

FORMALBEGRIFFSANALYTISCHE UNTERSUCHUNG AUSGEWÄHLTER AROMATIZITÄTSKRITERIEN

Hans-Georg Bartel

Walther-Nernst-Institut für Physikalische und Theoretische Chemie
der Humboldt-Universität zu Berlin
Fachbereich Chemie, Hessische Straße 1-2, D-10115 Berlin
(received: February 1993)

Im März 1993 begingen zwei Wissenschaftler ihr 60. Geburtstagsjubiläum, die sich in besonderem Maße um die Herausbildung und die Entwicklung der Theoretischen Chemie im Ostteil Deutschlands durch ihre Tätigkeit als Forscher, akademische Lehrer und Organisatoren der Wissenschaft verdient gemacht haben:

Prof. Dr. HANS MÜLLER (Friedrich-Schiller-Universität Jena)
und

Prof. Dr. CORNELIUS WEISS (Universität Leipzig).

Aus diesem Anlaß und aus herzlicher Dankbarkeit für die unschätzbare große Unterstützung, welche auch ich sowohl in wissenschaftlicher als auch in menschlicher Hinsicht von ihnen in mehr als in einem Vierteljahrhundert erfahren durfte, möchte ich ihnen diesen Aufsatz widmen.

Es ist immer noch lohnend, sich mit dem als Aromatizität bezeichneten Phänomen zu beschäftigen, denn jeder diesbezügliche Beitrag kann sowohl der praktischen organischen Chemie als auch der Theorienentwicklung dieser Disziplin nützlich sein. Dabei kommt es nicht mehr so sehr darauf an, den Klassifikator aufzufinden, der die Menge der Verbindungen in aromatische und nichtaromatische zerlegt, sondern die Aromatizität als Merkmal zu erfassen, das die Aromaten zu ordnen versteht. Das erfolgt durch das Ermitteln von Maßzahlen, die nach dem jeweiligen physikalischen Inhalt, vom welchem man bei ihrer Festlegung ausgeht, in geometrische, energetische und magnetische ¹ eingeteilt werden können. Eine zweite Aufgliederung könnte von ihrer Herkunft ausgehen, d.h. von ihrer Erzeugung in einem Experiment oder durch quantenchemisch-theoretische Berechnungen auf der Grundlage der elektronischen Struktur der jeweiligen Moleküle.

Als eine jüngere Arbeit zu diesem Fragenkomplex ist diejenige von K. JUG und A.M. KÖSTER ² zu nennen, in welcher neue Aromatizitätskriterien entwickelt wurden, um sie in die Untersuchung in Hinblick auf eine Klassifikation solcher Kriterien einzubeziehen. Mit Hilfe der auch schon von A.R. KATRITZKY et al. ¹ verwendeten Hauptkomponentenanalyse konnte unter anderem gezeigt werden, daß die Aromatizität zumindest ein zweidimensionales Problem darstellt.

Auch in der vorliegenden Untersuchung wurden Kriterien und Daten aus der Arbeit ² herangezogen (s.u.). Um das Wesen und die Wirkungsweise der von R. WILLE und seiner Schule entwickelten Formalen Begriffsanalyse (FBA) ^{3,4} aufzuzeigen, wurde diese als Methode der Datenexploration und damit der Untersuchung ausgewählter Aromatizitätskriterien herangezogen. Dabei sollte besonderer Wert auf die Betrachtung der Komplexität und Dimensionalität dieser Kriterien und deren Gruppierung sowie auf eine extensionale und zugleich intentionale Klassifikation der als Anwendungsbeispiele benutzten zwölf aromatischen Verbindungen gelegt werden.

1. Einige Grundlagen der Formalen Begriffsanalyse

Da die auf der Begriffsverbandstheorie beruhende FBA in der chemischen Wissenschaft wohl noch kaum bekannt ist (s. aber ^{5,6} und eventuell ⁷), seien hier einige grundlegende Definitionen und Zusammenhänge wiederholt. Um den Zusammenhang der Darlegungen nicht zu stören, werden in einem Anhang (Abschnitt 7) einfache Demonstrationsbeispiele aufgeführt.

An den Anfang ist die Definition eines Kontexts zu setzen:

(D1) Ein (einwertiger) Kontext ist das Tripel (G, M, I) , wobei G und M Mengen sowie die Teilmenge I von $G \times M$ eine binäre Relation sind. G heißt die Gegenstands- und M die Merkmalsmenge. I ist die binäre Relation gIm mit $g \in G$ und $m \in M$: "Der Gegenstand g hat das Merkmal m ".

Ein derartiger Kontext kann als Tabelle aufgeschrieben werden, in welcher die Gegenstände die Zeilen und die Merkmale die Spalten bezeichnen. Der einzige Wert, z.B. ein "X", erscheint am Schnittpunkt der g -Zeile mit der m -Spalte, wenn gIm wahr ist. Die Menge der Begriffe wird bezogen auf einen Kontext ermittelt, wobei ein Begriff definiert ist als:

(D2) Ein (formaler) Begriff des Kontexts (G, M, I) ist das geordnete Paar (A, B) , in welchem die Teilmenge A von G der Umfang und die Teilmenge B von M der Inhalt des Begriffs heißen und für welches gilt, daß alle Gegenstände aus A alle Merkmale aus B gemeinsam haben und alle Merkmale aus B zu allen Gegenständen aus A gehören.

Diese Definition kann unter Verwendung der Operatoren

$$A' = \{ m \in M \mid \forall g \in A: gIm \} \text{ und } B' = \{ g \in G \mid \forall m \in B: gIm \},$$

wobei A wiederum eine Teilmenge von G und B eine Teilmenge von M bedeuten, vorteilhafter auch folgendermaßen aufschreiben:

(D2a) Ein (formaler) Begriff des Kontexts (G, M, I) ist das geordnete Paar (A, B) , in welchem die Teilmenge A von G der Umfang und die Teilmenge B von M der Inhalt des Begriffs heißen und für welches gilt: $A' = B$ und $B' = A$.

Die Übereinstimmung mit der ersten Formulierung ergibt sich daraus, daß A' die Menge aller Merkmale aus M , die die Gegenstände von A haben, darstellt und B' die Menge aller Gegenstände aus G , welche alle Merkmale von B besitzen.

Zwischen allen Paaren der Begriffsmenge wird eine Halbordnungsrelation \leq , die Unter/Oberbegriffsrelation, festgelegt:

(D3) (A_1, B_1) heißt Unterbegriff zum Oberbegriff (A_2, B_2) , wenn $A_2 \supseteq A_1$ gilt: $(A_1, B_1) \leq (A_2, B_2)$.

Die genannten Operatoren A' und B' haben die Eigenschaft, daß aus $A_2 \supseteq A_1$ die Relation $A_1' \supseteq A_2'$ bzw. aus $B_1 \supseteq B_2$ die Relation $B_2' \supseteq B_1'$ folgt. Daher lautet eine zu (D3) gleichbedeutende Formulierung:

(D3a) (A_1, B_1) heißt Unterbegriff zum Oberbegriff (A_2, B_2) , wenn $B_1 \supseteq B_2$ gilt.

Diese formale Festlegung entspricht den allgemeinen Vorstellungen von der Beziehung zwischen Begriffen. So ist der Begriff der Kohlenwasserstoffverbindungen ein Oberbegriff zu demjenigen der Alkane. Tatsächlich gibt es mehr Verbindungen mit der Bruttoformel C_nH_x als solche, für welche C_nH_{2n+2} richtig ist. Dafür kommt beim Unterbegriff der Alkane zu dem gemeinsamen Merkmal, daß die Moleküle nur aus C- und H-Atomen zusammengesetzt sind, noch das der Sättigung hinzu, das im Begriff der Kohlenwasserstoffe fehlt, da es nicht allen reinen C-H-Verbindungen gemeinsam ist.

Es kann gezeigt werden, daß die halbgeordnete Menge $(\mathcal{B}(G, M, I), \leq)$ ein Verband ist, wobei $\mathcal{B}(G, M, I)$ die Menge aller Begriffe zum Kontext (G, M, I) und \leq die in (D3) eingeführte Relation bedeuten:

(D4) Die halbgeordnete Menge $(\mathcal{B}(G, M, I), \leq)$ wird als Begriffsverband zum Kontext (G, M, I) bezeichnet.

