TOSCANA — ein Werkzeug zur begrifflichen Analyse und Erkundung von Daten

Wolfgang Kollewe, Martin Skorsky,
Frank Vogt, Rudolf Wille

Inhalt

1 Probleme der Datenanalyse und –erkundung

Die zunehmende Komplexität in unserer Welt findet ihren Niederschlag in immer mehr Daten, die es zu erkunden und zu analysieren gilt. Deshalb werden in großem Umfang Methoden und Instrumente der Datenanalyse und -erkundung entwickelt und eingesetzt. Die Fülle der dabei zu bewältigenden Probleme und der vorherrschende Drang nach Automation bringt immer wieder mit sich, daß die entwickelten Systeme die in den Daten kodierten Informationen unangemessen verkürzen und manipulieren. Gebraucht werden problem- und menschengerechte Werkzeuge der Daten- und Wissensverarbeitung, die komplexe Zusammenhänge auf angemessene Weise transparent und verstehbar machen können. Das Programm TOSCANA, das in diesem Artikel vorgestellt werden soll, ist als ein solches Werkzeug entwickelt worden.

Was sind nun die zu bewältigenden Probleme der Datenanalyse und -erkundung? Antworten hierauf sollen zunächst an zwei Beispielen diskutiert werden. Als erstes wählen wir ein Beispiel aus der medizinischen Datenanalyse: die *Auswertung von Repertory Grid Befragungen eßgestörter Patientinnen* (s. [Spangenberg 1990]). Die Daten, die Ergebnis einer derartigen Befragung sind, liegen in der Regel in Form einer Zahlentabelle vor. Die Zahlen in einer Tabellenzeile geben für ein Gegensatzpaar von Adjektiven an, wie die Patientin Personen ihrer engeren sozialen Umgebung bezüglich dieses Gegensatzpaares bewertet. Ziel der Analyse solcher Daten ist, die sozialen Konflikte bewußt zu machen, die die Patientin psychisch belasten. Da die Datentabellen durchaus mehr als 200 Zahlen enthalten können, ist eine formale Auswertung nötig, um die mit den Daten erfaßten Zusammenhänge offenzulegen und eine stimmige Interpretation zu ermöglichen. Damit wird eines der zentralen Probleme der Datenanalyse angesprochen, und zwar die Frage: *Wie kann man vorliegende Daten formal so auswerten, daß die Dateninterpretation in Hinblick auf den gesetzten Zweck möglichst weitgehend unterstützt, jedoch nicht negativ beeinflußt wird?* Für die Auswertung von Repertory Grid Befragungen werden überwiegend Methoden der *Multivariaten Stastistik* angewandt, mit denen die Zahlentabellen verrechnet werden, um eine geometrische Darstellung der Zusammenhänge zu erhalten; dabei werden nur zu oft die in den Daten kodierten Informationen unzulässig verkürzt und nicht selten Artefakte produziert (vgl. [Spangenberg, Wolff 1991]). Will man derartige Probleme vermeiden, muß darauf geachtet werden, daß Begriffe und Beziehungen, an die die Daten im Rahmen der Zwecksetzung gebunden sind, auch nach der formalen Auswertung noch rekonstruiert werden können. Das ist auch deshalb wichtig, weil eine solche inhaltliche Rekonstruktion das Feststellen und Korrigieren von Fehlern in den Ausgangsdaten erleichtert; so kann bei der Diskussion einer unverkürzten Darstellung der Zusammenhänge die betroffene Patientin als fehlerhaft eingeschätzte Bewertungen immer noch erkennen und abändern. Ein Grund für die Unbrauchbarkeit von Auswertungen liegt häufig in der Mißachtung des sogenannten *„Meßniveaus“* der Daten, das die Zulässigkeit von Auswertungsverfahren einschränkt; da die Patientinnen bei den Repertory Grid Befragungen Bewertungen auf ordinalem Niveau angeben, lassen sich die Methoden der Linearen Algebra, wie sie die Multivariate Statistik verwendet, nicht rechtfertigen. Hier wird das allgemeine Problem sichtbar, daß stets zu reflektieren ist: *Wieweit und in welcher Art liegen qualitative bzw. quantitative Daten vor und welche Auswertungsverfahren dürfen angewandt werden?* Zur Diskussion dieses und verwandter Probleme soll hier stellvertretend auf [Orth 1974], [Lamnek 1988] und [Henning 1994] verwiesen werden.

Als zweites Beispiel soll die *Literaturrecherche in einer Bibliothek* diskutiert werden. Traditionell liefern dafür die Bibliothekskataloge die Daten, die überwiegend nach Autoren oder nach Such- bzw. Schlagworten angeordnet sind. Zunehmend werden bibliographische und die zugehörigen kategorialen Daten mit Retrievalprogrammen zugänglich gemacht, die eine automatische Suche in Autoren- und Schlagwortindex sowie Freitextsuche gestatten, wobei auch logische Kombinationen der Suchbegriffe gebildet werden können (vgl. [Gödert 1988]). Das zentrale Problem ist natürlich: *Wie kann eine Rechercheaufgabe mit bestmöglichem Erfolg gelöst werden?* Eine der Hauptschwierigkeiten bei der Recherche ist die Umsetzung der semantischen Vorstellungen zu einem Thema in die syntaktische Suche nach Zeichenketten. Ist diese Umsetzung zu eng gewählt, erhält man zu wenig Information; ist sie zu weit, wird eine zu hohe Trefferquote erzielt. Trefferzahlen werden häufig noch dadurch gesteigert, daß auch Teilabschnitte der eingegebenen Zeichenfolge zur Dokumentsuche herangezogen werden und Suchworte verschiedene Bedeutungen haben können. Alles in allem ist es nicht leicht, eine angemessene Menge an Information zu bekommen, die dem vorgegebenen Thema gerecht wird und sich brauchbar differenzieren läßt. Schwierigkeiten bereitet ferner, daß bei der Fülle der erfaßten Texte und Dokumente die gespeicherten Daten auch bei sorgfältiger Verschlagwortung nur unvollkommen das Inhaltliche wiedergeben, auf das viele Recherchen zielen. Noch grundsätzlicher ist das Problem, daß viele Bibliotheksbenutzer nur ungenau sagen können, was sie suchen. Erst im Prozeß der Erkundung lernen sie zunehmend das zu präzisieren, was sie wollen. Dieser Lernprozeß wird in der Regel durch einzelne Suchworte zu wenig unterstützt. Erst reichhaltige *Begriffsnetze*, die vielfältige Zusammenhänge sichtbar machen, können die wünschenswerte Orientierung im notwendigen Umfang liefern (vgl. [Vester 1988]).

An den zwei Beispielen ist schon ein großes Spektrum an Problemen der Analyse und Erkundung von Daten deutlich geworden, das sich noch auf vielfältige Weise erweiteren und vertiefen ließe. Bei fast allen diesen Problemen wird die Frage berührt: *Wie kann die notwendige formale Behandlung reichhaltiger Daten mit dem inhaltlichen Denken der betroffenen Menschen angemessen verbunden werden?* Antworten auf diese Frage werden immer dringlicher angesichts des starken Informations- und Wissenszuwachses in einer komplexer werdenden Welt, in der die Datenmengen in Umfang und Unüberschaubarkeit weiter zunehmen (Beleg hiefür ist die jüngste Ausschreibung der NASA, Projekte für die Aufbereitung und Behandlung ihrer riesigen Datenmengen vorzuschlagen). Die Lösung in einer weiteren Computerisierung zu sehen, ist zweischneidig, da sie den Trend einer fortschreitenden Mechanisierung des Denkens fördert, was die Gefahr eines Abbaus menschlicher Selbstbestimmung und Verantwortung in sich birgt. Daher gilt es, menschengerechte Methoden begrifflicher Datenanalyse und -erkundung zu entwickeln und einzusetzen, bei denen der Computer nur in der Rolle des beherrschbaren Werkzeugs verwendet wird. Derartige Methoden sind im Forschungsgebiet der *Formalen Begriffsanalyse* entstanden und in den letzten Jahren durch das Programm TOSCANA für ein breites Anwendungsspektrum verfügbar gemacht worden.

2 Analyse von Zusammenhängen: ein TOSCANA-Beispiel

In diesem und dem nächsten Abschnitt soll das Programm TOSCANA mit je einem Anwendungsbeispiel vorgestellt werden, um danach substantieller über die allgemeine Konzeption von TOSCANA informieren zu können. Als erstes wird über eine Analyse von *Produktdaten von PCs* berichtet, die zu Demonstrationszwecken für die CeBIT'94 erstellt wurde. Die Daten stammen aus der Zeitschrift [PC Magazine 1993] und beschreiben 74 PCs (mit 486/66 Prozessor) durch Eigenschaften wie Preis, Festplattengröße, Leistungsfähigkeit bei verschiedenen Tests, Gehäuseformen, Bustyp sowie Art und Anzahl der Schnittstellen. Die an-

