Chairman's Column ...................................................... 1
Notes From the Vice Chairperson................................. 4
SIGART Notice............................................................ 6
ICOD-2 Notice............................................................ 7
In Search of Warm Bodies............................................. 9
How the DBA Stole Christmas...
   by D. Reiner and A. Rosenthal............................... 11
Adding Database Management to Ada
   by A.V. Hall.......................................................... 13
Comparison of Four Relational Model Specifications
   by W. Lamersdorf.................................................. 18
An Informal Approach to Formal Specifications
   by A. Furtado....................................................... 32
Research Issues in Database Specification
   by M. Brodie........................................................ 42
Interfacing a Query Language to a CODASYL DBMS
   by R. Tagg.......................................................... 46
The Devolution of Functional Analysis
   by K. Meyer and J. Doughty................................. 65
SIGMOD RECORD

A Quarterly Publication of the Association for Computing Machinery (ACM)
Special Interest Group on Management of Data (SIGMOD)

SIGMOD Officers

Chairperson
Michael Stonebraker
M265 - Cory Hall
University of California
Berkeley, CA 94720
(415) 642 - 5799

Vice-Chairperson
Patricia G. Selinger
IBM Corp. - K53/282
5600 Cottle Road
San Jose, CA 95193
(408) 256 - 1295

Secretary-Treasurer
John L. Berg
Comp. & Info. Tech.
Standard Oil
MC: 1801
PO Box 5910-A
200 E. Randolph Drv
Chicago, IL 60680
(312) 836-4896

SIGMOD NEWSLETTER EDITOR

Thomas J. Cook
Cook Data Services, Inc.
6510 Abrams Road, Suite 410
Dallas, TX 75231
(214) 349 - 6900

CONTRIBUTIONS to the SIGMOD RECORD may be sent to the Editor. They should be camera-ready and typed single spaced. All papers printed herein are unrefereed working papers, unless otherwise noted.

MEMBERSHIP IN SIGMOD (at $3 per year) is open to ACM members, associate members and student members; non-ACM members may belong to SIGMOD (at $11 per year) provided their major professional allegiance is in a field other than information processing or computing. Library subscriptions are available for $9 per year.

CHANGES OF ADDRESS, SUBSCRIPTION INQUIRIES and OTHER BUSINESS INQUIRIES should be directed to ACM SIGMOD, 11 West 42nd Street, New York, NY 10036.

REGULAR BACK ISSUES of the SIGMOD RECORD (formerly FDT) are available from ACM Headquarters for $2 per copy prepaid. SIGMOD WORKSHOP PROCEEDINGS are also available for different prices.
Chairman's Column

With a little luck this will be my last chairman's column since a new slate of officers will take office in July (ACM headquarters willing). Hence, it seems like I should try to write something significant, instead of my usual dribble. I feel like I have this space once each quarter to spout off on anything. Usually I think "Oh my gosh -- how am I going to fill up a couple of pages?" It is refreshing for a change to have something that I am just itching to spout off about.

If asked to draw up my list of the outstanding academic non-tenured faculty members doing database research, it would be a "no brainer" decision to put Phil Bernstein at the top of the list. Although I have not done a formal canvas of other folks, I suspect other people would also rank Phil at the top of their lists. Hence, I was extremely upset to learn that Phil was denied tenure at Harvard. I am hard pressed to think of a person who has played "the tenure game" any better or any harder than Phil. He publishes prolifically, turns out excellent students, reliably performs service to the field as a reviewer, journal editor, etc. and generally is a credit to the database field. It appears that such qualifications, sufficient to obtain tenure at most universities, are not worthy of tenure at Harvard. The apparent reason is that Harvard wants to attract a tenured "systems type", and Phil is considered too much of a theorist.

Frankly, I am appalled at this miscarriage.

My other parting shot is to force on you my view of interesting and not-so-interesting research problems.

Questions I would most like to see answered:

1) The AI community views data as a collection of "facts" or "rules". Example facts include:
   John is-a boy
   John likes candy
   boys like girls
   John hates dogs
   John plays a guitar
   guitars have 6 strings
   some guitars have steel strings
   girls like dolls

A semantic net representation (or something equivalent) is often proposed for such information. The first question is to design a relational schema (or a schema in your favorite data model) for the above data. Then, figure out a way to express:

a) tell me everything you know about John
b) does John like girls? ... which requires an inference to answer

2) Suppose you want to write a "portable" DBMS, say one that could, for example, run successfully on OS/MVS, VM/370, UNIX, and CP/M. Suppose your system required the following user services: authorization, system time available, sophisticated accounting, transaction management, etc. To convince you this is non trivial, note that VM/370 has limited interprocess communication, UNIX has no efficient way of forcing a page to disk (to obtain a write ahead log), and CP/M has an address space of 64K or 128K (which is not large enough for a full function system).
The problem is to develop a serious approach (i.e. write the system in FORTRAN doesn't count) to portability.

