References

[ADAB80]

[ADIB77]

[ADIB80a]

[ADIB80b]

[AHO79]

[ALSB75]
Alsberg PA, “Space and time savings through large database compression and dynamic restructuring”, Proc. IEEE, 63,8, August 1975, 1114–1122

[ANDE77]

[APER83]

[ARDI79]

[ARM80]

[ASTR76]

[ASTR80]

[BANC81]


[BATO82a] Batory DS, "Optimal file designs and reorganization points", *ACM Transactions on Database Systems*, 7,1, March 1982, 60–81


[BATO84b] Batory DS, "Modeling the physical structures of commercial database systems" to appear in *ACM Transactions on Database Systems*, also TR-83-21, Department of Computer Sciences, University of Texas at Austin, 1983


BERN81d

BERN83

BERT83a

BERT83b

BITT83

BLAC82

BLAS77

BLAS79

BLAS81

BOBR77

BONF83

BOOT76

BOR84

BOWE82

BOYC74

BRAC79


[CHAM76]
Chamberlin DD et al, “SEQUEL2: A unified approach to data definition, manipulation, and control”, *IBM J. Res. and Devel.*, 20,6, November 1976, 560–575

[CHAM81]

[CHAM81a]

[CHAN77a]

[CHAN77b]

[CHAN82a]

[CHAN82b]
Chang JM, “A heuristic approach to distributed query processing”, *Proc. 8th VLDB Conference*, Mexico City, September 1982, 54–61

[CHAN83a]

[CHAN83b]
Chan MHC, McDonell KJ, “An algorithm for translating updates on relational user views”, (submitted for publication), 1983

[CHEN76]
Chen PP, “The entity-relationship model: toward a unified view of data”, *ACM TODS*, 1,1, March 1976, 9–36

[CHEN83]

[CHEU81]

[CHEU82]

[CHIA82]

[CHIU80a]

[CHIU80b]


[CICS] IBM Corporation, CICS/VS General Information Manual, IBM Form No. GC33-0066


[CODD79] Codd EF, “Extending the database relational model to capture more meaning”, ACM Transactions on Database Systems, 4,4, December 1979, 397–434


[DANI82a] Daniels D, “Query compilation in a distributed database system”, *IBM Research Laboratory RJ3423*, San Jose, Calif., 1982


[DAYA85] Dayal U, “Query processing in a multidatabase system”, this volume,


[GHOS83b] Ghosh SP, “Future of the consecutive retrieval property”, in [GHOS83a]


[HAMM82] Hammond R, technical discussion, (Statistics Canada), 1982


[HEVN80] Hevner AR, “The optimization of query processing on distributed database systems”, Ph. D. Dissertation, Department of Computer Science, Purdue University, Lafayette, Indiana, 1980


[H0FF75b] Hoffer JA, “A clustering approach to the generation of sub-files for the design of a computer data base”, Ph. D. Dissertation, Department of Operations Research, Cornell University, January 1975, 276 pages


[H0FF80] Hoffer J, “Database design practices for inverted files”, Information and Management, 3, 1980, 149–161


[IBM] IBM Pub., SQL/Data System Application Programming, SH 24-5018
IBM Corporation, “Introduction to direct access storage devices and organization methods”, 
*Student Text, No. GC20-1649-8*, IBM Corporation, White Plains, NY, 1974


IEEE, *Database Engineering*, Batory D, (ed), March 1984


Jarke M, “Common subexpression isolation in multiple query optimization”, this volume


Kambayashi Y, "The buffer-limited quasi-consecutive retrieval file organization and its application to logic circuit layout", in [GHOS83a], 1983, 165–178

Kambayashi Y, "Processing cyclic queries", this volume


Kerschberg L, Ting PD, Yao SB, Optimal Distributed Query Processing, Bell Laboratories, Holmdel, NJ, 1980


Kim W, “On optimizing an SQL-like nested query”, ACM Transactions on Database Systems, 7,3 September 1982, 443–469. Also available as IBM Research Laboratory RJ3063, San Jose, California

Kim W, “Global optimization of relational queries: a first step”, this volume
[KING74]

[KING81]

[KLUG81]

[KLUG82a]

[KLUG82b]

[KNUT73]

[KONI83]
Konishi O, Kambayashi Y, “Improvement of the KWIC index utilizing the consecutive retrieval property”, in [GHOS83a], 1983, 289–299

[KO0I82]
Kooi R, Frankforth, D., “Query optimization in INGRES”, [REIN82b], 1982

[KOWA79]

[KOWA81]

[KUNI82]

[LAFU83]

[LAME84]

[LAND82]

[LANG78]

[LIND79]

[LIND81]
Lindsay BG, “Object naming and catalog management for a distributed database manager”, *Proc. 2nd International Conference on Distributed Computing Systems*, Paris, France, April 1981. Also available as *IBM Research Laboratory RJ2914*, San Jose, California, April 1984

[LIPS76] Lipski, Jr W, "Information storage and retrieval – mathematical foundation II (combinatorial problems), Theoretical Computer Science, 3, 1976, 183–212

[LIPS78] Lipski, Jr W, "Generalization of the consecutive ones property and related NP-complete problems", Rept. ACT 9, Coordinate Sci. Lab., University of Illinois, September 1978

[LIPS83a] Lipski Jr W, "The consecutive retrieval property, interval graphs and related topics – a survey", in [GHOS83a], 1983

[LIPS83b] Lipski Jr W, "A note on decomposing a query set into subsets having the consecutive retrieval property", in [GHOS83a], 1983

[LOHM84] Lohmann GM, Daniels D, Haas LM, Kistler R, Selinger PG, "Optimization of nested queries in a distributed relational database", IBM Research Laboratory RJ4260, San Jose, California

[LONC83] Lonc Z, Traczyk T, Truszczynski M, "Optimum f-graphs for the family of all k-subsets of an n-set", in [GHOS83a]