Nun enthalten die zu untersuchenden Tabellen oder Datenmatrizen, wie etwa auch die Zusammenstellung mehrerer Aromatizitätskriterien für eine bestimmte Anzahl von Verbindungen, mehrere unterschiedliche Einträge. In jeder Spalte treten im allgemeinen verschiedene Werte bzw. Angaben auf. Eine solche Tabelle ist also als mehrwertiger Kontext aufzufassen, welcher folgendermaßen definiert ist:

(D5) Ein mehrwertiger Kontext ist ein Quadrupel (G, M, W, I) , in welchem G und M dieselbe Bedeutung wie in (D1) haben. W heißt die Menge der Werte (Ausprägungen). I ist eine ternäre Relation $(g, m, w) \in I$ mit $g \in G$, $m \in M$, $w \in W$: "Der Gegenstand g hat beim Merkmal m den Wert (die Ausprägung) w.", wobei die Gültigkeit von sowohl $(g, m, w_1) \in I$ als auch $(g, m, w_2) \in I$ die Gleichheit $w_1 = w_2$ impliziert. Für $|W| = n$ heißt der Kontext n-wertig.

Die für die Ermittlung des Begriffsverbandes notwendige Überführung eines gegebenen mehrwertigen Kontexts in einen einwertigen wird begriffliches Skalieren genannt⁸. Dabei wird jedem m, d.h. jedem (mehrwertigen) Merkmal, eine begriffliche Skala zugeordnet:

(D6) Es sei $m \in M$, wobei M die Merkmalsmenge eines mehrwertigen Kontexts (G, M, W, I) ist. Eine zu m gehörige begriffliche Skala ist ein einwertiger Kontext (G_m, M_m, I_m) , in welchem die Gegenstandsmenge der Teilmenge von W gleich ist, welche alle Werte (Ausprägungen) von m enthält. M_m und I_m werden so ge-

wählt, daß der Kontext (G_m, M_m, I_m) bzw. dessen Begriffsverband $(\beta(G_m, M_m, I_m), \leq)$ im Sinne der Interpretation so strukturiert sind, daß sie eine bestimmte begriffliche Bedeutung widerspiegeln. Tatsächlich ist mit der Skalierung bzw. der Festlegung der einzelnen Skalen eine Interpretation des Datencharakters der Merkmale verbunden. Die damit verbundene Willkür ist aber erfahrungsgemäß bedeutend geringer, als man im ersten Moment vielleicht glaubt. In sehr vielen Fällen kennt man die begriffliche Struktur der Merkmale oder die Anzahl der unterschiedlichen Möglichkeiten, die die Merkmalsstruktur sinnvollerweise aufweisen könnte, ist so gering, daß alle Fälle leicht nacheinander "ausprobiert" werden können. Für diese Feststellungen spricht auch die Praxis der Skalierung nominaler (klassifikatorischer) ⁹ und ordinaler Variabler ¹⁰ in der Taxonomie bzw. Cluster-Analyse beim Vorhandensein qualitativer Merkmale, welche der Wahl einer eindimensionalen nominalen bzw. ordinalen Skala in der FBA entspricht.

Für die hier vorzustellenden Untersuchungen genügt die Betrachtung bzw. Verwendung ordinaler und interordinaler Skalen, die daher im folgenden Abschnitt etwas näher dargestellt werden. Zuvor aber soll noch die graphische Darstellung der Begriffsverbände in Form der Linien-, Halbordnungs- oder Hasse-Diagramme erläutert werden.

Es sei (X, \leq) eine halbgeordnete Menge. Dann gelten zu ihrer graphischen Darstellung folgende Festlegungen: (a) Die $x \in X$ werden als Knoten $v(x)$ abgebildet. (b) Gilt $x_1, x_2 \in X$ und $x_1 \leq x_2$ mit $x_1 \neq x_2$, so werden die Knoten von x_1 und x_2 derart durch eine Kante verbunden, daß $v(x_1)$ unterhalb $v(x_2)$ gezeichnet wird, wobei die aus der Reflexivität von \leq resultierenden Schlingen und die Kanten, die sich aus der Transitivität von \leq ergeben, nicht mitgezeichnet werden. Letzteres ist dadurch begründet, daß die sich aus $x_h \leq x_i$ und $x_i \leq x_j$ ergebenden Kanten ohnehin die Gültigkeit von $x_h \leq x_j$ anzeigen.

Ist $(X, \leq) = (\beta(G, M, I), \leq)$ ein Begriffsverband, so ist aus Gründen der Übersichtlichkeit folgende Bezeichnungsweise der Knoten (Begriffe) im Hasse-Diagramm vorteilhaft: Jeder Gegenstand $g \in G$ und jedes Merkmal $m \in M$ erscheint im Diagramm nur einmal. Das Merkmal $m \in M$ erscheint an dem Knoten (gewöhnlich oberhalb), der zu dem Begriff gehört, der den kleinsten von allen Inhalten hat, welche m enthalten. Der Gegenstand $g \in G$ wird an den Knoten geschrieben (gewöhnlich unterhalb), dessen zugehöriger Begriff den kleinsten Umfang mit g als Element besitzt. Der Inhalt eines beliebigen Begriffs ist folglich so zu ermitteln, daß man zu den an seinen Knoten angeschriebenen Merkmalen alle diejenigen hinzufügt, welche an Knoten oberhalb des betrachteten stehen und von diesem über ständig steigende Kantenzüge erreichbar sind. Für die Vervollständigung des Umfangs eines bestimmten Begriffs muß man alle Gegenstände zu den am Knoten vorhandenen hinzufügen, die an Knoten stehen, welche auf nach unten führenden Kantenzügen erreicht werden können. So wird der Umfang des obersten Begriffs (1-Element des Verbandes) von der vollständigen Gegenstandsmenge G des Kontexts gebildet, sein Inhalt informiert also über die

Merkmale, die allen Gegenständen gemeinsam sind. Dieser Inhalt kann leer sein. Zum untersten Knoten (0-Element des Verbandes) gehört der Begriff, dessen Inhalt alle Merkmale aus M umfaßt, so daß sein Umfang die Gegenstände nennt, die alle Merkmale des Kontexts haben, und folglich auch leer sein kann.

2. Über ordinale und interordinale Skalierung

Wie oben bemerkt, werden hier vornehmlich eindimensionale ordinale und interordinale Skalen von Bedeutung sein. Die ersteren, die die Bedeutung einer Anordnung haben, lassen sich allgemein durch den Kontext (N, N, \geq) beschreiben, wobei N eine Teilmenge der natürlichen Zahlen ist ($N = \{1, 2, \dots, n\}$). Soll also ein Merkmal $m \in M$ eines mehrwertigen Kontexts (G, M, W, I) , dessen Ausprägungsmenge die Teilmenge G_m von W bildet, ordinal skaliert werden, so ist eine Abbildung von G_m in eine solche Menge N vorzunehmen, um die ordinale Skala (N, N, \geq) zu erhalten. Das Hasse-Diagramm einer solchen Skala ist eine Kette.

Eine eindimensionale Interordinalskala hat die begriffliche Bedeutung eines (linearen) Dazwischenseins und wird durch den Kontext $(N, N \times \{1, 2\}, \$)$ gekennzeichnet, wobei N wieder die genannte Teilmenge der natürlichen Zahlen ist. Die Relation $\$$ wird durch

$$i \$ (j, 1) := i \geq j \text{ und } i \$ (j, 2) := i \leq j \quad (i, j \in N)$$

beschrieben. In Abbildung 1 wird als Beispiel die Interordinalskala mit der Teilmenge $N = \{1, 2, 3, 4\}$ dargestellt.