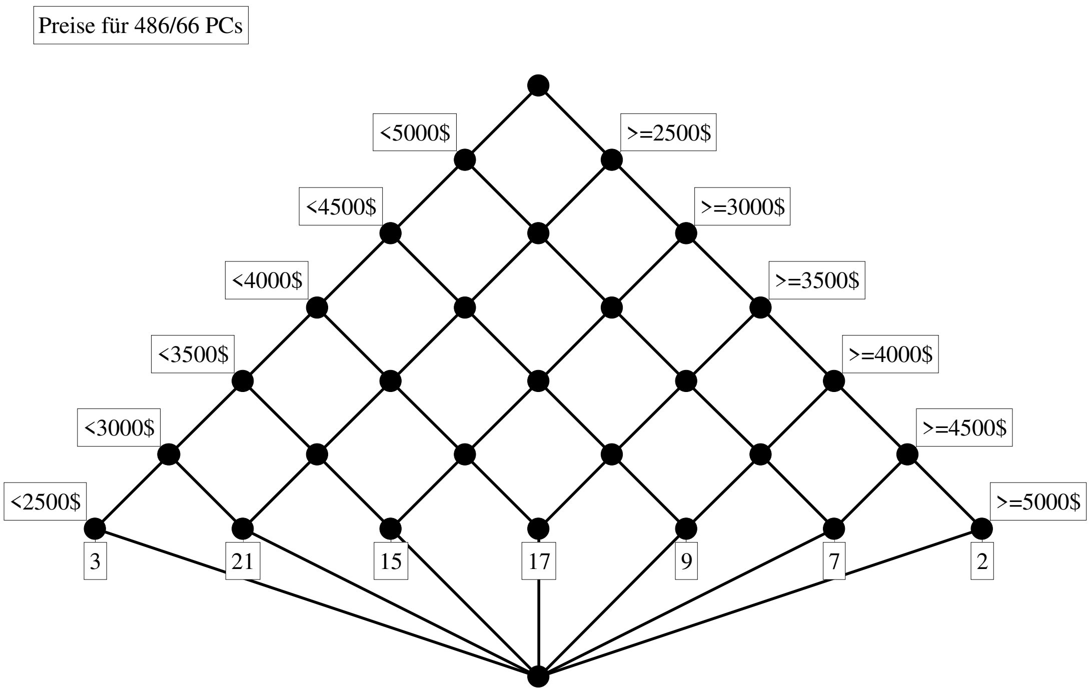

Abbildung 1: Preisklassen bei 486/66-PCs dargestellt durch ein Liniendiagramm mit Kontingentszahlen

gegebenen Ausprägungen der Eigenschaften sind entweder Anzahlen, Meßwerte, Ja/Nein-Urteile oder Benennungen (z. B. hat die Eigenschaft 'Gehäuseform' die Ausprägungen 'Desktop', 'Tower', 'Minitower'). *Aufgabe und Zweck der Analyse* ist die Unterstützung von Produktauswahl und Kaufentscheidung. Angemerkt werden soll, daß derartige Analysen auch mit Vorteil bei der Plazierung von Produkten im Markt genutzt werden können; dazu hat man die marktrelevanten Eigenschaften herauszuarbeiten, durch welche sich das eigene Produkt von anderen Produkten abhebt.

Um die Arbeit mit TOSCANA im Grundansatz demonstrieren zu können, genügt es, sich auf einen Ausschnitt der vorliegenden Daten zu beschränken. Im folgenden sollen deshalb nur die Eigenschaften 'Preis' und 'Bustyp' behandelt werden. Zunächst werden graphische Darstellungen für die einzelnen Eigenschaften diskutiert, woraus dann die Zusammenschau beider Eigenschaften gewonnen wird. Grundlegend sind dabei die graphischen Darstellungen durch *Liniendiagramme*, die seit vielen Jahren in der Formalen Begriffsanalyse erfolgreich eingesetzt werden (vgl. [Wille 1984], [Wille 1989]). Es ist eine der Hauptleistungen von TOSCANA, aufgrund gespeicherter Daten informative Liniendiagramme zu

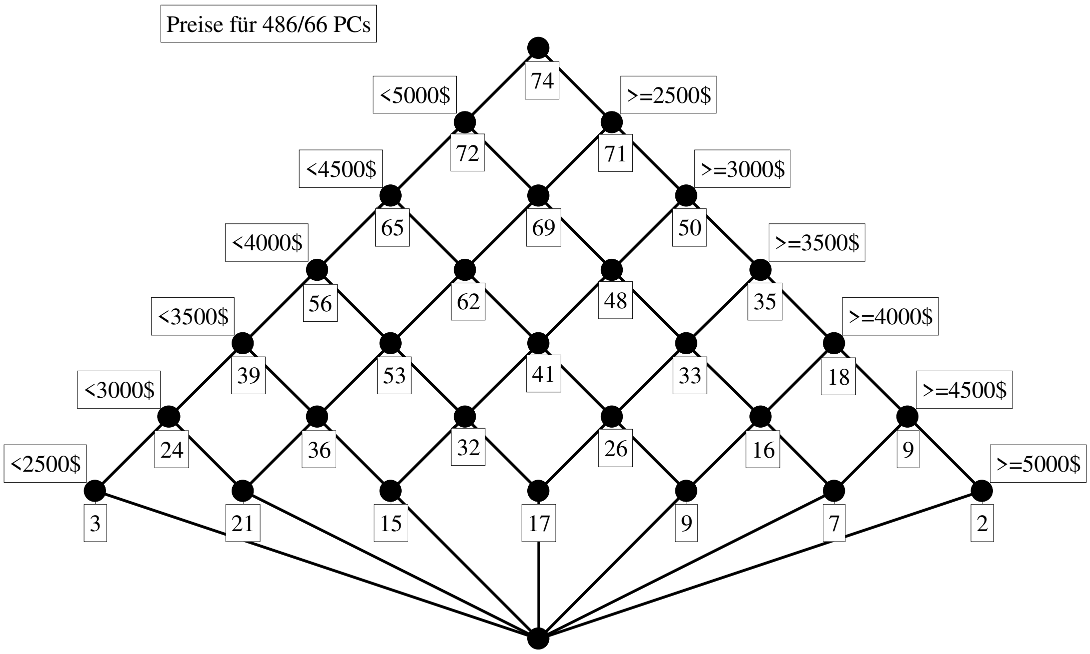

Abbildung 2: Preisklassen bei 486/66-PCs dargestellt durch ein Liniendiagramm mit Umfangszahlen

produzieren und damit Begriffshierarchien zu veranschaulichen.

Abbildung 1 zeigt ein Liniendiagramm für die Eigenschaft 'Preis' in einem Originalausdruck von TOSCANA. An dem Diagramm kann man die Verteilung der 74 PCs auf verschiedene Preisklassen ablesen: 3 PCs haben einen Dollar-Preis kleiner als 2500, 21 zwischen 2500 und 3000, 15 zwischen 3000 und 3500, 17 zwischen 3500 und 4000, 9 zwischen 4000 und 4500, 7 zwischen 4500 und 5000 sowie 2 von 5000 und mehr. Damit sind aber nur die Preisintervalle genannt, die durch die schwarzen Kreise dargestellt werden, an denen Anzahlen der PCs stehen. Auch die übrigen Kreise (außer dem untersten) stellen Preisintervalle dar. Die obere Grenze des zu einem Kreis gehörigen Preisintervalls ist der niedrigste Preis, den man von dem Kreis durch einen aufsteigenden Streckenzug auf der linken Seite des Diagramms erreicht; die untere Grenze ist entsprechend der höchste Preis erreichbar auf der rechten Seite. So erhält man beispielsweise für den Kreis in der Mitte zwischen oberstem und unterstem Kreis das Preisintervall von 3000$ bis 4500$. Die Anzahl der PCs, die in ein solches Preisintervall fallen, ist gerade die Summe der Anzahlen, die man von dem Kreis durch einen absteigenden Streckenzug erreicht; im Fall unseres Beispielkreises erhält man so die Anzahl 41.

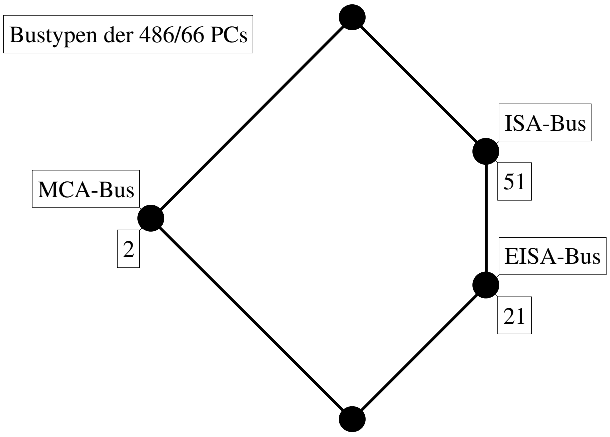

Abbildung 3: Bustypen bei 486/66-PCs

Auch wenn die Anzahlsummen in einem derart übersichtlichen Liniendiagramm einfach zu bilden sind, ist es für manche Zwecke doch vorteilhaft, wenn diese Zahlen wie in Abbildung 2 (ebenfalls Originalausdruck von TOSCANA) schon im Liniendiagramm eingetragen sind. In der Terminologie der Formalen Begriffsanalyse werden Anzahlen wie in Abbildung 1, wo jeder PC in genau eine der Anzahlen eingeht, *Kontingentszahlen* genannt; die Anzahlen wie in Abbildung 2, wo jeder PC entsprechend seiner Zugehörigkeit in eine oder mehrere Anzahlen eingeht, werden *Umfangszahlen* genannt.