[LOWE82] Lowenthal E, "Database systems for local nets", DATAMATION, August 1982

[LUK80] Luk WS, Luk L, "Optimal query processing strategies in a distributed database system," Department of Computer Science, Simon Fraser University, Burnaby B.C., Canada, 1980


[LUM70] Lum VY, "Multiattribute retrieval with combined indices", Comm ACM, 13,11, November 1970


[MARC83a] see [MARC83]


[MARC84] March ST, “Physical database design: techniques for improved database performance”, this volume


[MEIE83]

[MERR77]
Merrett TH, “Database cost analysis: a top down approach”, ACM SIGMOD, Toronto, 1977, 135–143

[MERR83]
Merrett TH, Kambayashi Y, “Join scheduling in a paging environment using the consecutive property” in [GHOS83a]

[MINK78]

[MINK83]

[MINS75]

[MITC76]

[MITO75]

[Moha83]

[MOON81]

[MOTR81]

[MUNZ79]

[MUTH83]
Muthuswamy B, Kerschberg L, “Distributed query optimization using detailed database statistics”, unpublished manuscript, University of South Carolina, May 1983

[MYLO84]

[NAKA78]

[NAU82]

[NAVA78]

[NDX81]
[NG82] Ng P, "Distributed compilation and recompilation of database queries", IBM Research Report RJ3375, January 1982


[OZS84b] Ozsoyoglu M, Ozsoyoglu G, “A query language for statistical databases”, this volume
338


PERE82] Pereira LM, Porto A. “A Prolog implementation of a large system on a small machine”. Departamento de Informatica, Universidade Nova de Lisboa


[ROTH77a]

[ROTH77b]

[ROTH80]

[ROUS82a]

[ROUS82b]

[ROWE79]

[SACC82]
Sacco GM, Schkolnick M, “A mechanism for managing the buffer pool in a relational database system using the hot set model”, Proc. 8th VLDB Conference, Mexico City, 1982, 257–262

[SAGI80]

[SAGI81a]

[SAGI81b]
Sagiv Y, Optimization of Queries in Relational Databases, UMI Research Press, Ann Arbor, 1981

[SANT81]

[SCH82]
Scheck HJ, Pistor P, “Data structures for an integrated data base management and information retrieval system”, Proc. 8th VLDB Conference, Mexico City, 1982

[SKCH75]

[SKCH77]

[SKCH79]

[SKCH79a]

[SKCH79b]


[SENK69] Senko ME et al, File Design Handbook, IBM San Jose Research Laboratory


[SPER75] Sperry Rand Corporation, DMS 1100 Schema Definition: Data Administrator Reference, 1975


[STAN73] Stanford University, Design of SPIRES: Vol I and II, Center for Information Processing, Stanford University, 1973


[STON80] Stonebraker M, “Retrospection on a data base system”, ACM Transactions on Database Systems, 6,3, September 1980


[TANA83] Tanaka K, “Tree-structured data organization with consecutive retrieval property”, in [GHOS83a]


[VANM79] van Melle W, “A domain-independent production rule system for consultation programs”, *Proc. 6th International Joint Conference Artificial Intelligence*, 1979


[VASS85] Vassiliou Y, Clifford J, Jarke M, “Database access requirements of knowledge-based systems”, this volume


[WHAN84b] Whang K, “The property of separability in physical design of network model databases,” *Information Systems*, 9,4


[WILM81] Wilms PF, Lindsay BG, “A database authorization mechanism supporting individual and group authorization”, *Distributed Data Sharing Systems*, van de Riet and Litwin (eds), North Holland, May 1981, Also available as *IBM Research Report RJ3137*

[WILM83] Wilms PF, Lindsay BG, Selinger PG, ‘‘I wish I were over there’: distributed execution protocols for data definition in R*, *Proc. ACM-SIGMOD*, San Jose, California, May 1983, 238–244. Also available as *IBM Research Laboratory, RJ3892, San Jose, California*


[YAO77b] Yao SB, “Approximating block access in database organizations”, *Communications ACM*, 20, April 1977, 260–261


[YU79a]
Yu CT, Lam K, Ozsoyoglu M, “Distributed query optimization for tree queries”, Dept. of Information Engineering, University of Illinois at Chicago Circle, October 1979

[YU79b]

[YU82a]

[YU82b]
Yu CT, Chang CC, Chang YC, “Two surprising results in processing simple queries in distributed databases”, IEEE COMPSAC, November 1982

[YU83]

[YUE75]

[ZANI79]

[ZLOO77]
List of Authors

Don S. Batory  
Dept. of Computer Science  
The University of Texas at Austin  
Austin, TX 78712  
USA

Marco A. Casanova  
Centro Cientifico de Brasilia  
IBM do Brasil  
Brasilia  
Brasil

Jim Clifford  
Graduate School of  
Business Administration  
New York University  
90 Trinity Place  
New York, NY 10006  
USA

Dean Daniels  
Carnegie-Mellon University  
Pittsburgh, PA 15213  
USA

Umeshwar Dayal  
Computer Corporation of America  
Four Cambridge Center  
Cambridge, MA 02142  
USA

Anthony L. Furtado  
Departamento de Informatica  
PUC/RJ  
Rio de Janeiro  
Brasil

Sakti P. Ghosh  
IBM Research Laboratory  
San Jose, CA 95193  
USA

Laura M. Haas  
IBM Research Laboratory  
San Jose, CA 95193  
USA

Matthias Jarke  
Graduate School of  
Business Administration  
New York University  
90 Trinity Place  
New York, NY 10006  
USA

Yahiko Kambayashi  
Dept. of Computer Science and Communication Engineering  
Kyushu University  
Hakozaki, Higashi, Fukuoka, 812  
Japan
Won Kim
MCC – Microelectronics &
Computer Technology Corporation
9430 Research Blvd.
Austin, TX 78759
USA