Es sei wiederum $N = \{1, 2, \dots, n\}$ die Teilmenge der natürlichen Zahlen, π_i bedeute eine der $|N| = n!$ Permutationen, d.h. Abbildungen von N auf N . Durch (N, N, \geq_i) werde eine ordinale Skala bezeichnet, in der \geq_i definiert ist als:

$$x \geq_i y := \pi_i(x) \geq \pi_i(y) \text{ mit } x, y \in N.$$

Nun werde eine Skala $\mathcal{J}(\pi_i) := (N, N \times \{1, 2\}, \$_i)$ als Verallgemeinerung der Interordinalskala betrachtet. Hier bedeutet:

$$x \$_i (y, 1) := x \geq y \text{ und } x \$_i (y, 2) := \pi_i(x) \geq \pi_i(y) \text{ mit } x, y \in N.$$

Es ist leicht zu sehen, daß die identische Abbildung $\pi_1(x) = x$ für alle $x \in N$ zu einer ordinalen Skala $\mathcal{J}(\pi_1)$ führt. Die Abbildung $\pi_n(x) = n+1-x$ stellt die völlige Umkehrung der Reihenfolge der natürlichen Zahlen $1, 2, \dots, n$ dar, so daß $\mathcal{J}(\pi_n)$ eine Interordinalskala ist. Alle Skalen $\mathcal{J}(\pi_i)$ mit $1 < i < n!$ liegen in ihrer Komplexität zwischen der linearen Ordinalität ($i = 1$) und der vollständigen Interordinalität ($i = n!$). Die Ursache für die zunehmende Komplexität der Begriffsverbände $(\mathcal{B}(\mathcal{J}(\pi_i)), \leq)$ ist also in der Anzahl der Umstellungen in der Reihenfolge der Gegenstände beim Vergleich ihrer Anordnung bezüglich der Merkmalsklasse $(x, 1)$ mit derjenigen bezüglich $(x, 2)$ zu sehen. Der Terminus "Komplexität" wird hier

qualitativ bzw. intuitiv aufgefaßt, so daß auf die Festlegung eines entsprechenden Maßes verzichtet werden kann.

	≥ 1	≥ 2	≥ 3	≥ 4	≤ 1	≤ 2	≤ 3	≤ 4
1	X				X	X	X	X
2	X	X				X	X	X
3	X	X	X				X	X
4	X	X	X	X				X

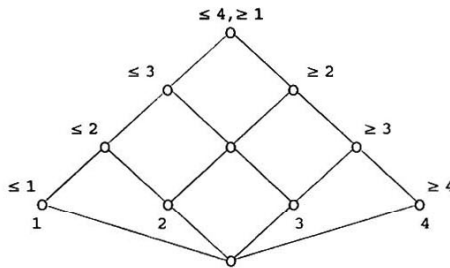


Abb. 1 Kontext $(\{1,2,3,4\}, \{1,2,3,4\} \times \{1,2\}, \$)$ und zugehöriger Begriffsverband (Die Merkmale $(i,1)$ wurden mit $\geq i$, die Merkmale $(i,2)$ mit $\leq i$ bezeichnet.)

Zur Illustration wurden die Begriffsverbände $(\mathcal{B}(\mathcal{J}(\pi_i)), \leq)$ zu allen 24 Permutationen $\pi_i = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \pi_i(1) & \pi_i(2) & \pi_i(3) & \pi_i(4) \end{pmatrix}$ der Menge $N = \{1,2,3,4\}$ in Abbildung 2

dargestellt, wobei auf die Kennzeichnung der Umfänge verzichtet wurde. Angezeigt wurde neben den entsprechenden Diagramm die Folge $\pi_i(1), \pi_i(2), \pi_i(3), \pi_i(4)$. Die Merkmale $(x,1)$ wurden mit x und die Merkmale $(x,2)$ mit $x+4$ ($x = 1,2,3,4$) bezeichnet.

Die Skala $\mathcal{J}(\pi_i)$ kann als das Ergebnis der ordinalen Skalierung beider Merkmale eines mehrwertigen Kontexts $(G, \{m_1, m_2\}, W, I)$ aufgefaßt werden. Bedeuten G_1 und G_2 die Ausprägungen der Merkmale m_1 und m_2 , so ist der skalierte Kontext mit $\mathcal{J}(\pi_i)$ identisch, wenn $|G| = |G_1| = |G_2| = n$ gilt. Die relativen

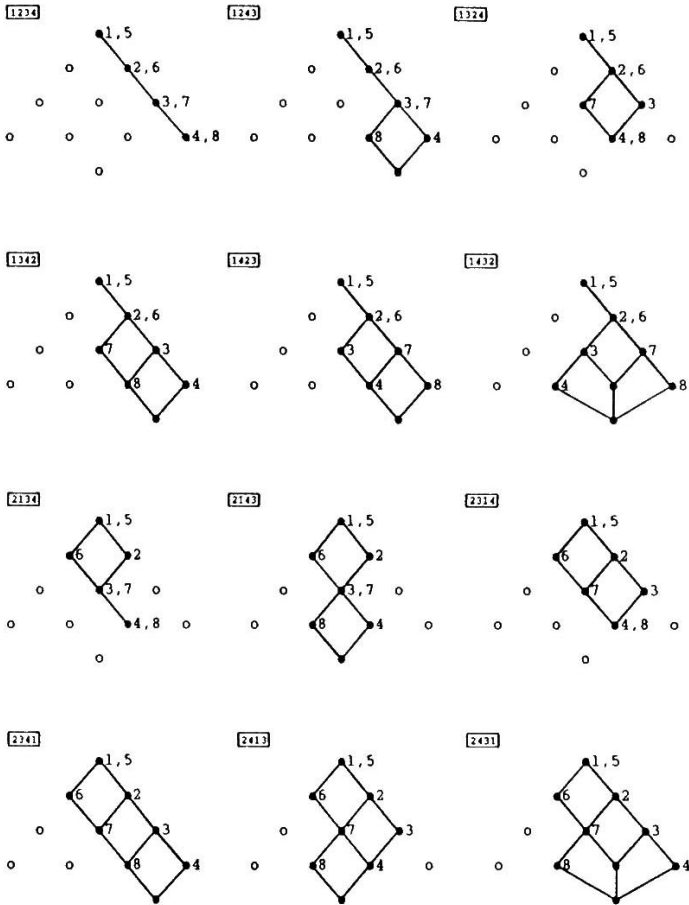


Abb. 2 Hasse-Diagramme der Begriffsverbände $(\mathcal{B}(J(\pi)), \leq)$ für $N = \{1, 2, 3, 4\}$
 (Bezeichnungen s. Text)

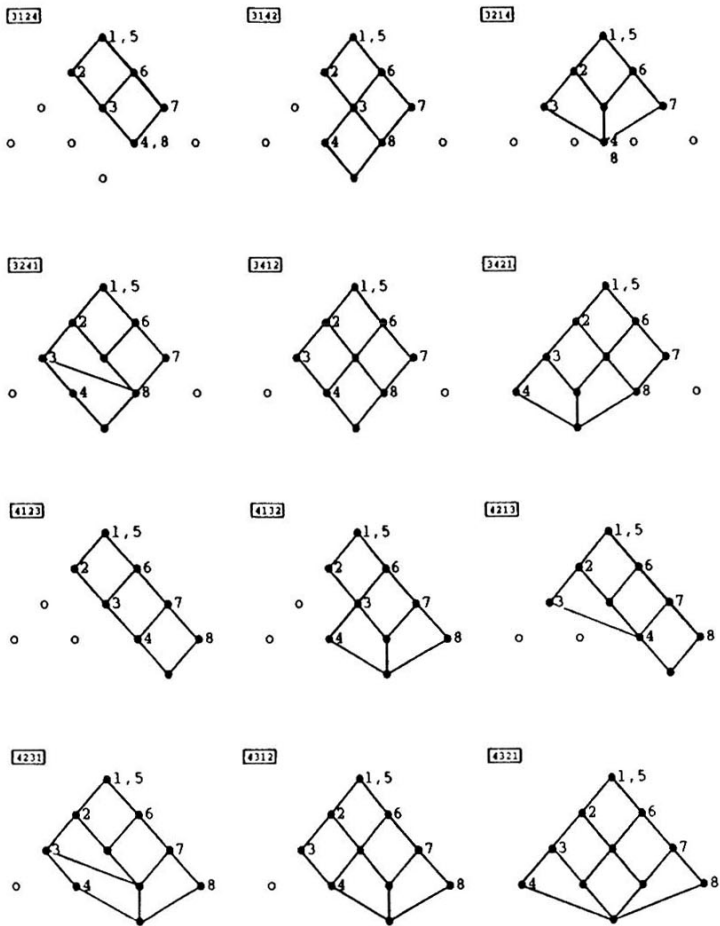


Abb. 2 Fortsetzung

Komplexitätsverhältnisse lassen sich aber allgemein in allen anderen Fällen sinnvoll erfassen, in denen $|G| \leq |G_1| + |G_2|$ ist.

3. Der mehrwertige Kontext

Für die einleitend genannten Ziele wird der in Tabelle 1 gegebene mehrwertige Kontext benutzt, der der Arbeit ² entnommen wurde.

