Formale Begriffsanalyse
(B. Ganter, R. Wille)
Teil 1

Einleitung und Definitionen

Definition von formalen Kontexten

- $K{:=}(G,M,I)$: formaler Kontext
- G : Menge der Gegenstände
- M : Menge der Merkmale
- $I \subseteq G \times M$: Inzidenzrelation ($(g,m) \in I$ $\Leftrightarrow$ „m trifft auf g zu“)

Formaler Begriff

$A' := \{ m \in M \mid g \, \mathrm{I} \, m \text{ für alle } g \in A \}$

$B' := \{ g \in G \mid g \, \mathrm{I} \, m \text{ für alle } m \in B \}$

(A,B) heißt formaler Begriff des Kontextes (G,M,I) mit Umfang A und Inhalt B, wenn $A \subset G, B \subset M, A' = B$ und $B' = A$.

$\mathfrak{B}(G, M, I)$: Menge der Begriffe des Kontextes (G, M, I)

Bemerkungen zu Begriffen

- (A'', A') ist stets ein Begriff
- A'' ist der kleinste Begriffsumfang, der A umfaßt
- $A \subseteq G$ ist genau dann ein Begriff, wenn $A=A''$

Begriffsverband

(A_1, B_1) ist Unterbegriff von (A_2, B_2),

falls $A_1 \subseteq A_2$.

Dies definiert eine Ordnung $\leq$ auf $B(G, M, I)$.

$(B(G, M, I), \leq)$ ist der Begriffsverband

zu (G, M, I).

Hauptsatz über Begriffsverbände

$B(G, M, I)$ ist ein vollständiger Verband.

Infimum: $\bigwedge_{t \in T} (A_t, B_t) = \left(\bigcap_{t \in T} A_t, \left(\bigcup_{t \in T} B_t \right)'' \right)$

Supremum: $\bigvee_{t \in T} (A_t, B_t) = \left(\left(\bigcup_{t \in T} A_t \right)'', \bigcap_{t \in T} B_t \right)$

Bedeutung von Begriffsverbänden

Ein Begriffsverband kann gesehen werden als:

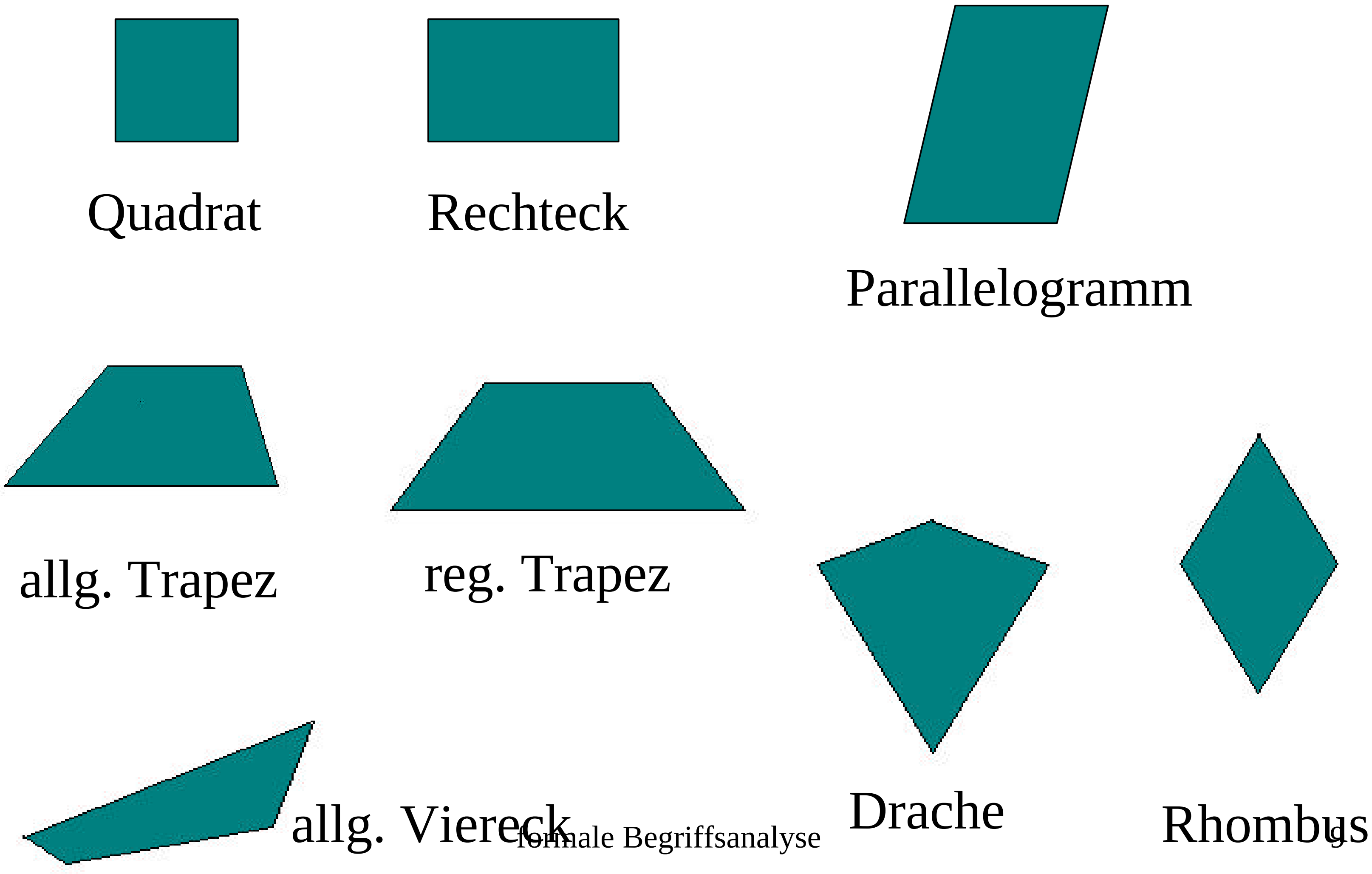
- Hierarchische Klassifikation von Gegenständen.
- System von Merkmalsimplikationen.
- Struktur zum Darstellen und Abfragen von Wissen.
- ...

Gegenstands- und Merkmalbegriffe

- Gegenstandsinhalt: $g' := \{m \in M \mid gIm\}$
- Merkmalumfang: $m' := \{g \in G \mid gIm\}$

- Gegenstandsbegriff: $gg := (g'', g')$
  (Begriff mit kleinstem Umfang, der g umfaßt)
- Merkmalbegriff: $mm := (m', m'')$