3) Suppose you have to write a Visicalc clone (a so-called Visiclon) and want it to use data base data. Propose an extended query language or some other support mechanism that would expedite this software module.

Questions I'd least like to see answered:

1) Generate another "efficient" concurrency control algorithm.

2) Propose yet another dependency (PYAD) and explore its properties.

3) Propose yet another distributed data base query processing algorithm based on "perfect" information (i.e. the sizes of all intermediate objects one might create).

Anyway, on to other things.....

The nominating committee persuaded the following slate of candidates to run for SIGMOD offices:

For Chairman:

Dr. Michael Blasgen  
IBM Communications Products Division  
18100 Frederick Pike  
Gaithersburg, Md.  
(301) 428-2438

Dr. Bea Yormark  
Coopers & Lybrand  
1800 M. St. N.W.  
Washington, D.C.  
(202) 822-4000

For Vice Chairman:

Dr. Murray Edelberg  
Sperry Research  
Sudbury, Mass.  
(617) 369-4000

Dr. Jeff Ullman  
Computer Science Dept.  
Stanford Univ  
Stanford, Ca.  
(415) 497-4079

For Secretary/Treasurer

Dr. Randy Katz  
Computer Science Dept.  
University of Wisconsin  
Madison, Wisc.  
(608) 262-0664
I am delighted that the committee conceived such a well rounded slate of excellent people. All have expressed a willingness to pick up the SIGMOD banner and push forward over the next two years. In the meantime the current officers plan to retire to the sidelines as quickly as possible.

Data Base Week/83 promises to be an exciting conference. David Dewitt and his program committee met last month to formally accept 24 papers from among 120 submitted. From the looks of the final program, it will be an great conference. In addition, Raymond Lorie and his committee received over 20 papers for the stream on engineering applications. His committee has chosen a collection of very good papers. Lastly, Eric Carlson has a full day and a half of business applications planned. Although SIGIR could not participate this year because of scheduling conflicts, they have expressed a desire to include an information retrieval stream next year.

Robin Williams, Corky McCord and the organizing committee have worked very hard to put on a good show. Now the next steps are up to you:
1) pick up pen
2) fill out registration form
3) make plane reservations
4) come to San Jose

Official SIGMOD meteorologist Pat Selinger promises clear skies and balmy weather. Moreover, the San Jose Hyatt (site of the conference) is not in Alviso, Ca. Hence, Pat has assured me that there is no danger that it will be under water.
NOTES FROM THE VICE CHAIRPERSON

This is my last column as vice chairperson. I have enjoyed serving the SIGMOD community for the last three years. These years have been a time of evolution for those of us in the database area. Database work three years ago focussed mainly on new "system" algorithms (e.g. new query optimization algorithms, new concurrency techniques). Today, in addition to this system-oriented work, we see a considerable emphasis on using databases in new, non-traditional ways (e.g. storing office or engineering data). Research and development is no longer entirely introspective, aiming at database system architecture, but also now emphasizes an external awareness as well.

I see this same evolution in SIGMOD itself, reflecting this same multiplicity of directions. We have more joint conferences with other SIGs, and the 1983 SIGMOD conference has expanded its scope to a "Database Week" with other sponsors which reflects databases as used in business, office, and engineering applications. In addition, the issues of the SIGMOD Record are now coming out regularly.

Here is a reminder of the various upcoming database and general conferences. For more details, see the ACM Communications.

CALLS FOR PAPERS


CALENDAR OF CONFERENCES


October 5-7, 1983: Third International Symposium on Data Analysis and Informatics, Versailles, France.
October 6-8, 1983: Third International Conference on Entity-Relationship Approach, Disneyland Hotel, Anaheim, Ca.


October 17-23, 1983: Informatica 83: Sixteenth Brazilian National Informatics Congress and International Exhibit, Sao Paulo, Brazil.


October 8-10, 1984: ACM 84, San Francisco, California.

October 14-16, 1985: ACM 85, Denver, Colorado.

Pat Selinger
IBM Research Lab K53/282
5600 Cottle Rd.
San Jose, CA 95193
Special Issue on AI and Database Research

A special issue of the SIGART Newsletter on AI and Database Research is scheduled for the October issue[?]. The purpose of the issue is to provide a survey of current research in artificial intelligence that has particular bearing on database representation and use. Suggested topics include, but are not limited to: database description; deductive question answering; very high level query languages; use of constraints on data values; representation of negative, incomplete, or hypothetical information in databases; natural language retrieval and updates; and the use of intelligent aids for locating and querying data.