Tabelle 1 Aromaten und Aromatizitätskriterien nach ²

Aromat	RC _v	$-10^6 \chi_{ZZ}^{\pi}/n_{\pi}$ [cm ³ /mol]	$10^{-4} \Delta E/E$	$-10^6 \chi$ [cm ³ /mol]
Benzen	1,437	10,05	6,8	54,8
Pyridin	1,422	9,67	6,2	49,2
Pyrimidin	1,411	9,28	5,7	43,1
1,3,5-Triazin	1,404	8,88	6,2	37,9
Pentaazin	1,407	8,18	1,7	
Hexaazin	1,426	7,86	0,3	
Pyrrrol	1,124	7,48	3,8	47,6
Furan	1,081	7,25	2,7	43,1
Pyrazol	1,096	7,43	3,7	42,6
Imidazol	1,075	7,17	3,0	44,3
1,2-Oxazol	1,032	6,98	2,4	37,5
1,3-Oxazol	1,042	7,03	2,2	39,2

Die aufgeführten zwölf aromatischen Moleküle bilden die Gegenstands-
menge. Die vier Merkmale sind Vertreter der Kriterienklassen nach KATRITZKY
et al. ¹:

(a) geometrisches Kriterium: RC_v (Abkürzung auch RCv) ist ein von JUG und
KÖSTER ² eingeführtes Kriterium und bedeutet die kleinste Bindungswalenz der
Bindungen zwischen den Atomen in einem monozyklischen Ring. Die Bindungs-
valenz V_{AB} zwischen den Atomen A und B wird aus den quantenchemisch be-
rechneten Dichtematrixelementen P_{μAνB} ermittelt.

(b) magnetische Kriterien: χ_{ZZ}^{π}/n_{π} (andere Abkürzungen: χ^{π} , chi) bedeutet die
quantenchemisch berechnete diamagnetische Suszeptibilität des π -Elektronensy-
stems senkrecht zur Ringebene χ_{ZZ}^{π} bezogen auf die Anzahl der π -Elektronen n_{π} .
Auch dieses Kriterium wurde von JUG und KÖSTER ² eingeführt. Dagegen ist χ
(Abkürzung auch Susz) die experimentell ermittelte Suszeptibilität.

(c) energetisches Kriterium: $\Delta E/E$ (weitere Abkürzungen: ΔE , dE/E) ist das Verhältnis der Energiedifferenz zwischen der Kekulé- und der Gleichgewichtsstruktur ΔE zur Energie der Gleichgewichtsstruktur E . Dieses aus quantenchemischen Rechnungen ermittelbare Kriterium wurde ebenfalls von JUG und KÖSTER¹¹ vorgeschlagen.

Tabelle 2 Angaben zu den aromatischen Molekülen

Molekül	Anzahl pro Ring			
	Atome	N-Atome	O-Atome	Bindungen zwischen Heteroatomen
Benzen	6	0	0	0
Pyridin	6	1	0	0
Pyrimidin	6	2	0	0
1,3,5-Triazin	6	3	0	0
Pentaazin	6	5	0	4
Hexaazin	6	6	0	6
Pyrrrol	5	1	0	0
Furan	5	0	1	0
Pyrazol	5	2	0	1
Imidazol	5	2	0	0
1,2-Oxazol	5	1	1	1
1,3-Oxazol	5	1	1	0

Neben dieser Einteilung lassen sich die ausgewählten Kriterien in folgende zwei Klassen zerlegen:

(A) quantenchemisch-theoretische Kriterien: RC_v , χ_{π}^2/n_{π} und $\Delta E/E$

(B) experimentell-magnetisches Kriterium: χ

Die Gegenstände (Moleküle) können durch die Angaben der Tabelle 2 charakterisiert werden.

Zur Demonstration der vorhergehenden Abschnitte wird in der Abbildung 3 der Begriffsverband des mehrwertigen Kontexts, welchen Tabelle 2 darstellt, wiedergegeben, wobei das erste Merkmal nominal (6-Ring, 5-Ring) und das zweite und vierte ordinal (1N, 2N, 3N, 5N, 6N, bzw. b1, b4, b6) skaliert wurden. Das dritte Merkmal wurde einwertig (1 O-Atom $\hat{=}$ X) aufgefaßt.

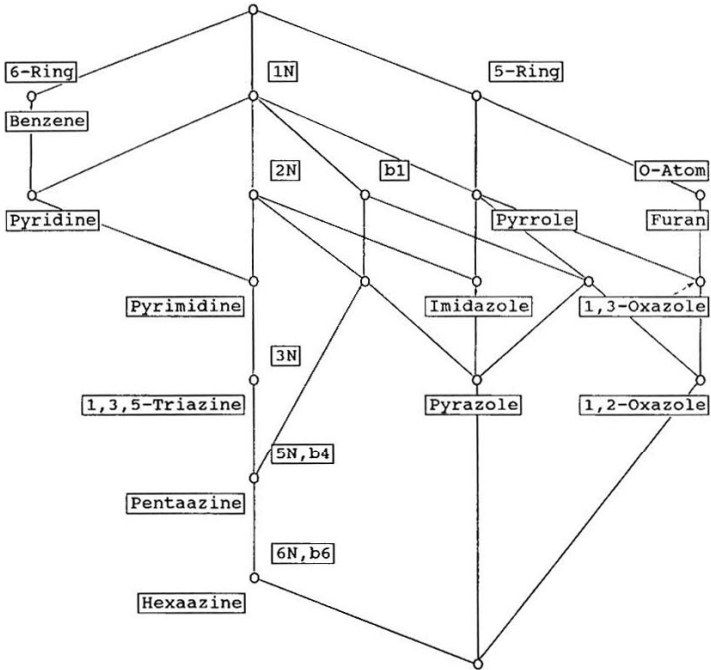


Abb. 3 Hasse-Diagramm des Begriffsverbandes zum mehrwertigen Kontext der Tabelle 2 (Skalierung s. Text)

4. Skalierung und einwertiger Kontext

Da die Werte in der Tabelle 1 die aromatischen Verbindungen bezüglich des jeweiligen Kriteriums und des Grades ihrer Aromatizität ordnen, bietet sich für alle Merkmale eine eindimensional ordinale Skalierung an, d.h., es werden

Skalen des Typs (N, N, \geq) mit $N = \{1, 2, \dots, n\}$ verwendet. Dabei ist $n = 12$ für RC_v und $\chi_{\frac{E}{E}/n_{\pi}}$, $n = 11$ für $\Delta E/E$ und $n = 9$ für χ . Natürlich wird die Abbildung der Ausprägungen (Werte) eines Merkmals auf die Menge N so vorgenommen, daß einem kleineren Absolutwert einer Ausprägung beim paarweisen Vergleich stets auch die kleinere natürliche Zahl beim entsprechenden Vergleich in der Bildmenge entspricht. Die Merkmale in den Skalen und damit auch im einwertigen Kontext sowie in den Hasse-Diagrammen der Begriffsverbände werden durch Vorsetzen der natürlichen Zahl aus N vor die jeweilige Benennung des Merkmals (s.o.) bezeichnet. In der Tabelle 3 ist der erhaltene einwertige Kontext wiedergegeben.

Tabelle 3 Der einwertige Kontext (Die Punkte "." dienen nur der Lesehilfe.)

	111	111	11	11
	123456789012	123456789012	12345678901	123456789
	RRRRRRRRRRRR	cccccccccccc	EEEEEEEEEEEE	uuuuuuuuuu
	CCCCCCCCCCCC	hhhhhhhhhhhh	//////////	ssssssss
	vvvvvvvvvvvv	iiiiiiiiiiii	EEEEEEEEEEEE	zzzzzzzz
Benzen	XXXXXXXXXXXXX	XXXXXXXXXXXXX	XXXXXXXXXXXXX	XXXXXXXXXXXXX
Pyridin	XXXXXXXXXXXXX.	XXXXXXXXXXXXX.	XXXXXXXXXXXXX.	XXXXXXXXXXXXX.
Pyrimidin	XXXXXXXXXXXXX...	XXXXXXXXXXXXX...	XXXXXXXXXXXXX...	XXXXXXXXXXXXX...
1,3,5-Triazin	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....
Pentaazin	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....
Hexaazin	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....
Pyrrol	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....	XXXXXXXXXXXXX.....
Furan	XXXXX.....	XXXXX.....	XXXXX.....	XXXXX.....
Pyrazol	XXXXX.....	XXXXX.....	XXXXX.....	XXXXX.....
Imidazol	XXX.....	XXX.....	XXX.....	XXX.....
1,2-Oxazol	X.....	X.....	XXX.....	X.....
1,3-Oxazol	XX.....	XX.....	XXX.....	XXX.....