In Abbildung 3 wird ein Liniendiagramm für die Eigenschaft 'Bustyp' gezeigt, die in den vorliegenden Produktdaten drei Ausprägungen hat: ISA (Industry Standard Architecture), EISA (Extended Industry Standard Architecture) und MCA (Micro Channel Architecture). Die für das Diagramm gewählte Begriffshierarchie drückt aus, daß ein EISA-Bus alle Leistungsmerkmale eines ISA-Bus besitzt und dazu noch mehr leistet; der MCA-Bus ist dagegen mit den beiden anderen Typen inkompatibel. Da in Abbildung 3 die Kontingentszahlen angezeigt werden, läßt sich am Liniendiagramm direkt ablesen, wieviele PCs welchen Bustyp haben. Da ISA-Steckkarten in einen PC mit EISA-Bus eingebaut werden können, zeigen die Umfangszahlen direkt, wieviele Rechner für die Ausrüstung mit ISA-, EISA- oder MCA-Steckkarten geeignet sind. Zieht man anstelle der Bustypen die Anzahlen der Steckplätze vom jeweiligen Typ heran, so ist eine differenziertere Analyse von Anzahl und Art der jeweils passenden Steckkarten

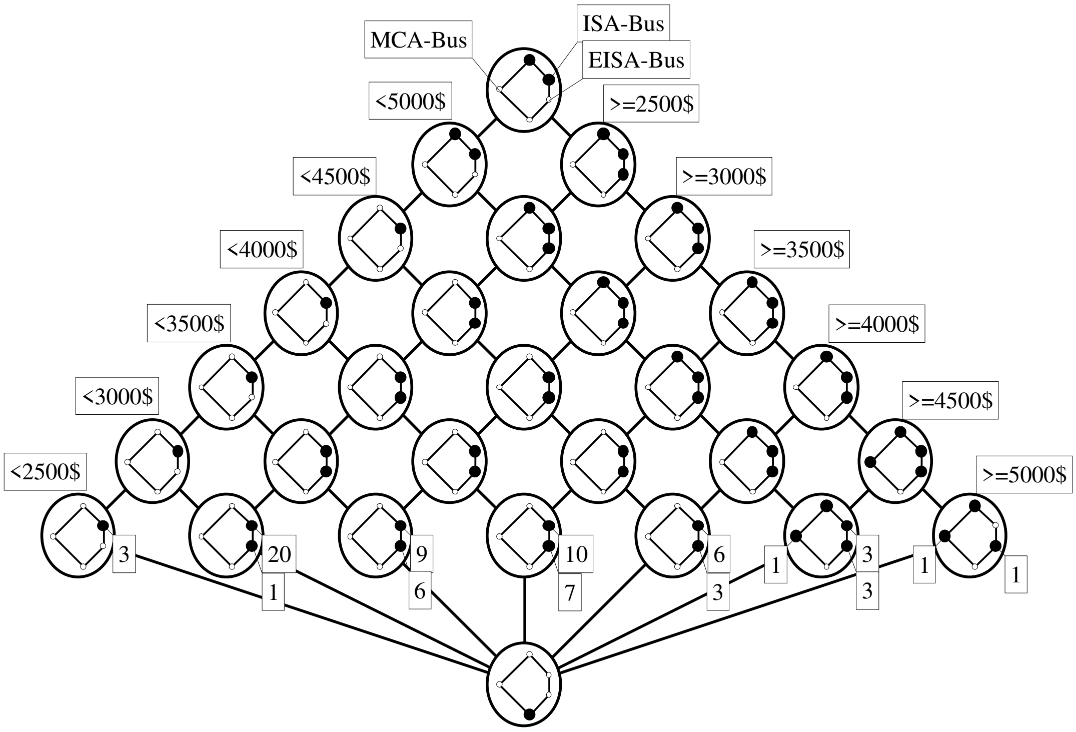

Abbildung 4: Zusammenhang von Preis und Bustyp bei 486/66-PCs

möglich, die jedoch den Rahmen dieser Ausführungen sprengen würde.

Die bisher gezeigten Liniendiagramme stellen sehr einfach strukturierte Begriffshierarchien dar, so daß man vielleicht nicht einsieht, wozu überhaupt solche Darstellungen nützlich sind. Dazu ist einerseits zu sagen, daß Begriffshierarchien in der Regel komplexer ausfallen und die einfachen Strukturen hier nur zum Zweck der Einführung gewählt worden sind. Zum anderen werden die Diagramme sofort reichhaltiger, wenn man Zusammenhänge zwischen mehreren Eigenschaften veranschaulichen will. Demonstriert werden soll das an der Frage: Wie hängt der Bustyp vom Preis ab? Für die Beantwortung derartiger Fragen haben sich zwei- und mehrstufige Liniendiagramme bewährt, in denen die Einzeldiagramme kombiniert werden. Mit TOSCANA können solche Kombinationen von Diagrammen mit großer Flexibilität durchgeführt werden. In Abbildung 4 wird ein zweigestuftes Liniendiagramm in einem Originalausdruck von TOSCANA wiedergegeben. In diesem Diagramm bildet offenbar das Liniendiagramm der Eigenschaft 'Preis' die *Grobstruktur*, während das Liniendiagramm zum 'Bustyp' jeweils die *Feinstruktur* liefert. Die Begriffshierarchie wird dabei nur durch die schwarzen Kreise dargestellt; die übrigen kleinen Kreise zeigen an, für welche Kombinationen von Ausprägungen es keine PCs gibt. Unschwer liest man am Diagramm ab, wieviele

PCs in den einzelnen Preisklassen welchen Bustyp haben. Hier bestätigt sich der plausible Zusammenhang zwischen Preis und Leistungsfähigkeit eines Bustyps; so trifft der weniger leistungsfähige ISA-Bus bei den niedrigeren Preisklassen häufiger auf. Ebenso ist ablesbar, daß der MCA-Bus, der eine Exotenrolle einnimmt, nur bei den hohen Preisklassen vorkommt.

Allgemein kann man mit TOSCANA durch Kombination geeigneter Eigenschaften die für den gesetzten Zweck wichtigen Zusammenhänge herausfinden, was durch eine auf Interaktion angelegte Benutzeroberfläche auf vielfältige Weise unterstützt wird. TOSCANA gestattet auch, in allen Diagrammen die Namen der PCs einzublenden. Damit kann auf jedes der untersuchten Einzelobjekte zugegriffen werden, wie es bei der Verwendung dieses *Begrifflichen Datensystems* zur Unterstützung von Produktauswahl und Kaufentscheidung notwendig ist. Im Fall der Produktauswahl besteht für den Benutzer eine Strategie darin, die für ihn relevanten Eigenschaften in der Reihenfolge ihrer Wichtigkeit anzuordnen und dann die Eigenschaften von der wichtigsten zur unwichtigsten abzuarbeiten. Für jede Eigenschaft wird im zugehörigen Liniendiagramm die gewünschte Kombination von Ausprägungen durch Auswahl eines (Begriffs-)Kreises fixiert. Durch eine derartige Auswahl selektiert TOSCANA alle PCs mit der damit gewählten Kombination von Ausprägungen und geht zum Liniendiagramm der nächsten Eigenschaft über. Auf diese Weise wird die Menge der noch möglichen Objekte sukzessiv verringert bis zur endgültigen Wahl.

3 Erkunden von Dokumenten: ein TOSCANA-Beispiel

Das zweite TOSCANA-Beispiel stammt aus dem Bereich des Bauwesens und zeigt exemplarisch auf, wie ein Bestand von Dokumenten begrifflich erkundet werden kann.

Für die Beurteilung eines Bauvorhabens wie auch für die Weiterentwicklung der gesetzlichen Regelungen ist es von grundsätzlicher Bedeutung, das Zusammenspiel der einzelnen rechtlichen Bestimmungen zu kennen. Die Vielfalt der baurechtlich relevanten Gesetze, Verordnungen und Erlasse wie auch der baubezogenen DIN-Normen ist mittlerweile so groß, daß es selbst für Experten vielfach sehr schwierig und zeitaufwendig ist, alle rechtlichen Zusammenhänge und ihr Zusammenwirken zu berücksichtigen. Aufgrund der aus inhaltlichen Belangen notwendigen Spezialisierung der einzelnen Experten sind die gesamten Zusammenhänge in der Regel nur unvollständig bekannt, so daß präzise Abfragen — wie sie in den Abfragesprachen der gängigen Datenbankverwaltungssysteme gefordert werden — nur in den selteneren Fällen möglich sind. Vielmehr ist ein Navigieren, durch das die interessierenden begrifflichen Zusammenhänge offengelegt werden, in der Wissenslandschaft des Baurechts und der Bautechnik unumgänglich. Das Programm TOSCANA unterstützt dieses Navigieren durch die Bereitstellung geeigneter Fragekomplexe, die jeweils inhaltlich abgegrenzte Teilgebiete in ihren begrifflichen Zusammenhängen als Liniendiagramme visualisieren. Der Bearbeiter bzw. der Auswerter gelangt mit Hilfe der bereitgestellten Fragekomplexe zu dem Wissensausschnitt, der für seine aktuelle Fragestellung von Bedeutung ist.

Mit dem Ministerium für Bauen und Wohnen des Landes Nordrhein-Westfalen wurde in einem Forschungsprojekt ein Prototyp eines Begrifflichen Datensystems aufgebaut, der das Navigieren in der Gesetzeslandschaft des Baurechts und der Bautechnik erlaubt (vgl. [Kollewe et al. 1994]). Da die Bestimmungen und gesetzlichen Regelungen des Baurechts wie auch der Bautechnik ein über Jahrzehnte gewachsenes System sind, das auch sprachlich nicht immer in sich konsistent ist, war es für den zu entwickelnden Prototyp notwendig, ein einheitliches Sprachsystem der entsprechenden Grobheitsstufe, die der Aufgabenstellung angemessen ist, zu entwerfen. In einem diskursiven Prozeß wurden die einzelnen gesetzlichen Regelungen analysiert und entsprechend ihren Inhalten den entsprechenden Kategorien zugeteilt. Hierbei waren vielfach tieferliegende Zusammenhänge zu erkennen und zu berücksichtigen, die bei einer einfachen Auswertung nach Schlüsselworten niemals zutage getreten wären. So entstand eine Kreuztabelle, deren Vorspalte von den einzelnen gesetzlichen Regelungen gebildet wird und deren Kopfzeile aus den Kategorien der neu geschaffenen Metasprache besteht. Die Tabelle bildete die Grundlage, auf der die Fragekomplexe entwickelt wurden. Merkmale, die inhaltlich in einem engen Zusammenhang stehen, wurden zu Fragekomplexen zusammengefügt. Insgesamt wurden über 30 Fragekomplexe für den Wohnhausbau und den Krankenhausbau erarbeitet, mit denen eine inhaltlich detailierte Erkundung des baurechtlichen Dokumentenbestandes durch TOSCANA möglich ist.