Ebene Vierecke



Kontext der ebenen konvexen Vierecke

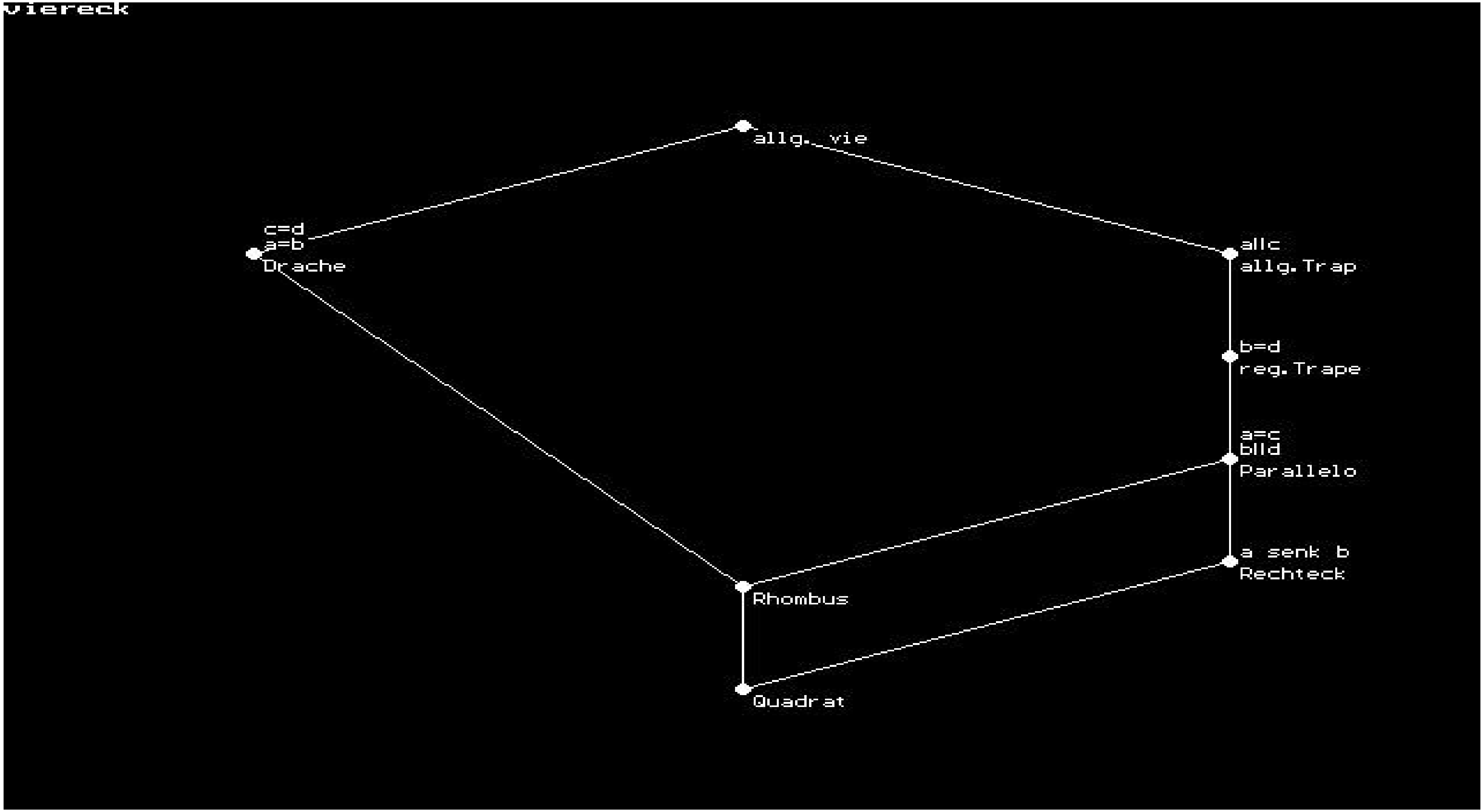
	a||c	b||d	a=b	c=d	a=c	b=d	a^b
Quadrat	x	x	x	x	x	x	x
Rechteck	x	x			x	x	x
Parallelo-gramm	x	x			x	x	
allg. Trapez	x						
reg. Trapez	x					x	
Drache			x	x			
Rhombus	x	x	x	x	x	x	
allg. Viereck							

Wie bestimme ich die Begriffsmenge?

- Jeder Begriffsumfang ist Durchschnitt von “Spaltenumfängen” m' ($m \in M$).
- Jeder Begriffsinhalt ist Durchschnitt von “Zeileninhalten” g' ($g \in G$).

- In Liniendiagrammen reicht es die Merkmal- und Gegenstandsbegriffe zu beschriften.

Begriffsverband der ebenen konvexen Vierecke



Kontextmanipulationen, die die Struktur des Verbandes erhalten

- Bereinigter Kontext (Gegenstände bzw Merkmale mit gleichem Informationsgehalt werden gestrichen)
- reduzierter Kontext (reduzible Merkmale, solche, die sich als Kombination anderer Merkmale schreiben lassen, werden gestrichen (analog: reduzible Gegenstände))

Bereinigter Kontext

Ein Kontext (G,M,I) heißt bereinigt, wenn für beliebige $g,h \in G$, $m,n \in M$ aus $g'=h'$ stets $g=h$ und aus $m'=n'$ stets $m=n$ folgt.