5. Die Begriffsverbände

Ausgehend von diesem Kontext wurden die Berechnungen mit dem Programm CONIMP von P. BURMEISTER (Technische Hochschule Darmstadt, Fachbereich Mathematik) ¹² durchgeführt. In Abbildung 4 ist das Hasse-Diagramm des erhaltenen Begriffsverbandes dargestellt.

Um die vier Aromatizitätskriterien (Merkmale des mehrwertigen Kontexts) paarweise vergleichen zu können, wurde der gesamte einwertige Kontext in sechs Teilkontexte zerlegt, von denen jeder ein Paar der skalierten Merkmale

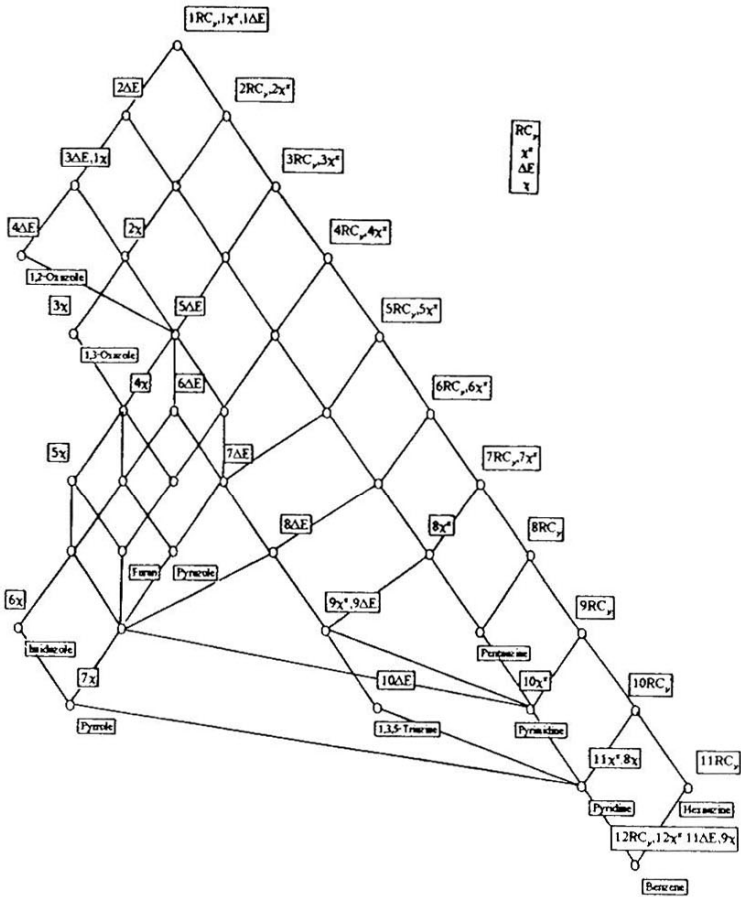


Abb. 4 Hasse-Diagramm des Begriffsverbandes zum Kontext in Tabelle 3

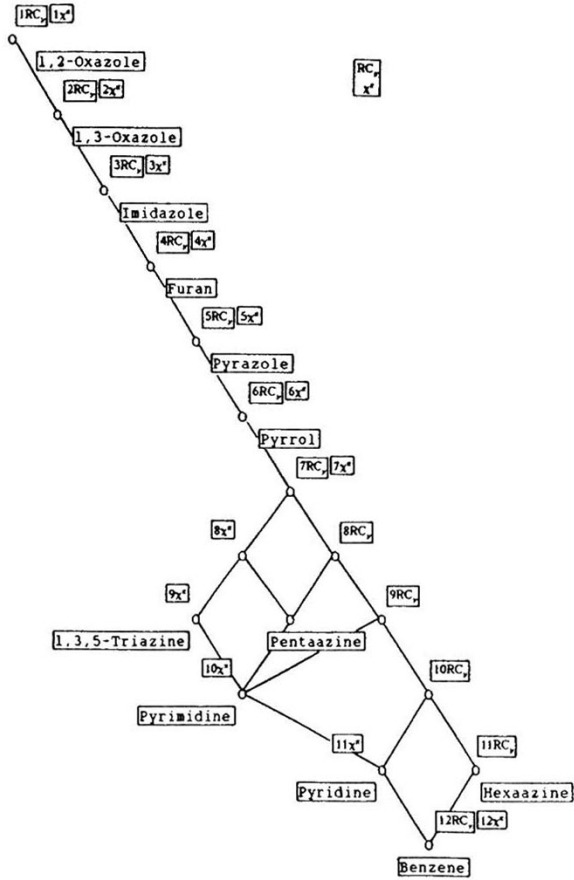


Abb. 5 Hasse-Diagrammen von Begriffsverbänden zu Teilen des Kontexts in Tabelle 3