Am Anfang eines Erkundungsvorgangs mit TOSCANA sucht der Bearbeiter entsprechend seinen Erkundungszielen, die ihm hierfür geeignet erscheinenden Fragekomplexe aus. Sollte sich während des Erkundungsvorgangs zeigen, daß ein Teil der ausgewählten Fragekomplexe für das Erkundungsziel nicht geeignet war, so lassen sich bei TOSCANA interaktiv die entsprechenden Fragekomplexe entfernen und neue auswählen. Mit den neu ausgewählten kann dann der begonnene Erkundungsvorgang fortgesetzt werden. Im Rahmen dieser Arbeit kann nur exemplarisch ein kurzer Erkundungsvorgang vorgestellt werden, der daher nur eingeschränkt die vielfältigen Möglichkeiten des Programms TOSCANA verdeutlichen kann. Hat ein Architekt oder Bauingenieur zur Aufgabe, eine Station im Krankenhaus zu planen, so wird er sich neben vielen anderen baurechtlichen Bestimmungen notwendigerweise auch mit der Brandsicherheit der Funktionsräume beschäftigen müssen. Der hierfür notwendige Erkundungsvorgang könnte folgendermaßen aussehen:

Der Planer wählt die Fragekomplexe 'Funktionsräume' und 'Betriebs- und Brandsicherheit' aus, von denen er sich zuerst nur den Fragekomplex 'Funktions-

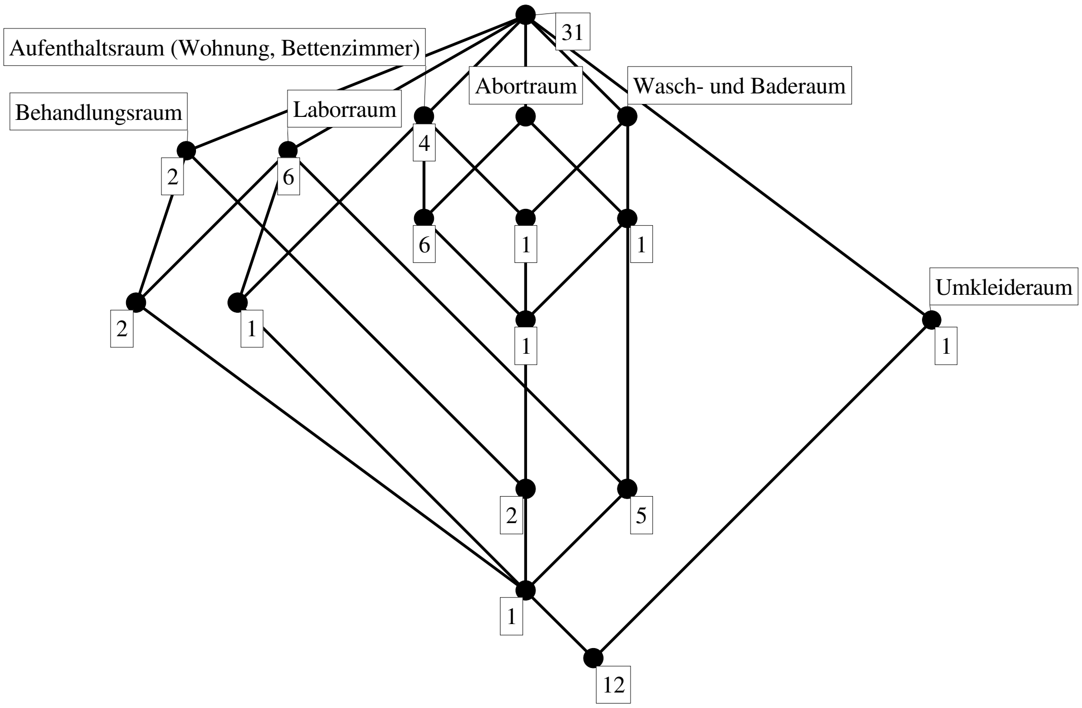

Abbildung 5: Baurechtliche Dokumente zu Funktionsräumen im Krankenhaus angezeigt durch Kontingentszahlen

räume' von TOSCANA anzeigen läßt. Der Fragekomplex 'Funktionsräume' (Abb. 5) im Krankenhaus umfaßt die Merkmale 'Aufenthaltsraum', 'Behandlungsraum', 'Laborraum', 'Abortraum', 'Wasch- und Baderaum' sowie 'Umkleideraum'. Er zeigt die Vernetzung der in das Begriffliche Datensystem einbezogenen rechtlichen Regelungen. Schon bei einem ersten Blick auf diesen Fragekomplex wird deutlich, daß der Umkleideraum eine Sonderrolle gegenüber den anderen Funktionsräumen spielt. Die Anzahlen, die im Liniendiagramm angezeigt werden, geben Auskunft darüber, wieviele gesetzliche Bestimmungen auf die einzelnen Begriffe (Kombinationen von Merkmalen) zutreffen. Beispielsweise gibt es vier baurechtlich relevante Bestimmungen in der Datenbasis, die sich ausschließlich mit dem 'Aufenthaltsraum' befassen. Hingegen gibt es keine Bestimmung die sich ausschließlich mit dem 'Abortraum' beschäftigt. Alle baurechtlichen Gesetze, Verordnungen und Normen, die auf den 'Abortraum' zutreffen, haben auch Bezug zu anderen Funktionsräumen. Insgesamt sind im Dokumentenbestand 12 baurechtliche und bautechnische Regelungen, die für alle Funktionsräume von Bedeutung sind.

Da es für einen Erkundungsvorgang in der Wissenslandschaft des Baurechts und der Bautechnik unbefriedigend wäre, nur über die Anzahlen der relevanten Bestimmungen informiert zu werden, ist es programmtechnisch in TOSCANA

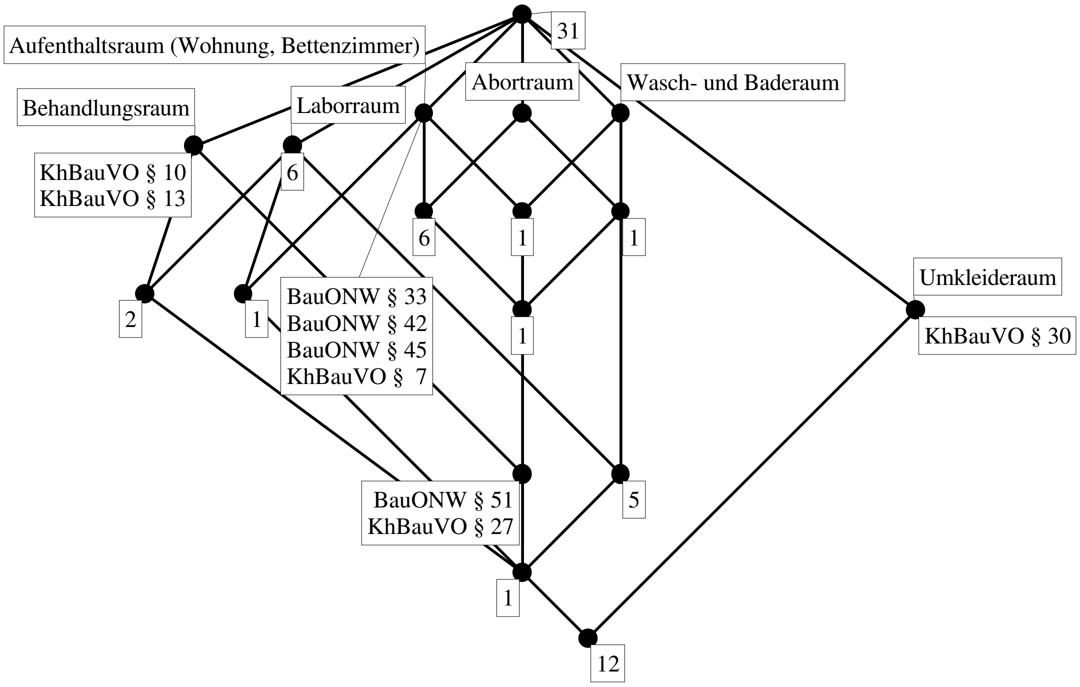

Abbildung 6: Baurechtliche Dokumente zu Funktionsräumen im Krankenhaus (teilweise) angezeigt durch ihre Titel

vorgesehen, sich die Titel der entsprechenden Bestimmungen anzeigen zu lassen. In Abbildung 6 sind für einen Teil der baurechtlichen Bestimmungen die entsprechenden Titel aufgeblättert. Nunmehr lassen sich die baurechtlichen Bestimmungen, die ausschließlich auf den Aufenthaltsraum zutreffen, durch ihre Titel identifizieren. Es sind die Paragraphen 33, 42 und 45 der Bauordnung Nordrhein-Westfalens sowie der Paragraph 7 der Krankenhausbauverordnung.

Der Architekt oder Bauingenieur hat nunmehr zwei Möglichkeiten, in TOSCANA den Fragekomplex 'Funktionsräume' durch den Fragekomplex 'Betriebs- und Brandsicherheit' begrifflich zu verfeinern. Interessiert ihn nur für einen speziellen Begriff die begriffliche Struktur der Brandsicherheit, so kann er sich durch Hineinzoomen für diesen das entsprechende Liniendiagramm anzeigen lassen. In Abbildung 7, einem Originalausdruck von TOSCANA, wurde der unterste Begriff des Liniendiagramms unter der obigen Fragestellung verfeinert. Von den zwölf baurechtlichen Bestimmungen, die für alle Funktionsräume des Krankenhauses relevant sind, befassen sich zwei Bestimmungen mit dem 'Brandschutz', während ein baurechtlich relevanter Text die 'Brandbekämpfung' regelt. Hingegen sind neun der zwölf Gesetze und Verordnungen unter diesem Gesichtspunkt ohne Bedeutung.