Jürgen Koch
Fachbereich Informatik
Johann Wolfgang Goethe Universität
Dantestraße 9
D-6000 Frankfurt 1
West Germany

Bruce G. Lindsay
IBM Research Laboratory
San Jose, CA 95193
USA

Guy M. Lohman
IBM Research Laboratory
San Jose, CA 95193
USA

Raymond Lorie
IBM Research Laboratory
San Jose, CA 95193
USA

Salvatore T. March
University of Minnesota
Minneapolis, MN 55455
USA

Dan McNabb
IBM Research Laboratory
San Jose, CA 95193
USA

Andreas Meier
IBM Research Laboratory
San Jose, CA 95193
USA

Present Address for Meier:
Institut für Informatik
ETH Zürich
CH-Zürich
Switzerland

C. Mohan
IBM Research Laboratory
San Jose, CA 95193
USA

Gultekin Ozsoyoglu
Department of Computer
Engineering and Science
Case Western Reserve University
Cleveland, OH 44106
USA

Z. Meral Ozsoyoglu
Department of Computer
Engineering and Science
Case Western Reserve University
Cleveland, OH 44106
USA

Wil Plouffe
IBM Research Laboratory
San Jose, CA 95193
USA

David S. Reiner
Computer Corporation of America
Four Cambridge Center
Cambridge, MA 02142
USA

Arnon Rosenthal
Computer Corporation of America
Four Cambridge Center
Cambridge, MA 02142
USA
Daniel Sagalowicz  
Framentec  
Monaco  

Joachim W. Schmidt  
Fachbereich Informatik  
Johann Wolfgang Goethe Universität  
Dantestraße 9  
D-6000 Frankfurt 1  
West Germany  

Patricia G. Selinger  
IBM Research Laboratory  
San Jose, CA 95193  
USA  

David Elliot Shaw  
Dept. of Computer Science  
Columbia University  
New York, NY 10027  
USA  

Michael Ubell  
Britton-Lee, Inc.  
14600 Winchester Boulevard  
Los Gatos, CA 95030  
USA  

Yannis Vassiliou  
Graduate School of  
Business Administration  

New York University  
90 Trinity Place  
New York, NY 10006  
USA  

Kyu-Young Whang  
IBM T. J. Watson Research Center  
Yorktown Heights, NY 10598  
USA  

Gio Wiederhold  
Stanford University  
Computer Science &  
Electrical Engineering  
Margaret Jacks Hall  
Stanford, CA 94305  
USA  

Paul F. Wilms  
IBM Research Laboratory  
San Jose, CA 95193  
USA  

C. T. Yu  
Dept. of Electrical Engineering and  
Computer Science  
University of Illinois at Chicago  
Chicago, IL 60680  
USA
Subject Index

ABE (Aggregates-by-Example) 171, 182, 183, 185

Abstract
data type 265, 272, 275
and views 127, 128, 132, 142
objects 266
records 265–266, 268, 271–272

Access
authorization statements 34, 36
cost 303, 305
configuration 297, 298, 309, 315
see also Access paths, cost computation of
in IDM 241
path
abstract representation 199
chained 283
defined by query 193
design 279, 280, 286, 292–295
hashed 283
primary 283, 294
selection 186, 295
structures 262
tree 312, 313, 314, 317