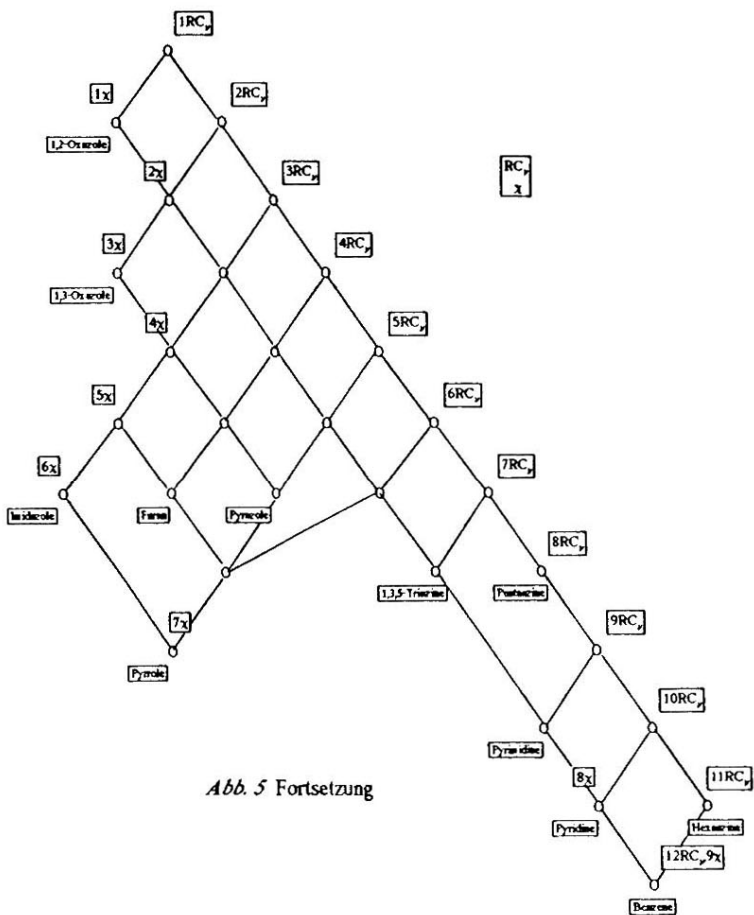


Abb. 5 Fortsetzung

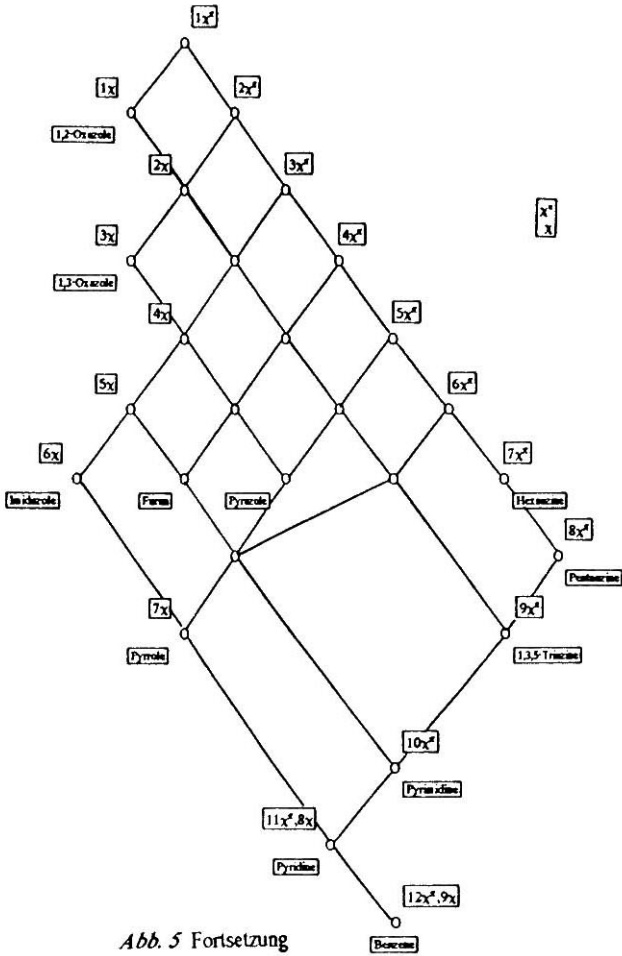


Abb. 5 Fortsetzung

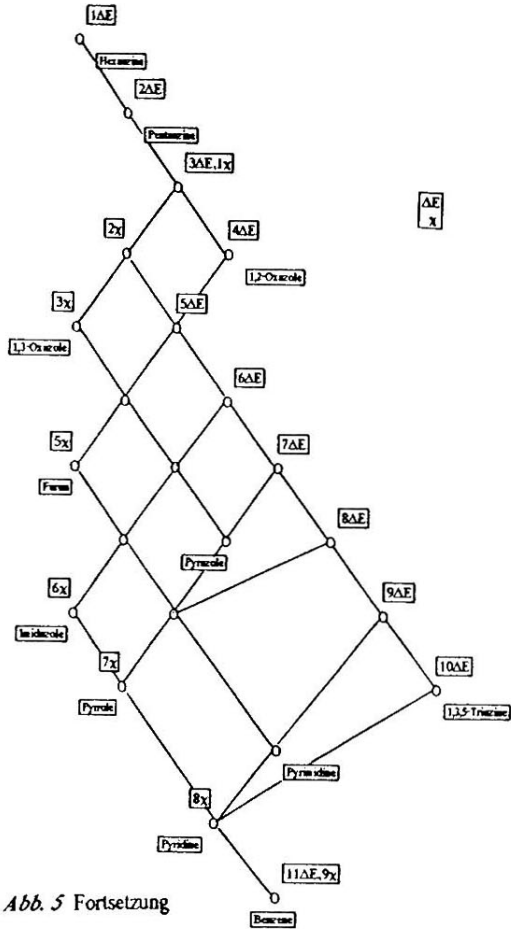


Abb. 5 Fortsetzung

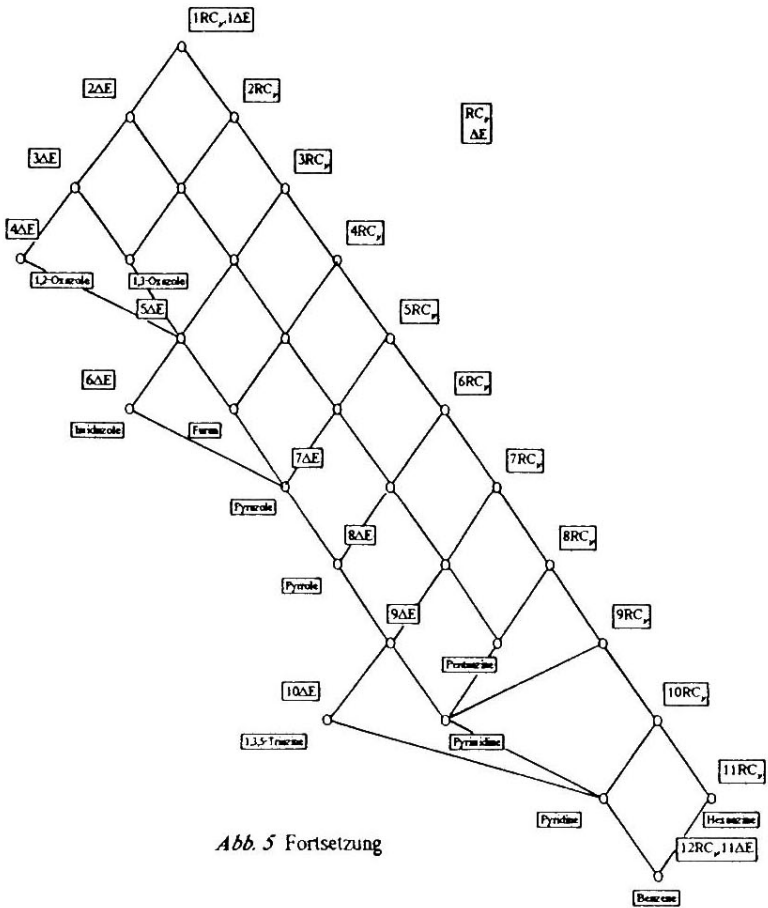


Abb. 5 Fortsetzung

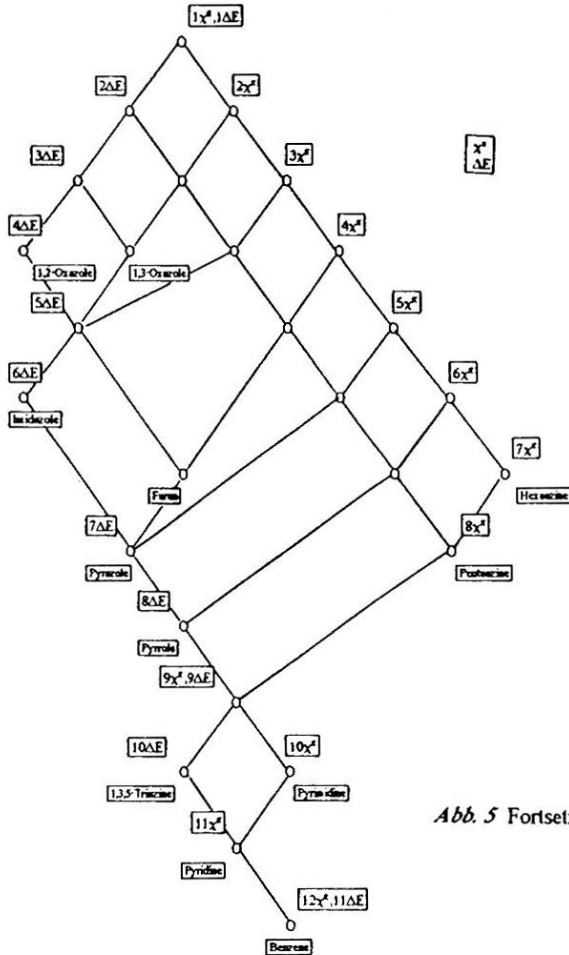


Abb. 5 Fortsetzung

enthält. Die ihnen entsprechenden Begriffsverbände sind in Form ihrer Hasse-Diagramme in den Abbildungen 5 gegeben. Das jeweilige Merkmalspaar wird in einer rechteckigen Rahmung angezeigt.

6. Gewonnene Aussagen

Da die Anzahl der Gegenstände in den sechs Teilkontexten stets kleiner ist als die der Merkmale, können die Aussagen des 2. Abschnitts über Begriffsverbände mit zwei eindimensional ordinal skalierten Merkmalen zur Anwendung kommen. Reine Ordinalität entspricht also Eindimensionalität bzw. Gleichgerichtetheit, während der andere Grenzfall, die Interordinalität, auf Zweidimensionalität bzw. "Orthogonalität" der verglichenen Merkmale hinweisen würde. In dieser Hinsicht können die sechs Paare von Aromatizitätskriterien in drei Klassen eingeteilt werden, wie man aus den Hasse-Diagrammen der Abbildungen 5 ersehen kann:

Klasse 1 : Sie enthält die im wesentlichen gleichgerichteten Kriterien RC_v/χ^π als einzige Kombination.

Klasse 2 : Hierher gehören alle drei Paare, die mit dem Kriterium χ gebildet werden und wesentliche Zweidimensionalität zeigen: $\Delta E/\chi$, RC_v/χ , χ^π/χ .