Verschiedene Gegenstände haben also verschiedene Gegenstandsinhalte und verschiedene Merkmale verschiedene Merkmalsumfänge.

Bereinigter Kontext der ebenen konvexen Vierecke

	a||c	b||d	a=b	b=d	a^b
Quadrat	x	x	x	x	x
Rechteck	x	x		x	x
Parallelo-gramm	x	x		x	
allg. Trapez	x				
reg. Trapez	x			x	
Drache			x		
Rhombus	x	x	x	x	
allg. Viereck					

Reduzierter Kontext

Ein Kontext heißt:

zeilenreduziert, wenn jeder Gegenstandsbegriff (g'',g') $\vee$-irreduzibel ist, d.h. (g'',g') hat genau einen unteren Nachbarn.

spaltenreduziert, wenn jeder Merkmalbegriff (m',m'') $\wedge$-irreduzibel ist, d.h. (m',m'') hat genau einen oberen Nachbarn.

reduziert, wenn er zeilen- und spaltenreduziert ist.

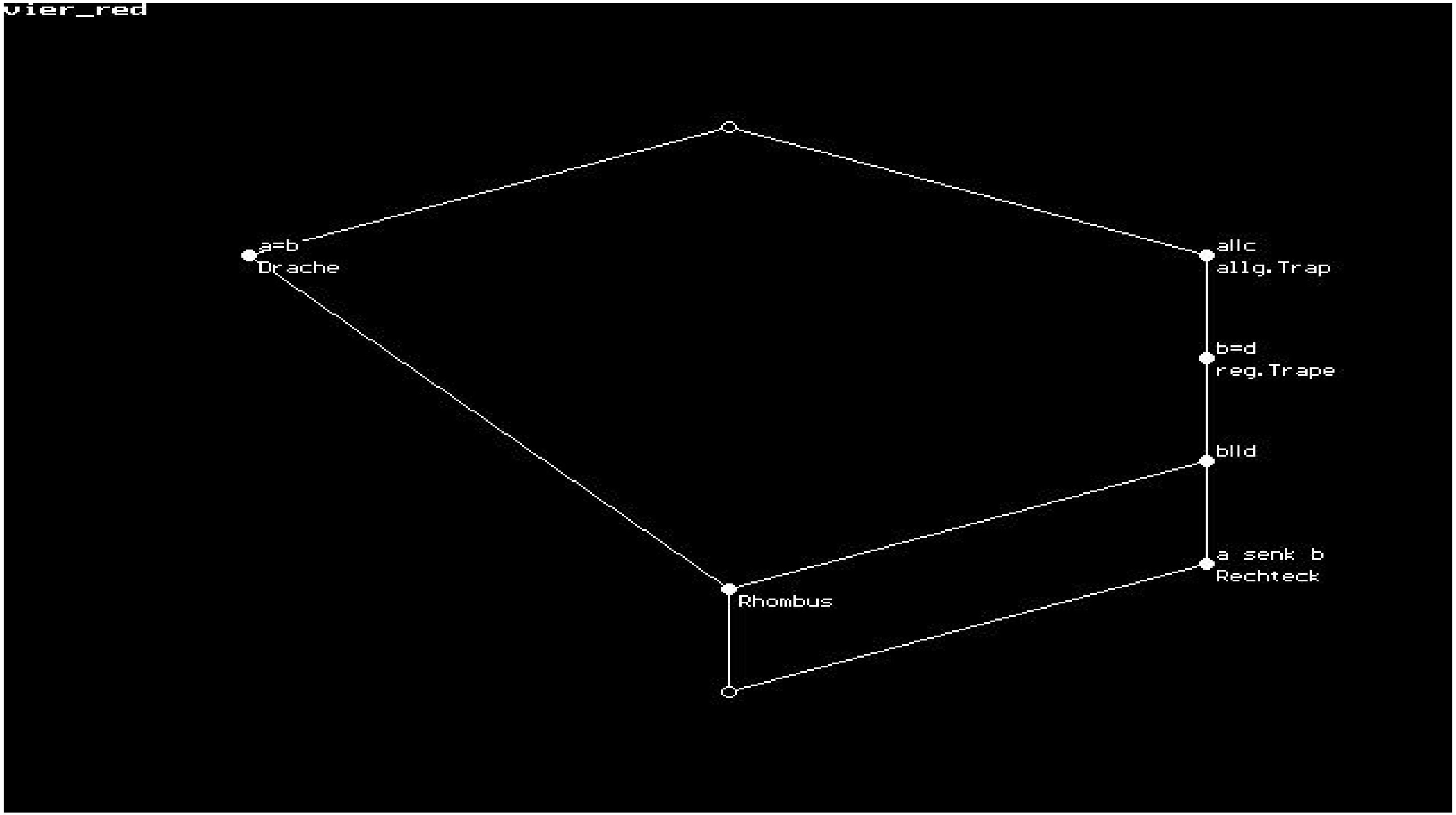
Standardkontext

- Zu jedem endlichen Verband, gibt es bis auf Isomorphie genau einen bereinigten, reduzierten Kontext, dieser heißt Standardkontext.
- Kontexte werden reduziert, indem man jeden Gegenstand, dessen Inhalt der Durchschnitt anderer Gegenstandsinhalte ist, streicht und analog mit den Merkmalen und ihren Umfängen vorgeht. Vollzeilen und Vollspalten werden ebenfalls gestrichen.