paths
for complex objects 150
CODASYL 103, 292, 302
cost computation of 103–106, 120, 219, 292–293
defined through selectors 202
in IMS 34
optimal 3, 43, 102, 150, 215, 280
physical 201, 282
and query optimization 10, 192
for repetitive queries 191
selection 185
for simultaneous processing 213
and subexpressions 194
planning 3, 14, 19, 22–24
of records, as basic operation 276
in R* 37, 41, 43
relation schema 110, 113
relations 109, 110, 113, 120
structure selection 297
structures 297, 299, 300, 310
in R* 37, 41
via set 121
see also Access paths, CODASYL
ADABAS 171, 265
ADAPLEX DDM,
and cost estimates 98
ADAPLEX LDM 85
Aggregate functions 88, 171, 172, 175, 177, 185, 240
Aggregates-by-Example see ABE
Aggregation 81, 99, 184, 282, 285, 287, 288, 296
AI see Artificial intelligence
Algebra see Relational algebra
Algebraic
methods vs nested loop solutions 11
simplification 119
Appendix, for complex object
implementation 148
Application domain knowledge 157, 158
oriented operations 132, 133
Apprentice site, in R* 39, 44
Artificial intelligence 26, 27, 156, 157, 158, 165
Assembly phase of distributed processing algorithm 53, 56
Associative processing 257
Attribute
addition, as procedure for cyclic queries 72–73, 77
descriptors 280
selectivity 51
Attributes
outside view 131
in summary table 173
Authorization 40, 41
Automated software development 275
Automatic
design of physical database 298
mandatory set
translation, of views 134
Backend machine 22–23, 237, 238, 249
Backjoin 201
Base relations 109, 110
Batching 181, 202, 206
Bicriterion programming 289, 291
Birth site, in R* 35
Block access 51, 244, 273, 284, 291, 300, 305, 316
Bloom filters 186
Branch and bound 227, 289, 295
Britton-Lee Corporation 237, 247
B+ trees 262, 268, 303
B-trees 120, 121, 242, 245, 246, 247, 249, 278
Buffer 10, 231, 232, 233
management, and physical databases 262
CAD 27, 164, 261
Cartesian product 11–13, 44
Catalog management, in R* 35–36, 40
Category attributes 173, 186
Chain queries, and access planning 24
Chained access path 283
CICS (IBM's Customer Information and Control System) 33
Clustering 192, 289, 297, 300, 303
CODASYL
and access paths 103, 292, 302
model 132, 285
see also Network model
query optimization for 81, 101
relational queries of 109
schemas, and integrity
information 116
as semantic network 160
tree queries 102, 103
Commercial database, physical designs 285
Common subexpressions 18, 26, 191, 193, 194–196, 197, 198, 204
Communication
cost 3, 46, 31, 49
host to IDM 238
steps, in multidatabase system 81
Compilation 45, 122
Compile stage 207
Compilers 275, 277
Compile-time analysis 215
Complex
fetch 151, 152, 153
object 145–150, 153–154
Computer Corporation of America see ADAPLEX
see also MULTIBASE
see also SDD-1
Computer systems engineering 248
Conceptual schema 127
Conceptual-to-internal mappings 261, 264, 265
Concurrency
control 238, 242, 262, 277
see also locking
Condition box 177
Conjunctive
normal form, and INGRES 16
queries 8, 49, 90–92, 198
relational calculus, 8, 183
subqueries 18
Conjuncts 119
Consecutive retrieval
    basic concepts of 218–221
    file organization 218, 220, 226, 228–230
    property 217, 221–224, 225
    see also Quasi-consecutive retrieval
Consistency, of data base update 128
Constraints 5, 130, 132, 133
Content addressable memory 253
Copy selection, in distributed query
    processing 58
Cost
    and access paths 103–106, 120, 219,
    292–293
    of CPU time 49, 150
    of data transmission in distributed
    systems 67
    and database machines 249
    and database maintenance 282
    of disk accesses 49
    estimation 25, 97, 103–104
    as linear combination 31, 43, 46
    models 105, 121, 128
    of page I/O 150
    of partial join 308
    of partial operation 315
    reduction 213–214
    and storage structure 303
    of transactions 299
    system operating 279
Coupled set 308
Coupling
    effect 304, 305, 306, 307, 308
    factor 307, 308, 310
CPU, and costs 31, 43, 46, 49
CREATABLE 265
Cyclic
    queries 21, 55, 64, 66
    query graph 55, 68, 69
    query processing 62, 67, 72, 73
DAPLEX 82, 83, 85, 90, 101, 102, 108
Data
    access
        of expert systems 165
        parallelism of 191
        paths 296
        requirements 163, 164, 166
        block 241
        composition 233
        definition statements 31, 36
        dependence, and nested queries 209,
        211, 212
        dependencies 64, 135
        dependent queries 212
dictionary 14, 186, 241
extraction 122
independent query 212
level, in expert system 157, 162
management functions 238
manipulation statements 31, 151
see also Queries
mapping 166, 276
models, recursive 27
reduction 100
redundancy 287
retrieval 165
security 296
sharing 244, 246
structure
    design 233
diagram 263, 264, 266, 273, 276
transfer, semi-joins to reduce 48
transmission 76, 77, 78, 97
volume, of knowledge bases 163
Database
    access, by expert system 164, 166–
    167, 168
    administration statements 31, 34
