

Informatik-Fachberichte 128

Herausgegeben von W. Brauer
im Auftrag der Gesellschaft für Informatik (GI)

Wolfgang Benn

**Dynamische
nicht-normalisierte Relationen
und symbolische Bildbeschreibung**



Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo

Autor

Dr. Wolfgang Benn
SCS Organisationsberatung und Informationstechnik GmbH
Oehleckerring 40, 2000 Hamburg 62

CR Subject Classifications (1985): H.2, H.2.1, H.2.7, I.4.8, I.4.0,
H.3.2, I.2.1

ISBN-13: 978-3-540-16823-2 e-ISBN-13: 978-3-642-71390-3
DOI: 10.1007/978-3-642-71390-3

CIP-Kurztitelaufnahme der Deutschen Bibliothek. Benn, Wolfgang: Dynamische nicht-normalisierte Relationen und symbolische Bildbeschreibung / Wolfgang Benn. - Berlin; New York; Tokyo: Springer, 1986.
(Informatik-Fachberichte; 128)
ISBN 3-540-16823-0 (Berlin ...)
ISBN 0-387-16823-0 (New York ...)
NE: GT

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically those of translation, reprinting, re-use of illustrations, broadcasting, reproduction by photocopying machine or similar means, and storage in data banks. Further, storage or utilization of the described programmes on data processing installations is forbidden without the written permission of the author. Under §54 of the German Copyright Law where copies are made for other than private use, a fee is payable to "Verwertungsgesellschaft Wort", Munich.

© Springer-Verlag Berlin Heidelberg 1986

Druck und Bindearbeiten: Weihert-Druck GmbH, Darmstadt
2145/3140-543210

Geleitwort

Die wissenschaftliche Entwicklung erhält ihre Impulse von Anforderungen, die aus neuen Fragestellungen erwachsen. Dabei können Anforderungen aus einem Gebiet Fragen in einem anderen Gebiet aufwerfen, die dort zu fortschrittlichen Lösungen führen. Es war aufregend zu beobachten, wie das Problem, bildbeschreibende Strukturen systematisch zu verwalten, seine Lösungsansätze mehr und mehr aus der Theorie der Datenbanken bezog.

Der Autor dieses Buches, Wolfgang Benn, nahm die Mühe auf sich, in zwei Gebieten der Informatik, Bildverarbeitung und Datenbanksysteme, zu Hause sein zu müssen. Es ist ihm gelungen, in beiden Bereichen den Stand der Kenntnisse zu mehren. Daß dabei ganz fremd erscheinende Ideen zusammengeführt wurden, hat uns beide außerordentlich gefreut.

Die Verwendung relationaler Strukturen zur Beschreibung von Bildern legte den Gedanken nahe, zu ihrer Verwaltung relationale Datenbanken einzusetzen. Dem Anspruch, komplexe, erst zur Zeit der Auswertung der zu beschreibenden Bilder bekannte Strukturen zu manipulieren, wuchs die Entwicklung der Theorie des „Non-First-Normal-Form“-Relationenmodells entgegen. Dies war ein Ansporn, der Wolfgang Benn zu einer Weiterentwicklung in Richtung dynamischer Modelle führte. Gleichzeitig bettete er damit den Umgang mit Relationengebilden zur Beschreibung in den Formalismus der relationalen Algebra ein.

Als ich (B.R.) noch in Hamburg zusammen mit dem Mitinitiator des Projektes, Hans-Hellmut Nagel, den ersten Antrag zur Unterstützung an die Deutsche Forschungsgemeinschaft formulierte, war mir nur die Notwendigkeit der Arbeit an dem Problem der Verwaltung von bildbeschreibenden Strukturen deutlich klar. Daß eine ganz allgemeine Lösung dieses Problems gefunden wurde, hat uns dazu geführt, Wolfgang Benn zu diesem Buch zu ermutigen.