Möchte hingegen der Architekt bzw. Bauingenieur für alle Begriffe des Frage-

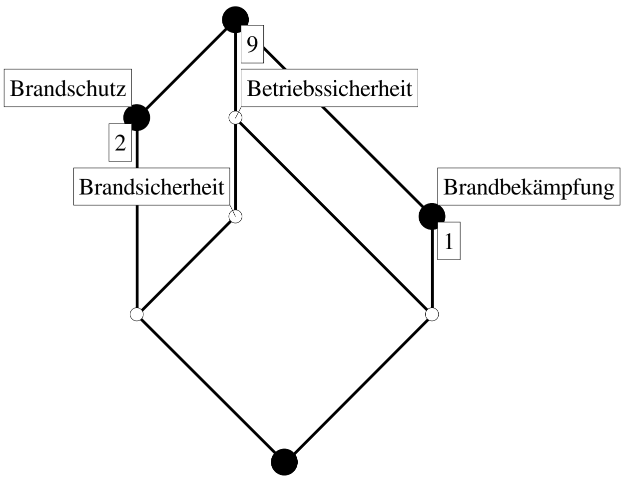

Abbildung 7: Baurechtliche Dokumente zur Betriebs- und Brandsicherheit

komplexes ‘Funktionsräume’ die begriffliche Verfeinerung durch den Fragekomplex ‘Brand- und Betriebssicherheit’ aufgezeigt bekommen, so stellt ihm das Programm TOSCANA die Möglichkeit zur Verfügung, ein zweigestuftes Liniendiagramm anzeigen zu lassen. Abbildung 8 — wie alle übrigen Liniendiagramme ein Originalausdruck von TOSCANA — zeigt diese programmtechnische Möglichkeit. Schaut man sich die unterste Ellipse des äußeren Liniendiagramms an, so findet man das Liniendiagramm in ihr wieder, was man auch durch den obigen Prozeß des Hineinzoomens erhalten hatte. Die schwarzen Kreise geben die Begriffshierarchie der Dokumente wieder. Die übrigen Kreise zeigen an, für welche Kombinationen es keine baurechtlichen und bautechnischen Dokumente gibt. Grundsätzlich lassen sich mit TOSCANA bis zu vier Fragekomplexe gleichzeitig auf dem Bildschirm darstellen. In vielen Erkundungsvorgängen hat sich das zweigestufte Liniendiagramm als guter Kompromiß zwischen Komplexität und Übersichtlichkeit bewährt.

Wie schon am dargestellten kleinen Erkundungsbeispiel ersichtlich wird, stellen die vielfältigen Funktionen von TOSCANA zusammen mit Fragekomplexen ein mächtiges Werkzeug dar, mit dem zielsicher und schnell ein großer Dokumentenbestand erkundet werden kann. Hierbei ist es nicht notwendig, alle Frageterme der Datenbank schon im voraus zu kennen, sondern vielfach reicht schon eine ungefähre Kenntnis des Erkundungszieles aus, um dennoch zu inhaltlich präzisen Ergebnissen zu gelangen.

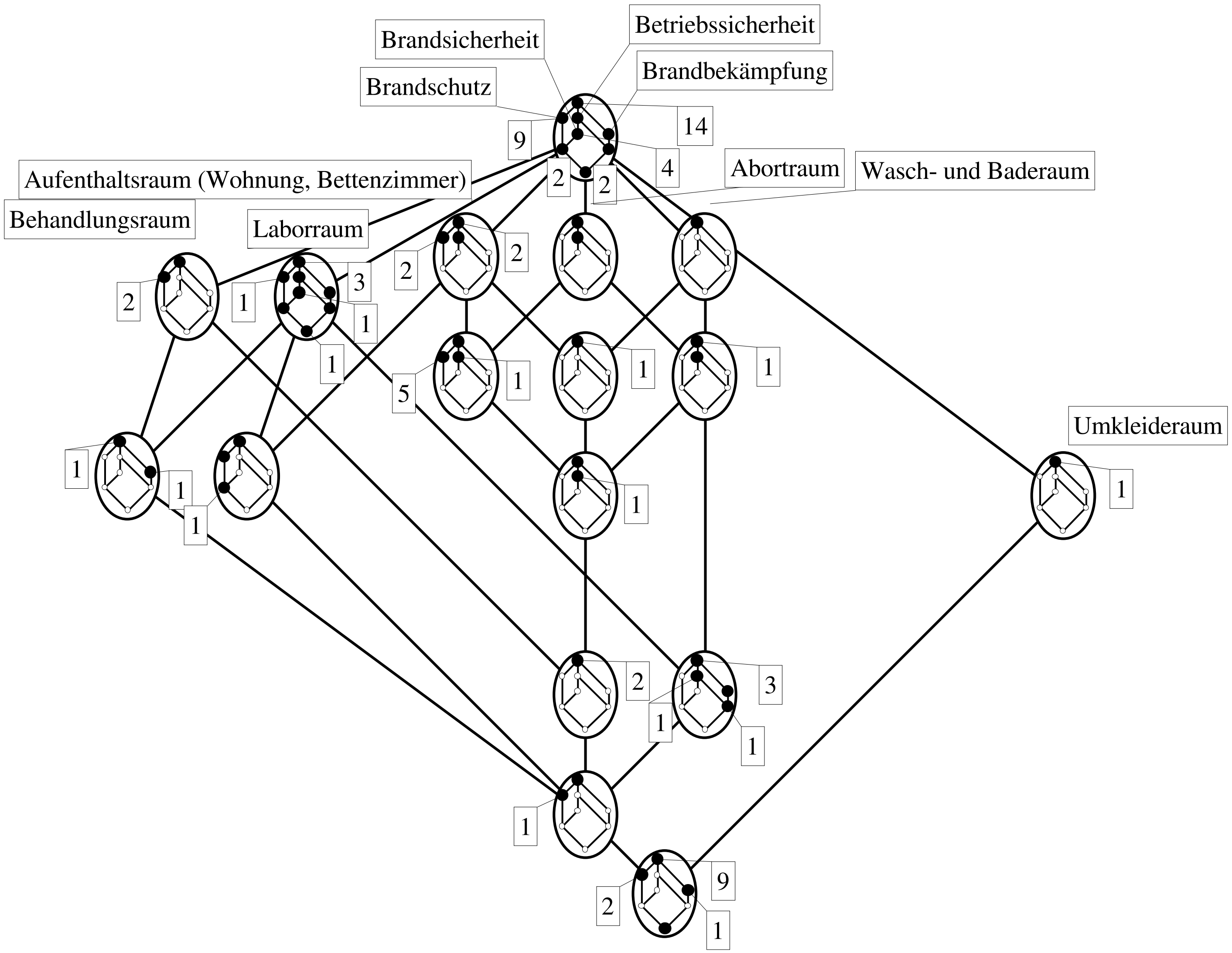

Abbildung 8: Gestuftes Liniendiagramm zu den Fragekomplexen 'Funktionsräume' und 'Betriebs- und Brandsicherheit'

4 Begriffliche Datensysteme: die Konzeption von TOSCANA

Wie in den beiden vorangehenden Abschnitten anhand von Anwendungsbeispielen erläutert worden ist, kann man mit TOSCANA Daten auf flexible Weise durch Liniendiagramme informativer Begriffsnetze veranschaulichen. TOSCANA fungiert dabei als ein allgemeines Verwaltungssystem, mit dem für konkrete Aufgaben Begriffliche Datensysteme erstellt werden können. Unter einem *Begrifflichen Datensystem* wird ein Programmsystem verstanden, das bezogen auf eine Datenbasis eine interaktive Analyse und Erkundung der gespeicherten Daten mit Hilfe sogenannter *Begriffsverbände* ermöglicht (s. [Scheich et al. 1993], [Wille 1992b]). Das im dritten Abschnitt beschriebene Erkundungssystem kann als ein Prototyp eines Begrifflichen Datensystems angesehen werden. Die in diesem Abschnitt zu erläuternde Konzeption von TOSCANA ist entsprechend ihrer Funktion untrennbar mit dem Konzept Begrifflicher Datensysteme verbunden, das deshalb zunächst ausführlicher diskutiert werden soll.

Die grundlegende Vorstellung von einem Begrifflichen Datensystem ist, daß es auf einer vorliegenden Datenbasis aufsetzt und ein begrifflich strukturiertes Abfragen von gespeichertem Wissen ermöglicht. Das Herzstück Begrifflicher Datensysteme besteht aus den *begrifflichen Abfragestrukturen*, mit denen Expertenwissen für die Datenanalyse und -erkundung aktiviert wird. Für das Beispiel des zweiten Abschnitts wurden die PC-Eigenschaften mit ihren Ausprägungen als begriffliche Abfragestrukturen aufbereitet, während im Beispiel des dritten Abschnitts die ausgearbeiteten Fragekomplexe das Wissen von Bauexperten für die Abfrage strukturieren. Nach der Terminologie der Formalen Begriffsanalyse fallen die begrifflichen Abfragestrukturen unter die *„begrifflichen Skalen"* (auch kurz *„Skalen"* genannt), mit denen die begrifflichen Beziehungen zwischen vorgegebenen *Gegenständen* und *Merkmalen* festgelegt werden. Begriffliche Skalen, deren begriffliche Struktur jeweils durch einen Begriffsverband gegeben ist, werden durch Liniendiagramme veranschaulicht (vgl. [Ganter, Wille 1989]). Um die Skalen möglichst flexibel verfügbar zu haben, arbeitet ein Begriffliches Datensystem mit drei Arten von Skalen: abstrakte, konkrete und realisierte Skalen. Eine *abstrakte Skala* stellt einen Begriffsverband mit zugehörigem Liniendiagramm bereit, wobei die aufgeführten Gegenstände und Merkmale abstrakt, d. h. ohne inhaltliche Bedeutung sind. Eine *konkrete Skala* entsteht aus einer abstrakten Skala, indem man die abstrakten Merkmale durch Merkmale mit inhaltlicher Bedeutung ersetzt und den abstrakten Gegenständen jeweils Abfrageterme zuordnet, mit denen *Objekte* aus der Datenbasis abrufbar sind (vgl. Abbildung 9); dabei soll ein abrufbarer Objekt genau diejenigen der eingebrachten inhaltlichen Merkmale haben, die ihm aufgrund der abstrakten Skala zukommen. Eine *realisierte* Skala entsteht aus einer konkreten Skala, indem man jeden abstrakten Gegen-