reduzierter Kontext der ebenen konvexen Vierecke

	a‖c	b‖d	a=b	b=d	a^b
Rechteck	x	x		x	x
allg. Trapez	x				
reg. Trapez	x			x	
Drache			x		
Rhombus	x	x	x	x	

Reduzierter Begriffsverband der ebenen konvexen Vierecke



Pfeilrelationen: Alternative zur Reduktion von Kontexten

$$g \swarrow m :\Leftrightarrow \{\neg gIm \text{ und falls } g' \subseteq h' \text{ und } g' \neq h',\ \text{dann } hIm\}$$

$$g \nearrow m :\Leftrightarrow \{\neg gIm \text{ und falls m}' \subseteq n' \text{ und } m' \neq n',\ \text{dann g}In\}$$

$$g \swarrow\!\!\!\!\nearrow m :\Leftrightarrow g \swarrow m \text{ und } g \nearrow m$$

- Ein bereinigter Kontext wird reduziert, indem alle Zeilen und Spalten, in denen $\swarrow\!\!\!\!\nearrow$ nicht vorkommt, gelöscht werden.

Versuch, die Pfeilrelationen natürlichsprachlich wiederzugeben.

- $g \nearrow m$ gilt nicht, wenn es ein Merkmal gibt, das auf zusätzliche Gegenstände zutrifft, nicht aber auf g.
- $g \swarrow m$ gilt genau dann, wenn g´ unter allen Gegenstandsinhalten, in denen m nicht vorkommt maximal ist; d.h. auf jeden Gegenstand, auf den mehr Merkmale zutreffen, trifft auch m zu.