Seiteneffekte entstehen in der Informatik nicht nur bei der Ausführung von Anweisungen in Programmen. In diesem Falle sehen wir den (nicht unwichtigen) Seiteneffekt des wissenschaftlichen Fortschritts in der Zusammenarbeit zwischen Disziplinen und nicht zuletzt Personen. Auch in diesem Sinne war das Projekt ein Erfolg.

Im Juli 1986

B. Radig, München

H.-J. Schek, Darmstadt

Vorwort

Die maschinell erfassbare Darstellung und Verwaltung von Daten und deren Bedeutungsgehalt ist die Domäne der semantischen Datenmodellierung. Arbeitsbereiche, die traditionell an intellektuelle Fähigkeiten eines Menschen gebunden waren, werden durch die geordnete Darstellung von Informationen und deren rasche Zugreifbarkeit in speziell hierfür erstellten Programmsystemen unterstützt und rationalisiert. Deutlich sichtbar wird diese Tendenz an dem wachsenden Interesse an Expertensystemen für den Einsatz im technisch-wissenschaftlichen und -industriellen Bereich. Faktenwissen und Sachverstand von Experten eines Aufgabengebietes sind in solchen Systemen als Wissensbasen und Regelwerke gespeichert. Automatisch erzeugte Entscheidungen auf dieser Grundlage dienen zur Verifikation des vorhandenen Wissens und zur Erprobung neuer Verfahren der Wissenserweiterung.

Wissensbasen enthalten im allgemeinen eine Vielzahl von Fakten, die nach den vorgegebenen Verwendungskriterien zu modellieren und zu ordnen sind, und die während der Systemanwendung rasch zugreifbar sein sollen. Traditionell bieten Datenbanken mit den ihnen unterliegenden Datenmodellen derartige Darstellungs-, Ordnungs- und Zugriffstechniken – hier liegt auch der Ursprung semantischer Datenmodellierung.

Das relationale Datenmodell nimmt eine besondere Stellung ein, weil es eine einfache Benutzeroberfläche anbietet und der Datenmodellierung nur wenige Einschränkungen auferlegt. Weitgehend parallel wurden daher in verschiedenen Arbeitsbereichen Methoden untersucht, relationale Datenbanken als unterstützendes Werkzeug einzusetzen. Bereits Ende der siebziger Jahre erkannte man jedoch, daß sich die Flexibilität des traditionellen relationalen Modells noch verbessern läßt. So entstanden im wesentlichen zwei Untersuchungsrichtungen mit ähnlichen Zielsetzungen:

1. Die Entwicklung eigener semantischer Datenmodelle, die weitgehend auf traditionellen relationalen Datenbanken aufbauen und besonderen Wert auf mächtige und komfortable Anwendungsoberflächen legen.
2. Erweiterungen des relationalen Modells um geeignete Datenstrukturen zur Modellierung von Applikationsdaten auf der Grundlage eines erweiterten mathematischen Modells.

Im vorliegenden Buch wird ein neuartiges erweitertes Datenmodell vorgestellt und seine Eignung zur Modellierung sogenannter *non-standard*-Applikationen gezeigt. Als Beispiel hierzu dient der Problembereich Bild- und Bildfolgendeutung mit seinen Forderungen nach der Erzeugung einer objektorientierten Datensicht und der Anpassung von Datenstrukturbeschreibungen an wechselnde Objektbeschreibungen in Realweltszenen. Durch die Integration strukturvarianter Attribute in das nicht-normalisierte relationale Datenmodell lassen sich problemorientierte Schema-Adaptionen darstellen, wie sie beispielsweise auch bei der Modellierung von Datenversionen in traditionellen relationalen Datenbanken auftreten, wenn sich eine Datenstruktur über die Zeit verändert hat, aber noch den gleichen Sachverhalt repräsentiert.

Die Anwendbarkeit dieses Modells für verschiedene Problemstellungen und die Möglichkeiten einer geeigneten Realisierung werden in den nächsten Jahren meine wesentlichen Arbeitsziele darstellen.