Klasse 3 : Sie ist durch eine Zweidimensionalität gekennzeichnet, welche im Gegensatz zu derjenigen der Klasse 2 als schwach bezeichnet werden kann, da sie von nur zwei Gegenständen, den Aromaten Penta- und Hexaazin, verursacht wird. Hierher gehören die beiden Kombinationen von ΔE mit den anderen theoretisch ermittelten Kriterien: $\Delta E/RC_v$, $\Delta E/\chi^\pi$.

Diese Resultate erhält man auch bei der Betrachtung des Begriffsverbandes des gesamten Kontexts (Abb. 4), der eine Dreidimensionalität augenscheinlich werden läßt, die von den beiden Kriterien χ und ΔE sowie der Kombination RC_v/χ^π gewissermaßen "aufgespannt" wird. Beim Weglassen der beiden extremen Verbindungen Penta- und Hexaazin würde im wesentlichen wieder eine Zweidimensionalität hergestellt werden, welche von dem experimentell-magnetischen Kriterium χ (Klasse B, s.o.) einerseits und der Klasse A der drei theoretisch-quantenchemischen Kriterien andererseits verursacht wird. Sieht man also von "ungewöhnlichen" Aromaten bei der Untersuchung ab, d.h. von solchen, bei denen das energetische Kriterium stärkere Umstellungen gegenüber den anderen quantenchemischen Kriterien aufzeigt, so ist es vorteilhafter, die Kriterien in theoretisch-quantenchemische (Klasse A) und experimentell-magnetische (Klasse B) einzuteilen. Der Unterschied ergibt sich offensichtlich aus einer stärker abweichenden Reihenfolge der betrachteten Aromaten in Hinblick auf die jeweilige (ordinale) "Stärke" der Aromatizität.

Betrachtet man die beiden Aromaten mit mehr als drei N-Atomen als eine besondere Klasse, so lassen sich die restlichen zehn Verbindungen, wie aus der Abbildung 4 ersichtlich ist, bezüglich der beiden Kriteriengruppierungen A und B, so in Klassen zerlegen, wie es die Tabelle 4 zeigt, wobei die Kriterienwerte jeweils in die drei Bereiche "klein", "mittel" und "groß" eingeteilt wurden.

Tabelle 4 Klassifikation der Aromaten mit weniger als vier N-Atomen

		Kriterium (A)		
		klein	mittel	groß
Kriterium (B)	klein	1,2-Oxazol 1,3-Oxazol		1,3,5-Triazol
	mittel		Furan Pyrrol Imidazol Pyrazol	Pyrimidin
	groß			Pyridin Benzen

Vergleicht man die hier erhaltenen Ergebnisse bzw. Aussagen etwa mit denen, welche JUG und KÖSTER² erhielten, so ergibt sich im wesentlichen eine Übereinstimmung. Die Anwendung der FBA erlaubt aber eine andere Sicht auf das untersuchte Datenmaterial, welche einige Zusammenhänge klarer hervortreten läßt.

Da es die Absicht der vorgelegten Arbeit war, die grundlegenden Züge und Möglichkeiten der FBA vorzustellen, kann auf eine vergleichende Betrachtung oder gar Wertung dieser Methode mit der von JUG und KÖSTER sowie KATRITZKY et al. angewendeten Hauptkomponentenanalyse verzichtet werden, zumal eine solche - allerdings in Verbindung mit einer medizinischen Fragestellung - bereits vorliegt¹³.

7. Anhang: Demonstrationsbeispiele zur Erläuterung der FBA

Um die Definitionen und das Wesen der FBA zu illustrieren, sei das folgende kleine Beispiel angefügt. Dabei sollen 13 organische Verbindungen als die zu betrachtende Gegenstandsmenge herangezogen werden:

$G_1 = \{\text{Essigsäure, Benzoessäure, Ethanol, Methylamin, Phenol, Anilin, Milchsäure, Salicylsäure, Glycin, Phenylalanin, Serin, Tyrosin, Benzen}\}$.

Die vier funktionellen Gruppen COOH, OH, NH₂ und C₆H₅ (= Ph, Phenyl) dienen als Merkmale, d.h.

$$M_1 = \{\text{COOH, OH, NH}_2, \text{Ph}\},$$

so daß die Relation $g_1 I_1 m_1$ lautet:

$g_1 I_1 m_1 =$ "Die Verbindung g_1 aus der Menge G_1 enthält in ihrer Struktur zumindest einmal die funktionelle Gruppe m_1 aus der Menge M_1 ".

Der aus G_1 , M_1 und I_1 gebildete (einwertige) Kontext BEISPIEL1 = (G_1, M_1, I_1) läßt sich in Form der Tabelle 5 aufschreiben, wobei ein "X" anzeigt, daß die Relation $g_1 I_1 m_1$ wahr ist, d.h., die funktionelle Gruppe m_1 in der Verbindung g_1 vorhanden ist.

Tabelle 5 Kontext BEISPIEL1

	COOH	OH	NH ₂	Ph
Essigsäure	X			
Benzoessäure	X			X
Ethanol		X		
Methylamin			X	
Phenol		X		X
Anilin			X	X
Milchsäure	X	X		
Salicylsäure	X	X		X
Glycin	X		X	
Phenylalanin	X		X	X
Serin	X	X	X	
Tyrosin	X	X	X	X
Benzen				X

Wie die Berechnungen mit der FBA zeigen, lassen sich aus diesem Kontext 14 formale Begriffe finden, von denen als Beispiele $(A_{11}, B_{11}) = (\{\text{Essigsäure, Benzoessäure, Milchsäure, Salicylsäure, Glycin, Phenylalanin, Serin, Tyrosin}\}, \{\text{COOH}\})$, $(A_{12}, B_{12}) = (\{\text{Milchsäure, Salicylsäure, Serin, Tyrosin}\}, \{\text{COOH, OH}\})$ und $(A_{13}, B_{13}) = (\{\text{Serin, Tyrosin}\}, \{\text{COOH, OH, NH}_2\})$ genannt seien. Der Begriff (A_{11}, B_{11}) enthält in seinem Umfang A_{11} alle Carbonsäuren, d.h. alle Verbindungen aus G_1 , die das Merkmal COOH und nur dieses gemeinsam haben, so daß der Inhalt dieses Begriffs $B_{11} = \{\text{COOH}\}$ ist. (A_{12}, B_{12}) nennt die vier Hydroxycarbonsäuren in seinem Umfang, sein Inhalt folglich die beiden Merkmale COOH und OH. Die Carbonsäuren stellen einen Oberbegriff zu den Hydroxycarbonsäuren dar, denn es ist $A_{11} \supset A_{12}$ und gleichbedeutend $B_{12} \supset B_{11}$. Analog sind die Hydroxyaminocarbonsäuren (A_{13}, B_{13}) ein Unterbegriff zu den Hydroxycarbonsäuren und somit auch zu den Carbonsäuren.

In der Abbildung 6 wird das Hasse-Diagramm des Begriffsverbandes zum Kontext BEISPIEL1 gezeigt, aus welchem sich außer den soeben erwähnten Begriffen und Ober/Unterbegriffsbeziehungen alle weiteren ablesen lassen, welche sich aus dem betrachteten Kontext ergeben. Bezüglich der Vervollständigung

der einzelnen Begriffsumfänge und -inhalte sei auf die Bemerkungen am Ende des ersten Abschnitts hingewiesen.