Reduktion mit Pfeilrelationen

	a||c	b||d	a=b	b=d	a^b
~~**Quadrat**~~	~~x~~	~~x~~	~~x~~	~~x~~	~~x~~
Rechteck	x	x	⤢	x	x
~~**Parallelo-gramm**~~	~~x~~	~~x~~	~~↗~~	~~x~~	~~↗~~
allg. Trapez	x		↗	⤢	
reg. Trapez	x	⤢	↗	x	
Drache	⤢	↙	x	↙	
Rhombus	x	x	x	x	⤢
~~**allg. Viereck**~~	~~↗~~		~~↗~~		

1. Algorithmus zur Bestimmung der Begriffe eines Kontextes

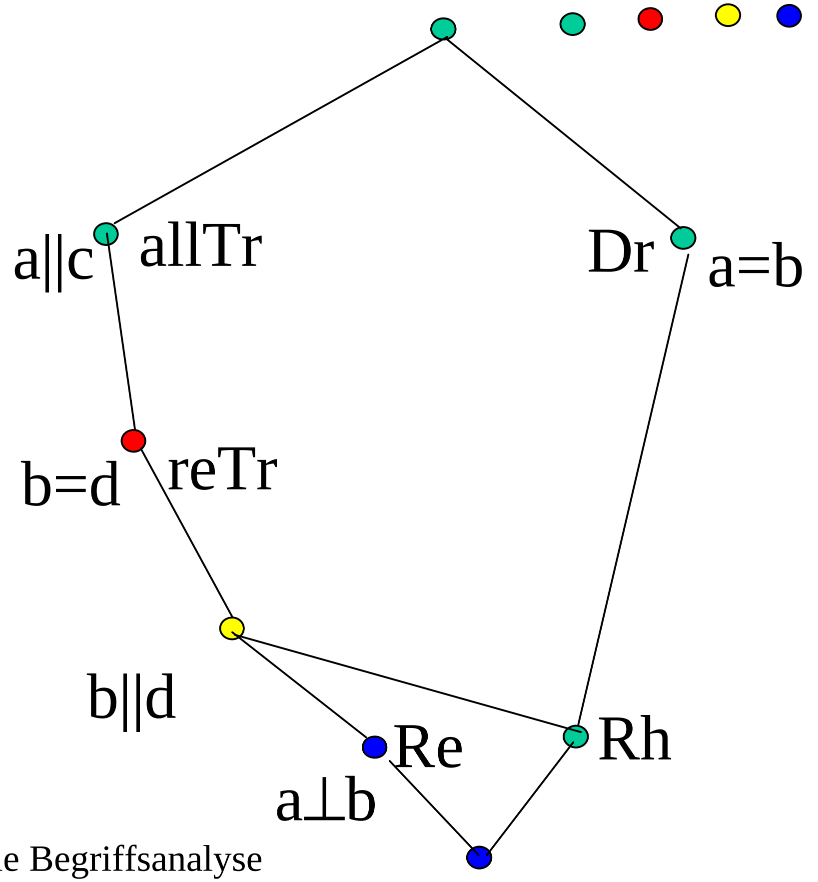
Satz: Jeder Begriffsumfang ist Durchschnitt von Merkmalsumfängen und jeder Begriffsinhalt ist Durchschnitt von Gegenstandsinhalten.

Schritt 1: Trage G in die Liste der Begriffsumfänge ein.

Schritt m ($m \in M$): Für jede Menge A, die bei einem früheren Schritt in die Liste eingetragen wurde, bilde die Menge $A \cap m'$ und trag sie, falls noch nicht vorhanden, ein.

Zeichnen des Begriffsverbandes mit Hilfe von Algorithmus 1

Schritt	**Umfang**
1	{Re, ..., Rh}
a‖c	{Re, allTr, reTr, Rh}
b‖d	{Re, Rh}
a=b	{Dr, Rh}, {Rh}
b=d	{Re, reTr, Rh}
a⊥b	{Re}



Besserer Algorithmus (Implementierung: ConImp)

Es gibt einen wesentlich schnelleren und leichter zu programmierenden *Algorithmus zur Erzeugung aller Hüllen eines gegebenen Hüllenoperators*, der auf dieses Problem anwendbar ist.

Implementierung: z.B. ConImp von Peter Burmeister