Wolfgang Benn

Inhaltsverzeichnis

Einleitung	1
1 Formen der Wissensrepräsentation	5
1.1 Repräsentationsformen	7
1.1.1 Semantische Netze	8
1.1.2 Generische Schemata – frames	13
1.1.3 Relationengebilde	15
1.2 Datenmodelle	17
1.2.1 Das hierarchische Modell	18
1.2.2 Das Netzwerkmodell	19
1.2.3 Das relationale Modell	20
1.2.4 Semantische Datenmodellierung	22
1.3 Zusammenfassung	24
2 Datenmodellanforderungen	26
2.1 Überblick und Historie	27
2.1.1 Der Ansatz von Lorie et al.	30
2.1.2 Das NF ² -Modell nach Schek et al.	31
2.1.3 Beurteilung der Ansätze	33
2.2 Relationengebilde im relationalen Datenmodell	35
2.2.1 Gruppierung primitiver Bildsymbole	35
2.2.2 Beziehungen zwischen Bildsymbolen	40
2.2.3 Beschreibungshierarchie und Objektupel	41
2.2.4 Multirelationale Teilgebilde	44
2.2.5 Zusammenfassung und Entscheidung	45
2.3 Objektvergleiche als Datenbankanfragen	46
2.3.1 Allgemeine Anfragemethoden	46
2.3.2 Anfragen nach Objekteigenschaften	47
2.3.3 Anfragen nach Eigenschaften der Beschreibungsstruktur	48
2.3.4 Tolerante Vergleiche als Datenbankanfragen	54
2.4 Zusammenfassung	59

3	Dynamische nicht-normalisierte Relationen – (NF^{2D}-Relationen)	60
3.1	Ausgewählte Modellerweiterungen	61
3.1.1	Codd's erweitertes RM/T-Modell	62
3.1.2	Das semantische Datenmodell GEM	64
3.1.3	Beurteilung der Ansätze	65
3.2	Definitionen	66
3.2.1	Wertebereich und Attribut	67
3.2.2	Zuordnungsfunktion δ	68
3.2.3	Attributbegriffe	69
3.2.4	Tupel und Relation	69
3.2.5	Schemadarstellung	70
3.3	Operationen auf NF ^{2D} -Relationen	71
3.3.1	Vereinigung $+$ und Differenz $-$	72
3.3.2	Selektion σ	73
3.3.3	Projektion π	75
3.3.4	Relationenverbund, Join \bowtie	77
3.3.5	Varianzexpansion ι	79
3.3.6	Nestung ν und Entnestung μ	81
3.3.7	Zusammenfassung	84
3.4	Nullwerte in strukturtoleranten Datenbankanfragen	84
3.4.1	Typbindung und verdeckte Felder	85
3.4.2	Wahrheitswerte als Diskriminanten	85
3.4.3	Tolerante Datenbankanfragen	86
3.5	Objektvarianten und Symbolklassen	87
3.5.1	Traditionelle Darstellungen	87
3.5.2	Eine Lösung im RM/T-Modell	88
3.5.3	Alternative Darstellung mit NF ^{2D} -Relationen	88
3.6	Bild- und Schematupel	90
3.7	Anfragemethoden für NF ^{2D} -Relationen	95
3.7.1	Query-by-Structure-Example (QSE)	99
3.8	Zusammenfassung	102

4	Forderungen an eine NF^{2D}-Datenbank für symbolische Bildbeschreibungen	103
4.1	Datenbankintegration in einem Programmiersystem	103
4.1.1	Dokumentation und Informationsübersicht	104
4.1.2	Elemente einer Ada-Programmierungsumgebung	105
4.2	Aktive Unterstützung der Programmphase	113
4.2.1	Laufzeitprüfungen und Datenintegrität	113
4.3	Stand einer Prototypimplementierung	118
4.3.1	Die DRAGON-Datenbank	119
4.3.2	Spezifikation einer Anfrage	122
4.3.3	Zusammenfassung	125
	Epilog	126
	Literaturverzeichnis	130
	A Ada-Programmbeispiele	139