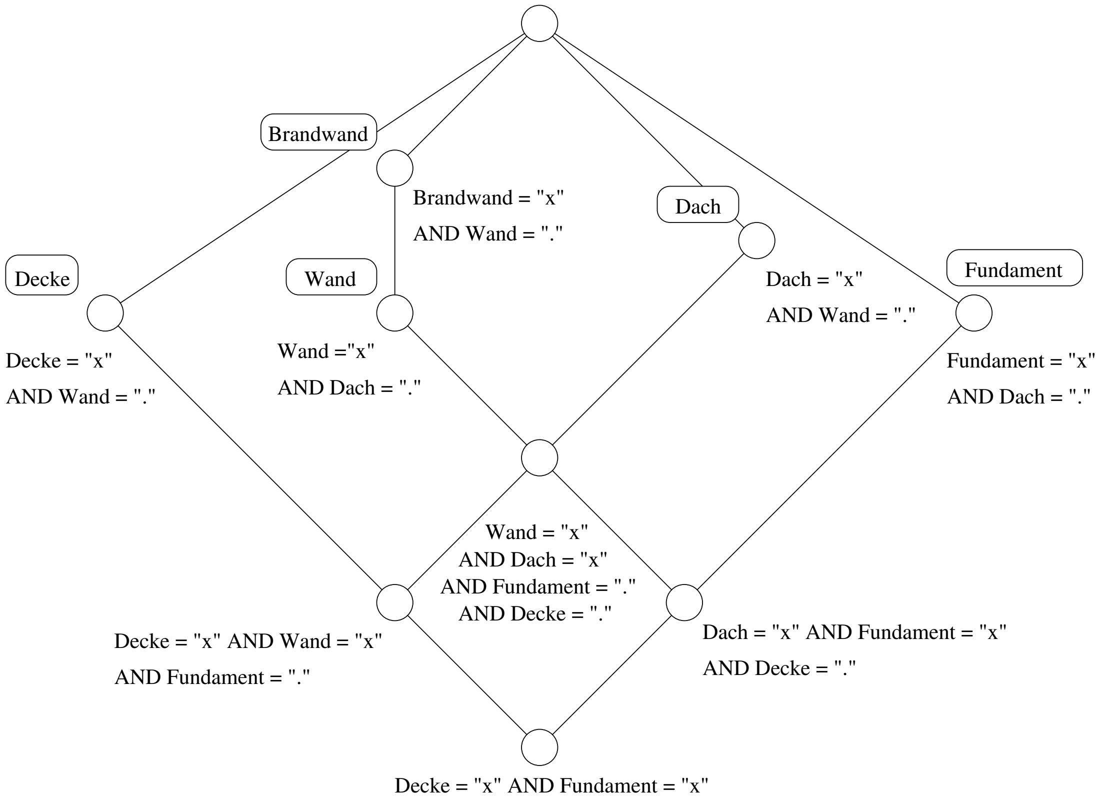

Abbildung 9: Beispiel einer konkreten Skala (z. B. bedeutet der Abfrageterm 'Brandwand = "x" AND Wand = ".": Suche alle Dokumente, die Brandwand betreffen und nicht Wand.)

stand durch alle die Objekte der Datenbasis ersetzt, die mit dem Abfrageterm des abstrakten Gegenstandes aus der Datenbasis abgerufen werden können.

Für abstrakte Skalen steht derzeit schon ein umfangreiches Repertoire an Begriffsverbänden und Liniendiagrammen zur Verfügung, das laufend erweitert wird. Auch für ein spontanes Einbringen neuer abstrakter Skalen in ein Begriffliches Datensystem liegen Verfahren vor. Mit den konkreten Skalen werden die relevanten Inhalte eines Anwendungsbereiches in Hinblick auf eine gegebene Aufgabestellung begrifflich strukturiert, womit das notwendige Theorie- und Erfahrungswissen für die Erkundung bzw. Analyse der vorliegenden Daten eingebracht wird. Den konkreten Skalen kommt deshalb bei der Erstellung Begrifflicher Datensysteme zentrale Bedeutung zu. Durch die Aufteilung in konkrete und realisierte Skalen werden Begriffliche Datensysteme relativ unabhängig von der Datenbasis; so bleiben die konkreten Skalen von Änderungen des Datenbestandes unberührt, sofern die Änderungen das kodierte Theoriewissen respektieren, und die realisierten Skalen können jeweils ohne Schwierigkeit aktualisiert werden. In einem

Begrifflichen Datensystem bilden die konkreten und realisierten Skalen zusammengenommen die *Begriffliche Datei*, die die entscheidenden Informationen für die interaktive Systemnutzung liefert.

Wie wird nun ein Begriffliches Datensystem interaktiv genutzt? Zunächst wählt der Benutzer entsprechend seines Interesses aus dem angebotenen *Menü* von realisierten Skalen eine Folge von Skalen aus, wobei er zusätzlich festlegt, wieviele der Skalen jeweils gleichzeitig gezeigt werden sollen. Das System präsentiert dann das n-gestufte Liniendiagramm der ersten n Skalen der ausgewählten Folge, wobei n die festgelegte Skalenanzahl ist. Im gestuften Liniendiagramm wird die erste Skala der Folge durch das gröbste Liniendiagramm dargestellt, die zweite durch das nächst feinere usw.; dabei wird der inhaltliche Bezug durch entsprechende Beschriftung auf der jeweils zugehörigen Stufe sichtbar. Will der Benutzer an gewissen Stellen des gestuften Liniendiagramms durch weitere Skalen der Ausgangsfolge feinere Einblicke nehmen, kann er solche Stellen anwählen und erhält dann wieder ein n-stufiges Liniendiagramm, bei dem die $n+1$-ste Skala der Folge die feinste Stufe bildet und dafür die erste Skala wegfällt. Auf diese Weise kann der Benutzer auf immer feinere Stufen des Informationsbestandes hineinzoomen, bis er die für ihn relevanten Antworten findet. Natürlich darf er vor jedem Schritt die gewählte Folge auch abändern, was aufgrund der schon gewonnenen Informationen als durchaus wünschenswert erscheinen mag. Manchmal ist es auch hilfreich, die in einem gestuften Liniendiagramm dargestellte Skalenfolge zu permutieren, da an dem neu entstehenden Diagramm manche Sachverhalte und Zusammenhänge leichter abgelesen werden können. Da die darzustellenden Objekte der Datenbasis in vielen Fällen sehr zahlreich sind, erlaubt ein Begriffliches Datensystem die Anzahlen solcher Objekte anzugeben, um dann erst auf feineren Stufen mit hinreichend kleinen Anzahlen die einzelnen Benennungen der zugehörigen Objekte einzublenden. Eine Reihe weiterer Funktionen von TOSCANA unterstützen die Systemnutzung, auf die hier jedoch nicht weiter eingegangen werden soll.

Wie schon erwähnt, ist TOSCANA ein allgemeines Werkzeug, mit dem das Konzept Begrifflicher Datensysteme im konkreten Fall realisiert werden kann. TOSCANA ist objektorientiert in C++ programmiert und benutzt die *Structured Query Language* (SQL) zur Abfrage von Datenbanken. Eine verfügbare Version von TOSCANA läuft unter MS-Windows™ (s. Oberfläche in Abbildung 10) und arbeitet mit dem relationalen Datenbankmanagementsystem MS-Access™ zusammen. Sind die zu bearbeitenden Daten in einer relationalen Datenbank vorhanden, braucht man nur noch die konkreten Skalen mit der dafür entwickelten Skriptsprache CONSCRIPT einzugeben, um ein funktionsfähiges Begriffliches Datensystem zu erhalten. Die realisierten Skalen werden aktuell für das Anzeigen der Liniendiagramme aus der Datenbank abgefragt.