    administrator 268
    conversion, as alternative to query
        translation 122
    inconsistency 81
    integration 82, 83, 84, 85
    see also Logical database integration
machines 23, 237, 249
object 34–35
programming language 5
reorganization 149
schema 134
state 128
statistics 24, 25
Database Accelerator, of IDM 237, 241, 242, 243, 244, 245
Database Processor, of IDM 244
Datagrams, asynchronous 42
Datasets 281, 284
DB2 32
DBTG model 297, 317
see also CODASYL model
DC* (Distributed Communications), of R* 33, 38
DDM see ADAPLEX DDM
Deadlock detection 42
Decomposition 10, 69–70
Dedicated environment 237, 242
Deductive databases 170
Degenerated multi-valued dependency (DMVD) 64, 75
Deletion 136, 252, 255, 276, 302
Dependencies 74–75
Dependency recording 41
see also type (Multi-valued,
Degenerated multi-valued, Data,
Placement)
Dependent entities 114
Derived relations 127, 132
Design
data, how stored 145
issues 262
objects 145
theory of separability 298
Differential files 262
Direct
access 149, 217, 221, 227
accesses 218, 220, 232
Directed graph see Graph
Disjunctive prenex normal form (DPNF) 8, 11
Disk
accesses 49, 103, 242
devices structure, and graph representation 225
storage 284
Distributed database
management systems 3, 101
and query processing 21–23, 76–77
system 62, 84
see also Heterogeneous distributed database systems
query
compilation 38–40
processing 22–23
Distributed INGRES 45, 46, 47
DMS–1100 (Sperry) 265, 286
Domain relational algebra 191, 192
calculus 195–196
Domain rules 158
DOS/VSE Operating system 32
Duplicate elimination 111
records, used by programmers 113
Duplicates Not Allowed (DNA) 112, 117
Enabling, selective 252
Encapsulation 133, 134
Encoded records 265, 273
Engineering databases 145, 261
Entity
cardinality 280
instances 280
types 82, 85, 86
Entity Relationship (E–R) model 112
Equi-join 63, 65, 154
Equivalence transformations 54, 66
Equivalent relational schema, of network schema 110
Estimation
algorithm, for distributed query processing 50–52
static vs. dynamic 52, 54
Ethernet 244
Execution
  phase, of R* processing 38–39 plans 104
Exhaustive
  enumeration 100 search 298
Expert Systems 26, 156, 158, 163, 165, 166
Expressions see Common subexpressions
Extended range expressions 199, 200
Extended Relational Algebra (ERA), 181
External
  projection 256 records 229 schema 127, 128
F-graphs 223, 224
Field definition diagram 263
File
  allocation 224, 233 horizontally partitioned 266
  modeling 297 organization 224, 282, 294 reorganization 262
  structures 261, 262, 263, 269, 278 see also Linear structure file
First-order logic 159
Fixed
  members, in CODASYL terminology 132
  variables 175, 176, 177, 179
Fragment processing 46, 48, 59
Fragmented distributed database environment 58
Frames 161
Free variables 175, 177, 182
Frequency
  of database entry 317 of query 227, 230 of traversals 313
FRL 162, 166
Full indexes 283, 292
Fully reduced relation 53
Function types, as edges in directed graph 82
Functional dependencies (FDs) 75, 129, 132, 135, 182
Generalization, for logical database integration 81
Global
  execution plan 81, 98 file organization design 295
  knowledge, of database 206 optimization strategy 84, 89–91, 207
  planning, in R* for inter-site 44 strategy, for processing queries 215
  view 82, 83
Global Data Manager (GDM), of MULTIBASE 84, 101, 108
Graft and prune operations 13
Graph
  acyclic 223 bipartite 221
  chain 223 for consecutive retrieval 217, 221–224
  cycle 223 deBruijin 221
degree constraint directed 223 directed 51, 174, 221
directed acyclic 178 of disk devices structure 225
  interval 221 planar 223
  regular 223, 224, 225 structures, vs linear structure 217
  see also Query graph, F-graph, View dependence, Object, Operator
Guard conditions 19
Hamiltonian path 221, 223
Hardware backend, see Backend machine Hashed
  access paths 283, 286 files 262, 263, 269, 292
  organizations 297
Hashing 120, 251, 255, 256, 257, 292
Heterogeneous databases 26, 81, 82, 101
Heuristics
Hierarchies
aggregation 287
data model 285
Hierarchical
processing, of query 77
sequential lists 263
Hierarchies
generalization, for logical database integration 81
in relational framework 146
Hill climbing 100, 107, 289
Homogeneous networks 24, 25
Horizontal
decomposition 71, 72, 75, 76
fragmentation 46, 59
Horizontally distributed databases 24, 25
partitioned 84, 92
Horn clauses 159
Host computers 237, 249, 251
Host Interface Channels, for IDM 242–243
Human resources, determining costs of 3
Hypergraph 221, 226, 228
IBM San Jose Research Laboratory
Distributed DB/DC Lab 34
ICL CAFS 237
IDM (Intelligent Database Machine) 237
architecture of 237
data management system 238–241
Disk Controller interfaces 245
hardware of 242, 243
Host Software 237, 238
performance statistics of 245
and query languages 240
IDMS 265
IDS (Honeywell) 297
IMS 164, 285, 287
Inclusion constraints 116, 117
Inconsistent databases 82
Independence assumptions 48
Independent entities 114