Any individual or group that would like to have their activities included is invited to submit a brief summary of their research. Submissions should include a title, names and affiliations of people associated with the research, and recent or selected bibliographical references. Suggested content would be a summary of the main research topics, accomplishments, current focus, and future goals. We would like the survey to be as complete as possible, so please feel free to include information on new or proposed projects.

Summaries must be limited to 3 double-spaced pages, including references and examples. Please be as brief as possible and avoid details of representation and techniques beyond short examples. Long submissions will be subject to editing. Please submit camera-ready copy of any figures. Submissions should be sent to Jonathan J. King at the address below (via Arpanet is preferable) prior to June 1.

Guest Editor, SIGART Special Issue on AI and Database Research:

Jonathan J. King
Symantec
306 Potrero Ave.
Sunnyvale, CA 94086

Arpanet address: KING@SRI-AI
ICOD-2  
SECOND INTERNATIONAL CONFERENCE ON DATABASES  
Churchill College, Cambridge, England  
August 30 - September 3, 1983  

PROGRAMME SUMMARY  
THE MAIN CONFERENCE (AUG 30 - SEPT 2, 1983)  

Invited Papers:  
2. Frank Manola and Alain Pirotte (Computer Corporation of America, USA and Phillips Research Laboratory, Belgium): An Approach to Multi-Model Database Systems  
3. Herbert Weber (Universitat Bremen, FRG): Object-Oriented DDBS: Design  

Contributed Papers:  
L. Bic, R. L. Hartmann, J. Todhunter (University of California, Irvine, USA): The Active Graph Database Model  
C. J. Date (IBM, San Jose, USA): The Outer Join  
A. Motro (University of Southern California, Los Angeles, USA): Interrogating Superviews  
S. Christodoulakis (University of Toronto, Canada): Access Files for Batching Queries in Large Information Systems.  
T. L. Anderson (Burroughs Corporation, USA): Modeling Events and Processes at the Conceptual Level  
E. Wong (University of California, Berkeley, USA): Semantic Enhancement Through Extended Relational Views  
D. Stemple, K. Ramamritham, S. Vinter, T. Sheard (University of Massachusetts, USA): Operating System Support for Abstract Database Types.  
M. A. Melkanoff, Q. Chen (University of California, Los Angeles, USA): An Experimental Database which Combines Static and Dynamic Capabilities  
E. Gelenbe (Universite Paris Sud and Centre Mondial, France): Incomplete Representations of Information in Data Bases  
R. Unland, U. Praedel, G. Schlageter (University of Hagen, F.R.G.): Design Alternatives for Optimistic Concurrency Control Schemes  
W. S. Luk, E. Luk (Simon Fraser University, Canada): Optimizing Semi-Join Programs for Distributed Query Processing
E. Bertino, C. Meghini, G. Pelagatti, C. Thanos (Politecnico di Milano and IEI-CNR Italy): The Update Problem in the Distributed Database System HERMES/I


N. H. Phuc, M. Becker, P. Sevray (CII - Honeywell Bull and Université Paris VI, France): Performance Comparison between B*-Tree and Prefix Binary Tree Index Organizations

Panel Sessions:
1. Database Machines - the last ten years?
2. Personal Databases - the first ten years?
3. Distributed Databases - the next ten years?

THE WORKSHOP ON NEW APPLICATIONS OF DATABASES (SEPT 2 AND 3, 1983)

TOPICS:

<table>
<thead>
<tr>
<th>Topic</th>
<th>Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image Databases</td>
<td>3</td>
</tr>
<tr>
<td>Text Oriented Systems</td>
<td>3</td>
</tr>
<tr>
<td>CAD/CAM Applications</td>
<td>2</td>
</tr>
<tr>
<td>AI in Database Management</td>
<td>2</td>
</tr>
<tr>
<td>High Level Interfaces</td>
<td>3</td>
</tr>
</tbody>
</table>

Open Panel Discussion: Databases - Where next?

ORGANISERS: The British Computer Society, University of Aberdeen, University of Cambridge, Université de Paris-Sud

SPONSERS: [acm] in cooperation with ACM (USA), AFCET (France), AICA (Italy), European Research Office of the US Army, IEEE (USA), GI (FRG), INRIA (France), Shell International Petroleum Company Ltd.