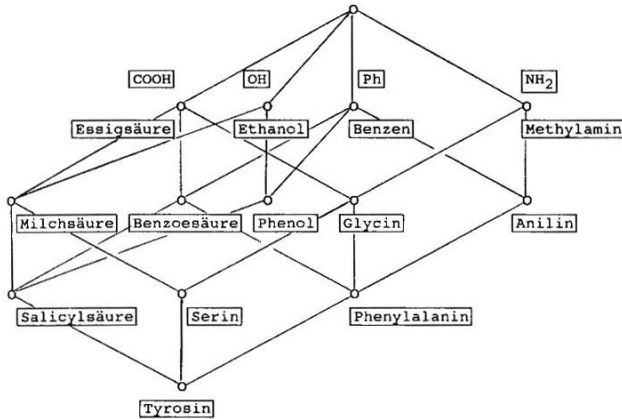


Abb. 6 Hasse-Diagramm des Begriffsverbandes zum Kontext BEISPIEL1

Um auch die Bearbeitung mehrwertiger Kontexte zu demonstrieren, möge von dem zweiwertigen Kontext BEISPIEL2 = (G_2, M_2, W_2, I_2) ausgegangen werden, welcher in Form der Tabelle 6 gegeben ist. Die Gegenstandsmenge G_2 wird also wiederum von chemischen Verbindungen,

$G_2 = \{\text{Benzen, Diphenyl, Essigsäure, Benzoesäure, Salicylsäure, } \alpha, \gamma\text{-Diphenylacetessigsäure, Oxalsäure, Äpfelsäure, Weinsäure, Phthalsäure, Phenyläpfelsäure, Benzilsäure, } \beta\text{-Resorcylsäure, Diphenylsäure, Glykol}\}$,
und die Merkmalsmenge M_2 von funktionellen Gruppen, die in den organischen Molekülen als Strukturen auftreten können,

$$M_2 = \{\text{Ph, OH, COOH}\}$$

mit $\text{Ph} = \text{C}_6\text{H}_5$ (Phenyl), gebildet. Die Menge der Merkmalswerte W_2 enthält zwei Elemente,

$$W_2 = \{1, 2\},$$

die die Anzahl der jeweiligen funktionellen Gruppen angeben, wenn das geordnete Tripel $(g_2 \in G_2, m_2 \in M_2, w_2 \in W_2)$ ein Element der dreistelligen Relation I_2 ist:
 $(g_2, m_2, w_2) \in I_2 =$ "Das Molekül der chemischen Verbindung g_2 aus G_2 enthält die funktionelle Gruppe m_2 aus M_2 w_2 -mal, wobei w_2 eine Zahl aus W_2 bedeutet".

Tabelle 6 Der zweiwertige Kontext BEISPIEL2

	Abk.	Anzahl		
		Ph	OH	COOH
Benzen	I	1		
Diphenyl	II	2		
Essigsäure	III			1
Benzoessäure	IV	1		1
Salicylsäure	V	1	1	1
α, γ -Diphenylacetessigsäure	VI	2		1
Oxalsäure	VII			2
Äpfelsäure	VIII		1	2
Weinsäure	IX		2	2
Phthalsäure	X	1		2
Phenyläpfelsäure	XI	1	1	2
Benzilsäure	XII	2	1	1
β -Resorcyllsäure	XIII	1	2	1
Diphensäure	XIV	2		2
Glykol	XV		2	

Tabelle 7 Die eindimensional ordinale Skala SC

	≥ 1	≥ 2
1	X	
2	X	X

Für die Umwandlung dieses Kontexts in einen einwertigen bietet sich eine eindimensional ordinale Skalierung aller drei Merkmale an. Die entsprechende

Skala ist ein einwertiger Kontext $SC = \langle G_S, M_S, I_S \rangle$, in welchem die beiden Werte des Kontexts BEISPIEL2 die Gegenstandsmenge bilden, $G_S = W_2 = \{1, 2\}$. M_S enthält zwei Elemente $M_S = \{\geq 1, \geq 2\}$, so daß die Relation I_S zu lesen ist als $g_S I_S m_S = \text{"}g_S \text{ ist } m_S\text{"}$ mit $g_S \in G_S$ und $m_S \in M_S$. SC läßt sich als Tabelle 7 schreiben. In diesem einfachen Fall gibt es nur die beiden Begriffe $(\{1, 2\}, \{\geq 1\}) \leq (\{2\}, \{\geq 1, \geq 2\})$, das Hasse-Diagramm ist folglich eine aus zwei Knoten bestehende Kette.

Den so skalierten Kontext zeigt die Tabelle 8. In ihm wurden die Merkmale mit einem Symbol "iY" bezeichnet, wobei $i \in \{1, 2\}$ und $Y \in M_2$ gilt. Es bedeutet "hat $\geq i$ funktionelle Gruppen Y in der Struktur". Die Begriffsmenge umfaßt 20 Begriffe. Das Hasse-Diagramm des Begriffsverbandes zeigt Abbildung 7.

Tabelle 8 Der mit SC skalierte Kontext BEISPIEL2

	Abk.	1Ph	2Ph	1OH	2OH	1COOH	2COOH
Benzen	I	X					
Diphenyl	II	X	X				
Essigsäure	III					X	
Benzoesäure	IV	X				X	
Salicylsäure	V	X		X		X	
α, γ -Diphenylacetessigsäure	VI	X	X			X	
Oxalsäure	VII					X	X
Äpfelsäure	VIII			X		X	X
Weinsäure	IX			X	X	X	X
Phthalsäure	X	X				X	X
Phenyläpfelsäure	XI	X		X		X	X
Benzilsäure	XII	X	X	X		X	
β -Resorcylnsäure	XIII	X		X	X	X	
Diphensäure	XIV	X	X			X	X
Glykol	XV			X	X		

Die Darstellung des Begriffsverbandes wurde so gewählt, daß die mit den Merkmalen Ph, OH und COOH verbundene "Dreidimensionalität" sichtbar wird. Diese Art von gewissermaßen intuitiver Dimensionalität ist bei der Diskussion der Begriffsverbände in den Abbildungen 4 und 5 gemeint.

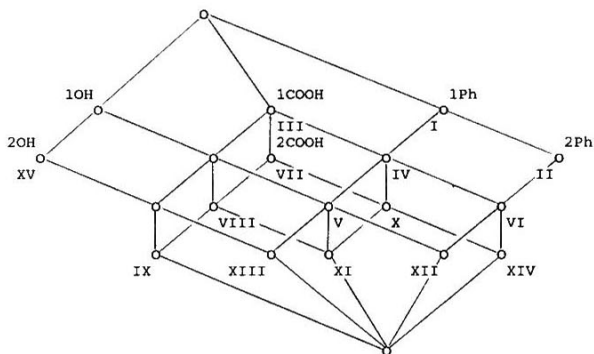


Abb. 7 Hasse-Diagramm zum Begriffsverband des skalierten Kontexts BEISPIEL2
(Abkürzungen s. Tabelle 6 und 8 sowieText)

8. Literatur

- ¹ A.R. KATRITZKY, P. BARCZYNSKI, G. MASUMARRA, D. PISANO and M. SZAFRAN, *J. Am. Chem. Soc.* **111**, 7 (1989)
- ² K. JUG and A.M. KÖSTER, *J. Phys. Org. Chem.* **4**, 163 (1991)
- ³ R. WILLE, *Studien zur Klassifikation* **15**, 32 (1984)
- ⁴ R. WILLE, *Bedeutungen von Begriffsverbänden*, in: B. GANTER, R. WILLE und K.E. WOLFF (Hrg.), *Beiträge zur Begriffsanalyse*, BI-Wissenschaftsverl., Mannheim, 1987, S. 161-211
- ⁵ J. NIPPERDEY, *Begriffsverbände - eine Anwendung in der Chemie*, Fachwiss. Abschlußarbeit, Martin-Luther-Universität Halle-Wittenberg, Sektion Mathematik, Halle/S., 1988
- ⁶ H.-G. BARTEL, *Informatik-Skripten VII*, Humboldt-Universität zu Berlin, Organisations- und Rechenzentrum, Berlin, 1988, S. 72-75
- ⁷ H.-G. BARTEL, *Selbstorganisation* **1**, 99 (1990)
- ⁸ B. GANTER and R. WILLE, *Conceptual Scaling*, in: F. ROBERTS (ed.), *Applications of Combinatorics and Graph Theory in the Biological and Social Sciences*, Springer, New York, 1989, pp. 139-167
- ⁹ L. TALKINGTON, *Syst. Zoology* **16**, 149 (1967)

¹⁰ R.R. SOKAL and P.H.A. SNEATH, *Principles of Numerical Taxonomy*, Freeman and Co., San Francisco, 1963

¹¹ K. JUG and A.M. KÖSTER: *J. Am. Chem. Soc.* **112**, 6772 (1990)

¹² P. BURMEISTER, *CONIMP (BAV3E, Version 3.06B)*, *Programm zur formalen Begriffsanalyse einwertiger Kontexte*, TH Darmstadt, 1991

¹³ N. SPANGENBERG and K.E. WOLFF, *Comparison between Principal Component Analysis and Formal Concept Analysis, Analysis of Repertory-Grids*, in: W. LEX (Hrg.), *Arbeitstagung Begriffsanalyse und Künstliche Intelligenz*, TU Clausthal, Clausthal-Zellerfeld, 1988, S. 127-134