Das Menü (s. Abb. 10) von TOSCANA bietet unter den vier Menütiteln

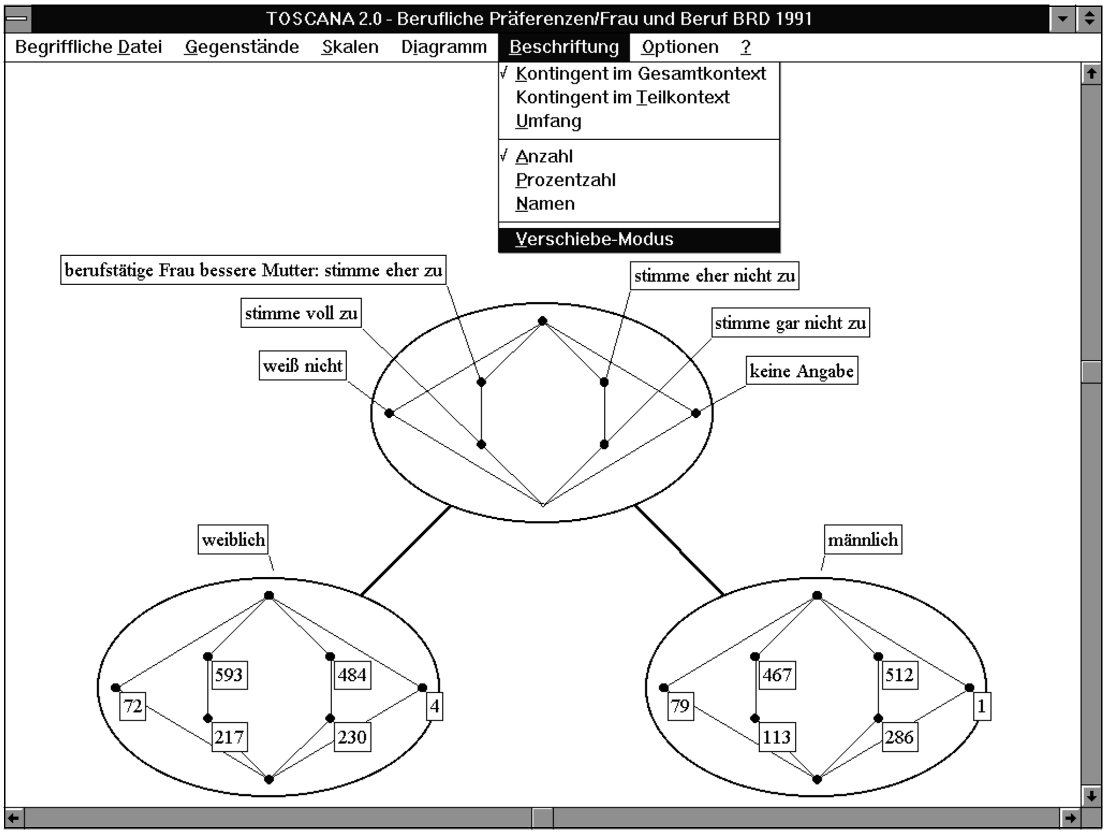

Abbildung 10: Oberfläche von TOSCANA mit dem Menü zur Auswahl der Beschriftung. Dem Begriffsverband liegt die Frage zugrunde: „Ist eine berufstätige Frau eine bessere Mutter?“ (vgl. [Allbus 1991])

‘Begriffliche Datei’, ‘Gegenstände’, ‘Skalen’, ‘Diagramm’ die Funktionsgruppen zum Arbeiten in Begrifflichen Datensystemen. Mit den Funktionen zur ‘Begrifflichen Datei’ werden Begriffliche Dateien ausgewählt und damit verschiedene Begriffliche Datensysteme aktiviert. Der Menütitel ‘Gegenstände’ ermöglicht das Auswählen einer Gruppe von Objekten aus der Gesamtmenge der Objekte in der Datenbank. Nach dem Aktivieren des Begrifflichen Datensystems wählt der Benutzer unter dem Menütitel ‘Skalen’ die realisierten Skalen aus, die im Liniendiagramm dargestellt werden sollen. Außerdem beinhaltet der Menütitel das Protokoll der bisher durchlaufenen Skalen. Das Anzeigen des Liniendiagramms nach den ausgewählten Skalen und den eingestellten Parametern (unter den Menütiteln ‘Beschriftung’ und ‘Optionen’) wird unter dem Menütitel ‘Diagramm’ aktiviert. Außerdem finden sich unter diesem Menütitel Funktionen zum Zurückspringen innerhalb der dargestellten Skalen und zum Vertauschen von dargestellten Skalen sowie zum Drucken des Diagramms.

5 Begriffliche Wissensverarbeitung: der pragmatische Ort von TOSCANA

TOSCANA ist im Rahmen der *Forschungsgruppe Begriffsanalyse* an der Technischen Hochschule Darmstadt entwickelt worden. Voraussetzung dafür waren mehr als zehn Jahre theoretischer und praktischer Erfahrungen mit der um 1980 in Darmstadt entstandenen *Formalen Begriffsanalyse* (s.[Wille 1982], [Wille 1992a]). Anstoß für die Entwicklung des Werkzeugs TOSCANA gaben Schwierigkeiten bei der Behandlung größerer Datensätze, die mit den verfügbaren Programmen zur Formalen Begriffsanalyse nur mit großem Aufwand befriedigend analysiert werden konnten. Diese Schwierigkeiten rührten vor allem daher, daß nach dem Grundverständnis der Formalen Begriffsanalyse bei der Datenbehandlung und -auswertung die Verbindung zu den Originaldaten stets transparent zu halten ist, um eine stimmige Interpretation der Daten zu ermöglichen. Den entscheidenden Schritt zur Entwicklung des neuen Werkzeugs brachte anhand konkreter Projekte die Einsicht, daß in Hinblick auf inhaltliche Fragen zum Untersuchungsgegenstand in der Regel nur begrenzte Teile der Gesamtdaten zugänglich zu machen sind. Mit der *Begrifflichen Skalierung* lag schon eine Methode vor, wie solche Teildaten im Sinne des inhaltlichen Verständnisses vom Untersuchungsgegenstand begriffsanalytisch aufbereitet werden können (s. [Ganter, Wille 1989]). Auch für die begriffsanalytische Verbindung und graphische Darstellung der aufbereiteten Teildaten bot sich mit den *gestuften Liniendiagrammen* eine schon vielfach bewährte Methode an (s. [Wille 1984], [Wille 1989]). So entstand die Konzeption der *Begrifflichen Datei* (s. [Vogt et al. 1991]), die für das Werkzeug TOSCANA grundlegend ist: In bezug auf die vorliegenden Daten repräsentiert sie inhaltliches Wissen in begrifflichen Skalen und deren Begriffsverbänden, für die Liniendiagramme zur graphischen Darstellung bereitgehalten werden. Ist schon die Konzeption der Begrifflichen Datei anhand von Anwendungsprojekten entwickelt worden, so ist ihr weiterer Ausbau zum Konzept des *Begrifflichen Datensystems* ebenfalls in Zusammenhang mit konkreten Projekten erfolgt (s. [Wachter, Wille 1992], [Scheich et al. 1993]). Für diese Projekte war es natürlich zwangsläufig, ein geeignetes Programm für die Rechnerbehandlung der erarbeiteten Datensysteme verfügbar zu machen. Als dieses Programm ist TOSCANA entstanden.

Von ihren Anfängen an ist für die Formale Begriffsanalyse die Auffassung grundlegend gewesen, daß Sinn und Bedeutung ihrer Methoden und Verfahren im angemessenen Bezug zur Wirklichkeit zu sehen sind; dabei gilt es Zwecke und Ergebnisse im rationalen Diskurs kritisch zu bestimmen. Hinter dieser Auffassung steht die von Ch. S. Peirce begründete *pragmatische Philosophie* (s. [Peirce 1991]), die ihre aktuelle Fortsetzung in der Diskursphilosophie von K.-O. Apel und J. Habermas findet (s. [Apel 1976], [Habermas 1981]). Nach der pragmatischen Philosophie ist Grundbedingung für den Umgang mit Daten, daß stets von einem lebensweltlichen Vorverständnis ausgegangen wird, von dem auch die formale Behandlung kodierten Wissens nicht abschneiden darf. Deshalb wird in der Formalen Begriffsanalyse so großen Wert daraufgelegt, daß die Verbindung zu den Ausgangsdaten erhalten bleibt und der inhaltliche Kontext bei der Dateninterpretation rekonstruiert werden kann. Gewißheits- und Gültigkeitsansprüche von Interpretationen können nach K.-O. Apel nur durch rationale Argumentation im Rahmen der intersubjektiven Kommunikationsgemeinschaft gesichert werden. Das hierbei postulierte *Apriori der Kommunikationsgemeinschaft als der sinnkritischen Bedingung der Möglichkeit und Gültigkeit aller Argumentation* betrifft die Argumentierenden auf doppelte Weise: *Wer nämlich argumentiert, der setzt immer schon zwei Dinge voraus: Erstens eine 'reale Kommunikationsgemeinschaft', deren Mitglied er selbst durch einen Sozialisationsprozeß geworden ist, und zweitens eine 'ideale Kommunikationsgemeinschaft', die prinzipiell imstande sein würde, den Sinn seiner Argumente adäquat zu verstehen und ihre Wahrheit definitiv zu beurteilen.* [Apel 1976, Bd. 2, S. 429] Da Interpretationen letztendlich inhaltlichen Charakter haben, bringt die beschriebene Verankerung in der menschlichen Kommunikations- und Argumentationsgemeinschaft wiederum mit sich, daß der inhaltliche Kontext und damit auch der Wirklichkeitsbezug bei der formalen Datenbehandlung nicht verloren gehen darf.

Bei umfangreichen Daten- und Wissenssystemen, die auf maschinellen Rechnern implementiert sind und somit ausschließlich formal arbeiten, ist die Gefahr der inhaltlichen Verkürzung und der Beschränkung von argumentativer Kommunikation besonders groß. Deshalb werden Werkzeuge begrifflicher Wissensverarbeitung benötigt, die die rationale Kommunikation und Interaktion auch bei großen Datenmengen ermöglichen. TOSCANA ist als ein solches Werkzeug entworfen und in ersten Anwendungen erprobt. Die zu diskutierende Frage ist, wieweit das Werkzeug TOSCANA dem in [Wille 1994] umrissenen Verständnis *Begrifflicher Wissensverarbeitung* entspricht. Dieses Verständnis sieht den denkenden und argumentierenden Menschen als konstitutiv für die Begriffliche Wissensverarbeitung, was auch für die Formale Begriffsanalyse zutrifft, zu deren Werkzeugen TOSCANA gehört. Grundforderung an die Begriffliche Wissensverarbeitung ist, daß bei ihr reichhaltige Verbindungen zwischen den thematisierten Inhalten und den formalen Hilfsmitteln bestehen, wobei die Schnittstellen zwischen Inhaltlichem und Formalem transparent modelliert sein sollen. Daß das bei TOSCANA in angemessener Weise der Fall ist, soll kurz skizziert werden: Die mit TOSCANA erfaßten *konkreten Skalen*, in denen inhaltliche Vorstellungen und Zusammenhänge des zu behandelnden Wissensbereich repräsentiert werden, sind die wichtigsten Schnittstellen in Hinblick auf die Eingabe von inhaltlichem Wissen in das formale System. Die elementare Ebene der Datenrepräsentation, auf die TOSCANA bezogen ist, stellt natürlich ebenfalls eine derartige Schnittstelle dar. Die *gestuften Liniendiagramme* zeigen als Schnittstellen der Ausgabe unverkürzt die mit den Skalen eingegebenen Inhalte in Verbindung mit den zugehörigen Daten. Durch die Möglichkeit der beliebigen Kombination von Skalen und der Navigation durch unterschiedliche Stufungen von Liniendiagrammen können die Schnittstellen flexibel aktiviert und damit auf vielfältige Weise inhaltliche Zusammenhänge aufgedeckt werden. Die einzige Beschränkung liegt darin, daß wegen ihrer großen Zahl gar nicht alle Kombinationen durchgespielt werden können. Das macht klar, daß eine erfolgreiche Nutzung natürlich einen kompetenten Benutzer voraussetzt, der in Hinblick auf das Inhaltliche lernend zu fragen versteht.