systems 165
Index
B-tree, in IDM 242
clustered, in IDM 242
encoding technique 271
on identifiers 149
scan 151, 303
selection 262, 292, 297
see also also Access path
Indexed
aggregate file structure 269
sequential files 263, 278, 283
vs. sequential access 43
Indexes
and costs 207, 215, 303
for complex fetch 152
in IDM 241
as query optimization method 10
secondary 206, 215
Inference engine 157
rules 163
Influence, mutual, among views 131
Information
processing requirements 279
systems analysis and design 157
INGRES 85, 171, 261, 262, 263, 264, 265, 266, 267–270, 278
and access planning 24
and conjunctive normal form 16
decomposition algorithm of 10
view definition in 127
see also Distributed INGRES
Inner
Outer-Loop Join method 301, 310
select statement, in IDM 241
table transfer strategy, in R* 44
Inorder traversal 262
Insertion 252, 255, 276
Instance diagram 263
set 180
trees 180
Integrated
    solutions 13
    systems 165
Integration, database 81
Integrity constraints, of data 4–5, 19, 109, 116–117, 128, 139
Intel iDBP 237
Intelligent
    Database Machine see IDM
    disks 248, 257
    head units 251, 256
Interface, IDM to host 238
Interference, among views 131, 132
Intermediate
    relations 12
    results 48, 50, 185, 189
    resolutions 298
Intermodel
    translation 112
    see also Relational queries of network database
Internal
    file, of physical database 263
    link, of physical database 263
    records 229, 276
INTERNIST 160, 163
Interpretation 122
    see also Compilation
Intersection, on NON-VON 252, 254
Inter-site communication 42, 43
Introduce Database statement 37
Inverted
    access paths 283, 292
    file 262, 263, 264, 283, 292
    list see Inverted file
I/O, as cost 31, 43, 46, 242
ISAM files 120, 286, 292
Iterative grouping refinement for record segmentations 289, 290
Join
    clauses, and conjunctive queries 90
    computing 185
costs 121
dependency, definition of 64
graph of qualification of query 52
and hardware devices 23
index 300, 304, 306, 308, 309, 317
inner/outer method of 301, 310
methods, 40, 310, 317
on NON-VON 252, 254
operator, of relational algebra 11
order of 20
predicate, as link relation 113
query, illustration of 98
in relational databases 318, 300
of relations, to convert query
    graphs 68–69
selectivity 306
simplification to restrictions 18, 19
site 44, 100
stage 209
strategies, in IDM 241
term 7, 15
unconditional 306
value-based 102
see also individual types (Semi, Equi, Natural)
Key 5, 113, 116
    for complex object 154–155
    disjoint partitioning 256
    in network Relational translation 117
oriented multiple query
    processing 193
selection, and inverted files 292
KLONE 167, 173
Knowledge
    bases 156, 157
    level, of expert system 157, 158, 162
    representation 158, 161, 162
KRL 162, 166
KWIC 233
Language
    power vs. fast response time 3
    processors 238
LDM see ADAPLEX LDM
Linear structure files 217, 225
Link
   predicates 110, 123
   relations 113
Linkage factor 312
Links
   actualized 266
   in network models 311
Linksets 263, 264, 272, 281
LISP 158, 173
Lists 121, 289, 292
Local
   data vs. remote data, in R* 32
   host schemas 82
   processing costs 98
   schemas, in multidatabase system 81
   selections, and data reduction 100, 102
Local Data Interface (LDI) 84, 85, 101
Locality of reference 217
Locking 204, 238, 258, 275, 277
   see also Concurrency
Logarithmic-stage networks 251
Logging transactions 258
Logic
   circuit layout 233
   first-order 159
   programming 167, 174
   see also Prolog
Logical
   database 81, 279, 280, 281
   object 297, 299, 311
   transformation to query, as
   optimization 13, 14
Loose coupling 165, 166
Machine resources, definition of 3
Mandatory sets, automatic 112, 117
Map
   of complex object 149, 150
   scans 151
   of tuple 149, 150
Mapping
   data 276
   of query 14, 109, 110
   relational updates over network
   data 112
see also Conceptual-to-internal mappings
Mass storage systems 225, 233
Master site, in R* processing 39
Mega-database, of access plans 14
Memory
   buffer 209, 217, 218
   see also Buffer
   system, of IDM 245
Merge
   condition 83, 86, 93
   join 12
   scan 185, 186
Metadata field 272
Meta-knowledge 158
Migrate Table statement 34, 37
Modeling, of physical database 278
Modular design, of data base
   schemas 134
Move operations, in multidatabase
   queries 84
MULTIBASE 25, 81, 82, 84, 85, 87, 97,
   100, 122
Multidatabase system see MULTIBASE
Multilists 263, 282, 292
Multiple
   accesses 225, 227
   copies of relations 48
   knowledge representation 162
   query optimization 26, 191, 193, 203,
   217
   queries, and consecutive retrieval
   files 218
   SIMD execution 251
Multiuser environment 246, 258
Multi-valued dependency (MVD) 63, 64
Multiversion database 205
MVS Operating System 31, 33
MVS/370 Operating system 32
MVX/XA Operating system 32
Naming, in R*, 34, 39
Natural
   equi-join 63, 65, 68
   language processing 160
Navigational database systems 81