FURTHER INFORMATION:

For further details, please contact:

Mr Peter Hammersley
ICOD-2 Secretary
Middlesex Polytechnic
Bounds Green Road
LONDON N11 2NQ

Tel: 01 - 368 1299
Telex: 8954762
Michael Stonebraker  
Chairman, ACM SIGMOD  
M265 Cory Hall  
University of California  
Berkeley, CA. 94720

Dear Michael:

I would like you to identify a prospective individual from SIGMOD who would be willing to serve as a member of the ACM Standards Committee with responsibility for the data management area. The individual would serve as a communications channel between ACM’s voting representatives (primarily on X3) and ACM members on topics in that area. That would be achieved by publishing information on relevant standards activities in the SIGMOD newsletter or CACM, and by serving as a focal point for input from ACM members on major items that come to a vote in X3. It would be additionally beneficial if the individual were a member of X3H2, if necessary sponsored by SIGMOD.

Attached are a couple of recent news releases on items which might be included in the newsletter. Please let me know if I can provide additional information.

Sincerely,

David C. Wood  
Chairman, Standards Committee

DCW/cpc

attachments

cc: Tom Cook, SIGMOD Newsletter Editor  
Jan Lee, X3 Rep.
October 5, 1982

ANSI DATA BASE GROUP SEEKS MEMBERS

Washington, D. C. -- The ANSI/X3/SPARC Data Base Systems Study Group (DBSSG) is responsible to X3 for planning the development of a family of standards for data base systems. In particular, the Study Group investigates and coordinates projects using an architectural framework for database standards. The group is seeking new members to help it in this challenging endeavor.

The family of standards will be based on work being done by DBSSG's Task Groups: Architecture Framework, User Facilities, and Glossary. DBSSG is also responsible for liaison with various international data base standards organizations. A new project on the relational model was recently initiated through this committee.

Persons who can actively participate in these interesting and important tasks are invited to contact the DBSSG Acting Chairman

John W. Young, Jr.
NCR Corporation
9900 Old Grove Road
San Diego CA 02131-1685
714/268-5358
How the DBA Stole Christmas, or
The Case for Separately Compiled Access Modules

David Reiner and Arnon Rosenthal
Sperry Research Center
December 23, 1982

This sad tale was found among some prematurely-discarded Christmas tree ornaments.

Gloomy Naive User of a relational DBMS (staring unhappily at his terminal): You said I didn’t need to know computerese with this relational database. But the program Joe wrote in 1979 to order booze for the Christmas party just bummed.

Database Administrator (at the next terminal): The word is "bombed". Let's see . . . "ERROR 0307 - INVALID ACCESS ATTEMPTED - INDEX DOES NOT EXIST".

GNU: I don't understand. They didn't teach us about this in our course last Tuesday.

DBA: Oh yeah, I had to remove some indices - they slowed update transactions too much. And of course the queries in your program won't work now that the underlying access paths have been modified.

GNU (muttering): Access paths?

DBA: Don't worry, just recompile and relink the program and it'll run fine.

GNU: Recompile, relink? What are those?

DBA: I guess I'll have to do it for you. Where's your source program stored?

GNU (blank look. Yells): Joe!

JOE (walking over from the nearby graphics terminal): Problems with the booze program? It's in the file ETHYL*WHOOPEE, but I encrypted the text so management wouldn't see it. I've lost the encryption key, so the source is gone. Besides, it was in FORTRAN V and we're using the ASCII FORTRAN compiler now. It wouldn't like all my calls to the FLD function. Isn't there a way to recompile just the database accesses?

DBA: Nope. I think you'll have to rewrite the whole booze program.

JOE: Just because some indices got changed? Some dynamically modifiable system this turned out to be! How many of my other programs are affected? Do I have to recompile everything?
DBA (typing in a command at his terminal): According to the data dictionary, 19 of your programs include queries that reference the relations which lost indices.

JOE (looking incredulous): I'd have to set up a weekend run to do them all! But some of them can't possibly be affected by index deletions. They just scan whole relations sequentially.

DBA: Well, the data dictionary doesn't have query-level resolution to determine the effect of a change. It only knows which programs reference which relations. You can recompile and relink everything to be safe, or just run the programs and see if they bomb. Some might still run, but be twenty times slower, because the clustering has changed since your queries were optimized.

As we take our leave, DBA is consoling GNU, who has begun to cry. JOE is gazing fixedly at his terminal, clenching and unclenching his fists. But this tragedy could have been avoided.

If the DBMS is given control over compiled access modules, including their possible recompilation, then query-level resolution is possible to determine the effect of changes. System R did this, so it's not a new idea. The object code corresponding to a given query is stored in a central area, and its validity can be checked before program execution to see if changes are necessary. The host program need not be recompiled or relinked at all, just the access module. And this can be done automatically, so the user never has any unpleasant surprises. In addition, the user's queries can be reoptimized to take advantage of new access paths, instead of being bound to the state of the database which existed at host program compilation time. Of course, this approach has to be designed in from the beginning.