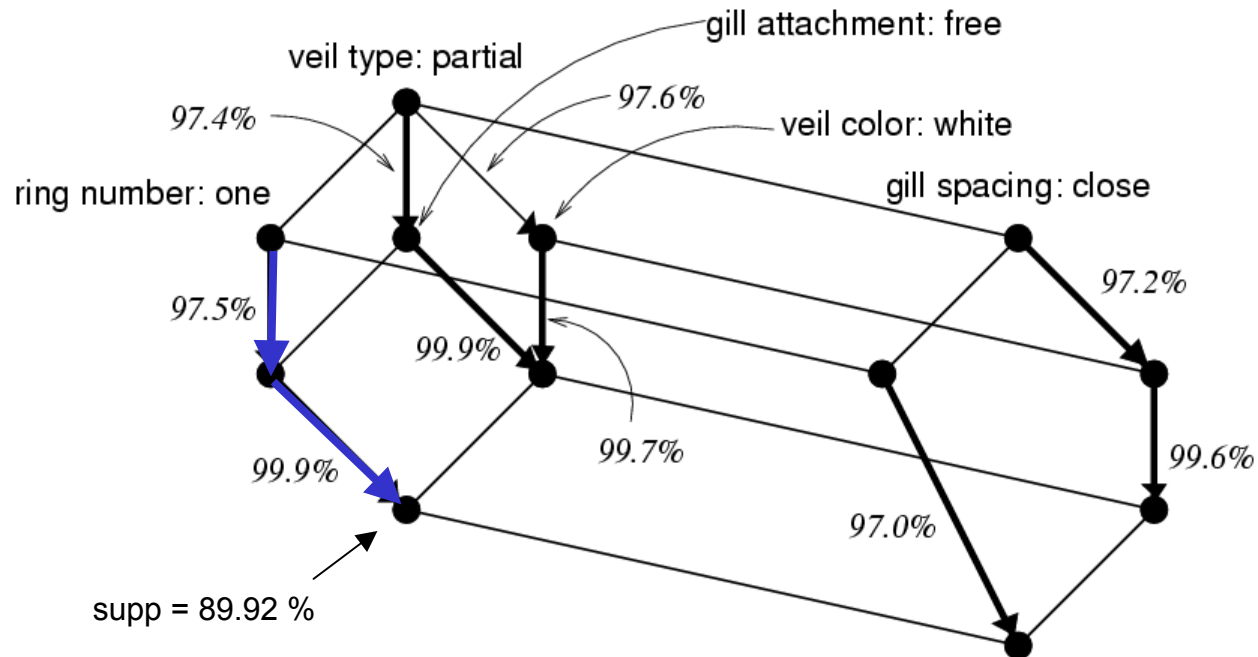
## VII.5 Formale Begriffsanalyse

**Def.:** The **Luxenburger basis** consists of all valid association rules  $X \rightarrow Y$  such that there are concepts  $(A_1, B_1)$  and  $(A_2, B_2)$  where  $(A_1, B_1)$  is a direct upper neighbor of  $(A_2, B_2)$ ,  $X = B_1$ , and  $X \cup Y = B_2$ .



Each arrow indicates a rule of the basis, e.g. the rightmost arrow stands for  $\{\text{veil type: partial, gill spacing: close, veil color: white}\} \rightarrow \{\text{gill attachment: free}\}$  (conf = 99.6 %, supp = 78.52 %)

## VII.5 Formale Begriffsanalyse



All other rules can be derived:

For instance  $\{\text{ring number: one}\} \rightarrow \{\text{veil color: white}\}$   
 has support 89.92 % (the support of the largest concept having both  
 attributes in its intent) and confidence  $97.5\% \times 99.9\% \approx 97.4\%$ .

## VII.5 Formale Begriffsanalyse

| Name       | Number of objects | Average size of objects | Number of items |
|------------|-------------------|-------------------------|-----------------|
| T10I4D100K | 100,000           | 10                      | 1,000           |
| MUSHROOMS  | 8,416             | 23                      | 127             |
| C20D10K    | 10,000            | 20                      | 386             |
| C73D10K    | 10,000            | 73                      | 2,177           |

### Some experimental results

| Dataset<br>(Minsupp) | Exact<br>rules | D.-G.<br>basis | Minconf | Approximate<br>rules | Luxenburger<br>basis |
|----------------------|----------------|----------------|---------|----------------------|----------------------|
| T10I4D100K<br>(0.5%) | 0              | 0              | 90%     | 16,269               | 3,511                |
|                      |                |                | 70%     | 20,419               | 4,004                |
|                      |                |                | 50%     | 21,686               | 4,191                |
|                      |                |                | 30%     | 22,952               | 4,519                |
| MUSHROOMS<br>(30%)   | 7,476          | 69             | 90%     | 12,911               | 563                  |
|                      |                |                | 70%     | 37,671               | 968                  |
|                      |                |                | 50%     | 56,703               | 1,169                |
|                      |                |                | 30%     | 71,412               | 1,260                |
| C20D10K<br>(50%)     | 2,277          | 11             | 90%     | 36,012               | 1,379                |
|                      |                |                | 70%     | 89,601               | 1,948                |
|                      |                |                | 50%     | 116,791              | 1,948                |
|                      |                |                | 30%     | 116,791              | 1,948                |
| C73D10K<br>(90%)     | 52,035         | 15             | 95%     | 1,606,726            | 4,052                |
|                      |                |                | 90%     | 2,053,896            | 4,089                |
|                      |                |                | 85%     | 2,053,936            | 4,089                |
|                      |                |                | 80%     | 2,053,936            | 4,089                |