Verzeichnis der Abbildungen

1.1	Bildfolge, schematisiert	5
1.2	Bildobjekt, Seitenansicht	9
1.3	Objektbeschreibung: Symbolebene	10
1.4	Objektbeschreibung: Generalisierungsebene	10
1.5	Einzelbildbeschreibung	10
1.6	Regelerzeugung	12
1.7	Zuordnung regelhafter Interpretationsergebnisse	12
1.8	Generisches Schema der Beispielbildfolge	14
2.1	Attributintegration durch <i>Nestung</i>	31
2.2	Schemaveränderung durch <i>Nestung</i>	32
2.3	Attributseparation durch <i>Entnestung</i>	33
2.4	Symbol im Entity-Relationship-Modell	36
2.5	Beziehungsdiagramm einer Bildsymbolklasse	36
2.6	Schemata von Beispielrelationen	37
2.7	Werkstück/Repräsentation: Strukturanaloge Darstellung	38
2.8	Werkstück/Repräsentation: Abstrahierte Darstellung	38
2.9	Flaches Relationenschema mit einelementigen Mengen	39
2.10	Schema mit ein- und mehrelementigen Mengenattributen	39
2.11	Schema ohne einelementige Mengenattribute	40
2.12	Redundanz in Symbolen 1	40
2.13	Redundanz in Symbolen 2	41
2.14	Redundanz in Symbolen 3	41
2.15	Problemtransformation: Bildobjekt - Objekttuplel	43
2.16	Homomorphismus	50
2.17	Modellierung und Schemata der Beispielrelationen	50
2.18	Relationenschemata als NF^2 -Relationen	51
2.19	Monomorphismus	52
2.20	Isomorphismus	52
2.21	Komorphismus	53

2.22	Attributtoleranz 1	55
2.23	Attributtoleranz 2	55
2.24	Attributtoleranz 3	56
2.25	Strukturtoleranz 1	57
2.26	Strukturtoleranz 2	57
2.27	Strukturtoleranz 3	58
3.1	Schemadarstellung: 1NF, NF ² , NF ^{2D}	71
3.2	Beispielrelation mit dynamischem Attribut	71
3.3	Schemadarstellung: Selektion	73
3.4	Selektion mit einelementiger Diskriminantenmenge als Ergebnis	74
3.5	Selektion mit mehrelementiger Diskriminantenmenge als Ergebnis	74
3.6	Schemadarstellung: Projektion	75
3.7	Projektion mit einelementiger Diskriminantenmenge	76
3.8	Projektion mit mehrelementiger Diskriminantenmenge	76
3.9	Schemadarstellung: Join-Operation	77
3.10	Zweite Beispielrelation zur Verbundoperation	78
3.11	Ergebnisrelation einer Verbundoperation	78
3.12	Schemadarstellung: Varianzexpansion	80
3.13	Schemadarstellung: Nestung	81
3.14	Ausgangsrelation mit einelementigen Mengenattributen	81
3.15	Basisrelation für Nestungsoperationen	82
3.16	Schemadarstellung: Nestung mit Varianzexpansion	82
3.17	Ergebnisrelation der verschiedenen Nestungsformen	82
3.18	Schemadarstellung: Entnestung	83
3.19	Ergebnisrelation einer unbeschränkten Entnestung	83
3.20	Klassenrelation	90
3.21	Bildtupel $Bild_N$	91
3.22	Bildtupel $Bild_{N+z}$	92
3.23	Schematupel	93
3.24	Schematupel der Beispielbildfolge	93
3.25	QBE-Anfrage	96
3.26	QPE-Anfrage	98

4.1	Programm/Datenbank: Getrennte Lösung	105
4.2	Programm/Datenbank: Integrierte Lösung	105
4.3	Integriertes Konzeptuelles Datenschema	108
4.4	Verarbeitungsvorgang zur Verbesserung der Objektrepräsentation	114
4.5	Private Datenbankkopien	114
4.6	Datenversionen	114
4.7	Systemübersicht	124