Begriffliche Wissensverarbeitung bezieht sich auf *(anspruchsvolles) Wissen*, das seine Bestimmung erst durch die intersubjektive Kommunikations- und Argumentationsgemeinschaft erhält (vgl. [Wille 1992b]). Daß sich auch auf dieser intersubjektiven Ebene mit TOSCANA menschliches Denken und Urteilen sowie rationale Kommunikation (im Sinne von [Habermas 1981]) aktivieren läßt, darauf weisen die reichhaltigen Erfahrungen mit Liniendiagrammen von Begriffsverbänden als wirksames Instrument der Wissenskommunikation hin. Schon bei der gemeinsamen Spezifikation und Erhebung von Daten und Wissen sowie deren Kontrolle haben sich die begrifflichen Liniendiagramme vielfach bewährt, so daß man beispielsweise damit begonnen hat, Begriffsverbände im Software Engineering einzusetzen (vgl. [Krone, Snelting 1993]). Mit TOSCANA steht nun ein variables und flexibles Werkzeug zur Verfügung, mit dem der zwischenmenschliche Verständigungsprozeß in Hinblick auf gemeinsame Interpretationen komplexer Zusammenhänge gefördert werden kann. Da solche Interpretationen stets zweckorientiert sind, sollten sie auf intersubjektiver Ebene abgestimmt und argumentativ ausgehandelt werden, wozu bei zunehmender Komplexität Werkzeuge wie TOSCANA immer hilfreicher werden. Hierbei kommt dem hinter TOSCANA stehenden Paradigma der Wissenslandschaft besondere Bedeutung zu, da es dem Offenen, Prozeßhaften und Diskursivem der Wissensfindung besser Rechnung trägt als das Paradigma vom Fakten- und Regelwissen. Es ist durchaus statthaft, bei TOSCANA (*Tools of Concept Analysis*) an die toskanische Landschaft zu denken, von der es in [Zimmermann 1980] heißt: *„Die von Natur stark strukturierte Landschaft wurde vom Menschen im Laufe der Jahrhunderte zu einem Kulturland gestaltet. (...) in diesem Land kamen sich Schönheitsempfinden und Verstand, Kunst und Wissenschaften näher als irgendwo anders“.*

Literatur

[Allbus 1991] Zentralarchiv/Zuma: Allbus Basisumfrage 1991 in Gesamtdeutschland. Codebook ZA-Nr. 1990. Zentralarchiv für empirische Sozialforschung, Köln 1992.

[Apel 1976] K.-O. Apel: *Transformation der Philosophie. Band 1: Sprachanalytik, Semiotik, Hermeneutik. Band 2: Das Apriori der Kommunikationsgemeinschaft.*

Suhrkamp–Taschenbuch Wissenschaft Band 164/165, Frankfurt 1976.

[Ganter, Wille 1989] B. Ganter, R. Wille: Conceptual scaling. In: F. Roberts (ed.): *Applications of combinatorics and graph theory to the biological and social sciences.* Springer-Verlag, New York 1989, S. 139–167.

[Gödert 1988] W. Gödert: Online-Katalog und bibliothekarische Inhaltserschließung. *Zeitschrift für Bibliothekswesen und Bibliographie*, Sonderheft 46 (1988), S. 280–302.

[Habermas 1981] J. Habermas: *Theorie kommunikativen Handelns. 2 Bände.* Suhrkamp, Frankfurt 1981.

[Henning 1994] H. J. Henning: Zur kontextualistischen Sichtweise und methodologischen Entwicklung in der psychologischen Datenanalyse. In: R. Wille und M. Zickwolff (Hrsg): *Grundfragen und Aufgaben der Begrifflichen Wissensverarbeitung*, B. I.–Wissenschaftsverlag, Mannheim, 1994, S. 301–321.

[Kollewe et al. 1994] W. Kollewe, M. Skorsky, R. Wille: Abschlußbericht des Projektes „Begriffliche Wissensverarbeitung auf dem Gebiet des Baurechts und der Bautechnik“, Phase I (in Zusammenarbeit mit dem Ministerium für Bauen und Wohnen des Landes Nordrhein-Westfalen). Darmstadt und Düsseldorf 1994.

[Krone, Snelting 1993] M. Krone, G. Snelting: On the inference of configuration structures from source code. *Informatik-Berichte 93-06*, Technische Universität Braunschweig, Braunschweig 1993.

[Lamnek 1988] S. Lamnek: *Qualitative Sozialforschung. Band 1: Methodologie.* Psychologie Verlags Union, München, Weinheim 1988.

[Orth 1974] B. Orth: *Einführung in die Theorie des Messens.* Kohlhammer, Stuttgart 1974.

[PC Magazine 1993] PC Magazine, Band 12, Nummer 2. Ziff-Davis Publishing Company, Boulder 1993, S. 110–247.

[Peirce 1991] C. S. Peirce: *Schriften zum Pragmatismus und Pragmatizismus.* Suhrkamp–Taschenbuch Wissenschaft Band 945, herausgegeben von K.–O. Apel. Suhrkamp, Frankfurt 1991.

[Scheich et al. 1993] P. Scheich, M. Skorsky, F. Vogt, C. Wachter, R. Wille: Conceptual data systems. In: O. Opitz, B. Lausen, R. Klar (eds.): *Information and classification.* Springer-Verlag, Heidelberg 1993, S. 72–84.

[Spangenberg 1990] N. Spangenberg: Familienkonflikte eßgestörter Patientinnen. Eine empirische Untersuchung mit der Repertory Grid Technik. Habilitationsschrift, Universität Gießen 1990; erscheint 1994 im Westdeutschen Verlag.

[Spangenberg, Wolff 1991] N. Spangenberg, K.-E. Wolff: Comparison between biplot analysis and formal concept analysis of repertory grids. In: H. H. Bock, P. Ihm (eds.): *Classification, data analysis, and knowledge organization.* Springer-Verlag, Berlin, Heidelberg 1991, S. 104–112.

[Vester 1988] F. Vester: *Leitmotiv Vernetztes Denken. Für einen besseren Umgang mit der Welt.* Heyne Verlag, München 1988.

[Vogt et al. 1991] F. Vogt, C. Wachter, R. Wille: Data analysis based on a conceptual file. In: H.-H. Bock, P. Ihm (eds.): *Classification, data analysis, and knowledge organization*, Springer-Verlag, Berlin, Heidelberg 1991, S. 131–140.

[Wachter, Wille 1992] C. Wachter, R. Wille: Formale Begriffsanalyse von Literaturdaten. In: DGD (Hrsg.): *Deutscher Dokumentartag 1991 – Information und Dokumentation in den 90er Jahren: Neue Herausforderung, neue Technologien*, Frankfurt 1992, S. 203–224.

[Wille 1982] R. Wille: Restructuring lattice theory: an approach based on hierarchies of concepts. In: I. Rival (ed.): *Ordered sets*, Reidel, Dordrecht–Boston 1982, S. 445–470.

[Wille 1984] R. Wille: Liniendiagramme hierarchischer Begriffssysteme. In: H.-H. Bock (Hrsg.): *Anwendungen der Klassifikation: Datenanalyse und numerische Klassifikation*, Indeks-Verlag, Frankfurt 1984, S. 32–51; engl. Übersetzung: Line diagrams of hierarchical concept systems. *Int. Classif.* **11** (1984), S. 77–86.

[Wille 1989] R. Wille: Lattices in data analysis: how to draw them with a computer. In: I. Rival (ed.): *Algorithms and order.* Kluwer, Dordrecht–Boston 1989, S. 33–58.

[Wille 1992a] R. Wille: Concept lattices and conceptual knowledge systems. *Computers & Mathematics with Applications*, 23 (1992), S. 493–515.

[Wille 1992b] R. Wille: Begriffliche Datensysteme als Werkzeug der Wissenskommunikation. In: H. H. Zimmermann, H.-D. Luckhardt, A. Schulz (Hrsg.): *Mensch und Maschine — Informationelle Schnittsstellen der Kommunikation*, Universitätsverlag Konstanz, Konstanz, 1992, S. 63–73.

[Wille 1994] R. Wille: Plädoyer für eine philosophische Grundlegung der Begrifflichen Wissensverarbeitung. In: R. Wille und M. Zickwolff (Hrsg.): *Grundfragen und Aufgaben der Begrifflichen Wissensverarbeitung*, B. I.–Wissenschaftsverlag, Mannheim 1994, S. 11–26.

[Zimmermann 1980] K. Zimmermann: *Toscana.* DuMont Buchverlag, Köln, 3. Auflage, 1980.