Nesting
abstract data types 265
aggregations 183, 184
expressions 20, 198, 200, 201–202
languages 240
loops 3, 9, 11, 185, 186, 242
queries 209, 210–211, 212

Nesting 202

NETL 160, 166

Network
model 109, 112, 298, 310
schema diagram 115
theory 291
topology 23, 46
see also CODASYL

Networks,
for expert systems 160
high-bandwidth 251
local vs. geographically dispersed 49
Node orders 103, 104, 107
Nonseparable join methods 310
NON-VON (Non vonNeumann) 248, 249, 250, 251–252, 257
Normal form, non-first 121, 172
Normalization, cost of 121
Normalized
DAPLEX non- 91
relations 186, 268
Null
transformation 32, 265, 269
undefined values in view 131
Nulls Not Allowed (NNA) 117, 178
Object graphs 15, 21, 200, 202
Ocsets see Owner-coupled sets
Offload 247
Operating kernel, in IDM 242
Operator graphs 15, 16
OPS–5 157
Optimization
of data access for expert system 168
of queries see Query optimization
Optimizer see Query optimization
see also Subsystems, query processing

Order
information-bearing 113
of operations 185
of records 221
Outerjoin 81, 86, 89, 99, 113, 118
and aggregation 87, 90, 91, 92
Overlap portion 85, 93, 95
Overlapping subsets 219, 220
Owner-coupled set 101, 103
Page
access 106, 242, 287, 290
fetches 209, 213–214
see also Random access
Parallel
local processing at multiple sites 59
machine, and query optimization 258
processing, in relational algebra
operations 13
string comparison 252
supercomputer 248
Parallelism, of data access 191
Parametric solutions 291
Parse tree 42, 185
Parsing 39, 241
Partial
join 308
result 65
see also Intermediate result
selection 96
Partitioning 317
Pascal/R 15
Performance
evaluation of R* 47
of IDM 244
prediction of 262
Personal
computer 243
databases 47
Physical
database 276, 282, 296
and costs 315
decomposed 263
design 279, 297, 317
modeling 278
software 261
structure 22
processing, for network data 120
record structures 281
Placement dependencies 60–61
PLANNER 158
Pointer arrays 121, 263, 278
POLYPHEME 46
Post-order traversal 240
Postprocessing queries, in
multidatabase 84
Precompilation 37
Precompilers 238
Predicates 119, 150, 196
Prenex normal form 8, 15
Preparation phase, of R* processing 38
Preprocessing 120, 208, 214, 241
Private
data area 244
portion 93, 95
PROBWELL 164
Procedural
knowledge representation 158
rules 163
Production rules 159, 165
Projection 11, 12, 63
and data reduction 100
and database integration 85
estimates of size of 25–26
and fragmentation 59
on NON-VON 252, 253–254
in relational databases 300
sequence of 19
in SQL 44
Prolog 157, 159, 162, 166, 192, 196–198
based expert system 169
Propagation of constraints 157
Prospector 160
Protection views 47
Prototype DBMS 262
Prototyping 157
QBE 6, 98, 127, 171
Qualification 49, 52
of records 239
Quantifier movement rule 16
Quasi-consecutive retrieval 217, 218, 231–232
QUEL 6, 192, 207
Query
block basis 186
chain 24
conjunctive 8, 49, 90–92, 198
evaluation algorithms 8
frequencies of 227
graphs 54, 62, 65, 68, 78, 90, 200
grouping 210
and join 151
language, IDM independent of 240
see also language name
locally processable 60
modification 83, 87, 89
optimization 3, 40, 41, 45, 46, 192
global 84, 89, 90
heuristics for 3, 99–101
and host machine 251
on parallel machine 258
in R* 43
subsystem 14
on views 17, 99
see also Query processing
parsing, and host machine 251
processing,
and automated software
development 275
in distributed databases 23, 48, 76, 77
84, 91
local, in multidatabase 84
and physical databases 262
subsystem 4, 206
see also Query optimization
processor 4, 5
as readset 193
repetitive 191
as request for data selection 3
set 226, 228, 231, 232
simplification rules 118
standardization and simplification
of 15–17
site 35
transformations 3, 14
translation 117, 118, 122
in update transactions 8
see also Data definition statement, Data manipulation statement
Query by Example see QBE
R* 31–34, 37–39, 47, 97, 98, 100, 164
R1, expert system 164
Random access environment 192
Range
box 177-179, 183, 184
formulas 181
nested expressions 20, 21
nesting 202
relation 7
variable 240
RAPID 261, 263–266, 271–275, 278
RDS* 33, 38
Record
access 106
– active 313, 314
augmented 265, 272
collected 265
devided 265
fetched level interface 241
fields extracted from 266
instance 281
linking mechanisms 261, 278
modification of 276
oriented 191, 192
segmentation 281, 282, 286, 288–290
selection criterion 283
structures 282, 285
structuring 279, 280, 286, 295, 296
types 110, 267, 276
Record Package 166
Record Storage System (RSS) 241
Recovery algorithms 262
Reduction phase 52, 209
Redundant records 225, 227
Reference value, for complex object 147
Relation
component, and data retrieval from 153
instance 173
scan 185, 242, 256, 303, 305
schema 128, 134, 173
skeleton 176, 179
Relational algebra
and aggregate functions 172
algorithms, for query execution 3
extended 86, 90
operators 253, 256, 257
set operations of 10, 11
and subexpressions 8, 191, 192, 194, 195
for view mapping 129
calculus
and aggregate functions 172
and queries 5–8, 11
query language, extended 181
and simplification rules 18
and syntactic query simplification 18
and tableau representation 15
completeness 8, 181
database management 237, 249
systems, and separability 298
interface schema 113
invariant, for view update 142
model 5, 146, 285
extended for statistical databases 171
queries, of network data 110, 111, 149
query modification algorithm 85
optimizer 123
schema, for a network database 112
views 109, 117
Relational Technology 267
Relations
multiple copies of 48
sorted copies of 206, 215
Relationship descriptors 280
Relationships, in view-relation
schemas 114
Replication 47, 59
Report writers 238
Research Storage Interface 238
Residual predicate 300
Response time 46, 76
Restricted set 308
Restriction
in SQL 44
index 300, 305
operations 11, 19
Restrictive terms, of selection
expression 7
Result
set 308
size 99
see also Intermediate result
Retrieval
and cost 279, 282, 284
of record 276
Ring lists 263
Root tuples 146, 148
RSS* 33, 37, 38
Run stage, of simultaneous
processing 207
Running times, for relational
operators 254, 255
Safe Boolean function 218
Scanning, via sets 120
Scans
on access relations 110, 120, 123
of map entries 151
see also Relation scans
Scatter tables 284, 292
Schema generating algorithm 121, 122
Screen
handlers 238
oriented query language 171, 172, 182
SDD-1 (Computer Corporation of
America) 15, 25, 45, 46, 48, 57,
100
Secondary
copies of relations 58
storage 163, 225, 284
Segment scan 150, 151
Segmented
field 272
records 265
Selection
correct distribution of 94–95
equation, as first order predicate 5
of entities, and database
integration 85
and horizontal fragmentation 59
on NON-VON 252, 253
partial local 95, 96
Selectivity, attribute 51
Selector Language Construct 198, 199,
201, 202
Semantic
integrity constraints 148
networks 160
query
optimization 10, 19
transformations 14
Semijoin 43, 121
advantages of 67
and data reduction 48, 49–50, 100
effect on amount of data
transmitted 76
estimation 52
and fragment processing 59
and hardware devices 23
illustration of 50
mapping of 21
operation 13
in R* optimization 46
reduction 93
sequences of 48, 55, 57
and table deletion 119
as tree path 77
vs. joins 62
Semiouterjoin 93–94, 95, 97, 99, 100
Separability 297, 298, 299, 300, 304, 310,
311, 316, 317
sequential
files 292
see also Indexed sequential files
processing, and batching 192–193
Set

difference, on NON-VON 252, 254–255
implementations 121, 311
key constraints, in access relation
schema 117
links 103
oriented query 193
of queries 206
restricted and partially restricted 308
scanning 120
type, as access relation 110
types of for separable database 308
valued attribute 171, 172, 173
valued functions 91
SGL 240
Simple
valued attribute 173
views, as single relation scheme 141
Simultaneous processing 206, 207
Single
relation 129, 209, 300
site queries 81, 84
system 32
Site
autonomy 33
logical database as 32
selection 53, 58
see also Birth site
Skeletons 174, 176, 179
Snapshots 47, 75, 208, 215
Software
automation 277
backend 23
design 296
Sophie 160
Sorted copies of relations 206, 215
Sorting
and cost 303
merge 300, 303, 304, 308, 309, 317
of relations 121, 207, 255
and sites 99
Spanning tree 62, 73, 76
Speech recognition 162
SPIRES 265
SQL 151, 182, 184, 192, 207
and access planning 24, 25
–, augmented by link predicates 120
extended to non-first normal form 121
extension to 147
from Prolog predicates 169
in R* 31, 34, 35
queries 6
standardization in 15
syntax, and new predicates 110
translation of to network database 109
view definition and 118, 135
SQL/DS 32, 127, 171
Star networks 24
State-space search 157
Statistical databases 27, 171, 261, 271
Statistics Canada 271
STBE 171, 180, 182, 183
Stepwise reduction, in nested
expressions 21
Storage
accesses, and intermediate results 26
see also Access plans
architecture 261, 266, 269–270, 273,
274, 275, 278
and cost 279, 284, 303
records 286
space, and consecutive retrieval file
organization 220
Store site, in R* 35
Stored relations 127, 131, 132
STROBE 162, 166
Subexpressions 191, 203
see also Common subexpressions
Subfile 281
Subqueries 21, 93, 174
Subrelations 10
Subset constraints see Inclusion
constraints
Subsets, predicted usage of 290
Subsystems 4, 206, 248, 249
Subtuple 300
Summary table 171, 173, 174, 176, 177, 178, 180, 181, 186
Summary-Table-by-Example see (STBE)
Supercomputers 248, 249, 252
Syntactic query 14, 17–18
SYSTEM 2000 265
System Network Architecture (SNA), 33
System operating costs 279
System R 85, 127, 185, 238, 241, 264, 265, 286
and access planning 24–25
compilation in 45
extended by R* 31, 33
optimizer of 150, 151
query language see SQL
TID in 149
view updates and 138
System R* see R*
System-for-Statistical-Databases (SSDB) 172
Table Producing Language (TPL) 171
Tableau representation 15, 18
Table
access 43, 44
deletable 119
replicated and partitioned 47
Target
component 49
list 5, 12, 118, 181, 184, 208, 239
Text processing 27
Textual databases 261
Throughput 246
Tight coupling 165, 166
TM* (Transaction Manager), of R* 33, 38
TOTAL DBMS 265
Transaction
abortion 205
commands, in DM 241
cost 297, 299
evaluation 300
file 192
logging 258
as logical unit of interaction 3
management 4, 203
processing 299, 317
Transformation functions 313
Transformation Model 261, 262, 264, 265, 266, 278
Translatability, of tuple insertions, deletions 141
Translatable views 140
Translation
algorithm, example of 11
into local host Languages 84
view 131
Translator induced by view 140
Transposed files 262, 276
Traversal strategy 123
Tree
graphs, and consecutive retrieval property 223
in IDM architecture 240
minimum distance 77
of processes, for inter-site communication in R* 42
query 21, 24, 48, 55, 66–67, 68, 102
query-graphs, algorithm for recognizing 55–56
structure and data organization 287
structured architecture 248
vs. cyclic queries 53–57
see also B-trees
trie 284
Trivial segmentation 290
Tuple
calculus, and common subexpressions 198
clustered, in complex object 146
heterogeneous 155
identifier field 113
links for 148
ordering of, in relational model 112
packed 253
parent 150
relational algebra 191, 192
spanned 253, 276
unique in sets 112
variables, in query simplification 119
Unifying Model 261, 262, 264, 276, 277, 278
Union, on NON-VON 252, 254
University of California, Berkeley 267
Unordered file structure 269
Update
costs 279, 296, 310, 316
log 241
operations 204
of redundant records 227
semantics 111
statement 208
templates 142
transactions 302
translation 109, 123
of tuple, on NON-VON 252, 255
Usage
catalog 41
information 312, 313, 317
User
environment 204
interface 251
and updates 133, 134
view, in multidatabase system 81
vs. machine cost trade-off 4
Value-based joins 102
Vertical
decomposition of relations 71–72
fragmentation 46, 59
Very large databases 156, 257
View
as abstract data type 127
based translations 113
complement 127, 135, 138–141
composition, by view definition 42
definition 127
and database integration 85
facility 127
by generalization 88
by joining 86
language, extended 128
mappings 128, 135
in R* 41–42
resulting from translation 130
site 41
specification of translation to stored
relations 132
dependency graph 136–138
integration, generalization for 81
interference 132
materialized 87
over views 142
processing see View substitution
queries, and common
subexpressions 197
for query and authorization 142
in R* 34
recursive 197
relation definition 113, 114, 117
single-relation 129
state, definition of 128
substitution 17, 27, 109, 118, 119
translation procedures 128
updates 112, 127, 128, 134, 135
Virtual
access paths 284, 292
circuit or session, for R* sites 42
Virtual Machine (IBM) 245
Virtual Sequential Access Method (VSAM), 33, 286
Virtual Telecommunications Access Method (VTAM), 33
VLSI chip, as design object 145
Windows 174, 177, 179
Write ahead log 241
On Conceptual Modelling:
Perspectives from Artificial Intelligence, Databases, and Programming Languages

Editors: M.L. Brodie, J. Mylopoulos, J.W. Schmidt


Contents: Artificial Intelligence, Database, and Programming Language Overviews. - Perspectives from Artificial Intelligence. - Perspectives from Databases. - Perspectives from Programming Languages. - Concluding Remarks from Three Perspectives. - References. - Authors and Symposium Participants. - Index.

Conceptual modelling relates to all areas of computer science, but especially to artificial intelligence, databases, and programming languages. Here is the first published collection of state-of-the-art research papers in these domains. Its purpose is to consider conceptual modelling as a topic in its own right rather than as an aspect of data modelling, and to present and compare research on knowledge representation, semantic data models, and data abstraction in this context.

The contributions consist of overviews and reports, each chapter having been written and edited for readers in all three areas. Also included are transcripts of symposium discussions which took place among the contributors during a workshop on conceptual modelling at Intervale; these interdisciplinary discussions of each paper clarify many aspects which might otherwise remain obscure to non-specialists. Key features of the book include introductions to pertinent concepts, and the integration of recent results; focus on twelve research projects, involving specific applications such as database design; and challenging suggestions for further research, especially in the concluding comments by leading experts in the three main fields of inquiry.
Relational Database Systems

Analysis and Comparison

Editors: J.W. Schmidt, M.L. Brodie


The book is the most comprehensive and detailed analysis of existing relational database management systems to date. It presents a generic characterization of an RDBMS (independently of specific RDBMSs) in terms of:

- relational database constituents
- relational functional capabilities
- definition, generation and administration facilities
- interfaces and DBMS architecture
- operational aspects

These features are then used as a common basis to analyze fourteen of the most important existing RDBMSs. The fourteen systems analyses are then compared, in a tabular format, with respect to the features, and system feature summaries are presented.

The book is introduced by a foreword written by Dr. E. F. Codd, the inventor of the relational approach to databases.

The book is intended to assist the reader in developing a detailed understanding of the Relational Data Model, Relational Database Management Systems, and the state-of-the-art in relational DBMS technology. It provides a comprehensive check list of features with which to evaluate RDBMSs or DBMSs in general.