

## IN DEPTH

# Does your DBMS run by the rules?

---

*To be "mid-80s" fully relational, a DBMS must support all 12 basic rules plus nine structural, 18 manipulative and all three integrity rules. There will be more requirements by the 1990s.*

---

By E. F. Codd

*Last week, the originator of the relational model described the 12 rules by which to measure any DBMS claiming to be relational. This week, Dr. E. F. Codd presents the practical consequences of his 12 rules as well as 30 additional features of a relational system. Then he asks vendors to measure up.*

## Part 2

No existing DBMS product that I know of can honestly claim to be fully relational at this time. The proposed ANSI standard does not fully comply with the relational model, so a DBMS' fidelity to the ANSI standard is no guarantee of relational capability. The standard could be modified, but already vendors are well advised to extend their products beyond the standard to support customers' DBMS needs fully.

In their ads and manuals, vendors have translated the term "minimally relational" to "fully relational," so more stringent criteria must be applied. Twelve rules (below) comprise a test to determine whether a product that is claimed to be fully relational is actually so. A grading scheme used to measure the degree of fidelity to the relational model follows.

A DBMS advertised as relational should comply with the following 12 rules:

1. The information rule.
2. The guaranteed access rule.
3. Systematic treatment of null values.
4. Active on-line catalog based on the relational model.
5. The comprehensive data sublanguage rule.
6. The view updating rule.
7. High-level insert, update and delete.
8. Physical data independence.
9. Logical data independence.
10. Integrity independence.
11. Distribution independence.
12. The nonsubversion rule.

---

E. F. Codd originated the relational model for data base management. He led the team that designed and implemented the first operating system with multiprogramming capability. This year he established two companies with Chris Date: The Relational Institute and the Codd & Date Consulting Group, both based in San Jose, Calif.

## IN DEPTH/RELATIONAL DBMS

These rules are based on a single foundation rule. I call it Rule Zero:

*For any system that is advertised as, or claimed to be, a relational data base management system, that system must be able to manage data bases entirely through its relational capabilities.*

This rule must hold whether or not the system supports any nonrelational capabilities of managing data. Any DBMS that does not satisfy this Rule Zero is

not worth rating as a relational DBMS.

But compliance with Rule Zero is not enough. Failure to support the information rule, guaranteed access rule, systematic nulls rule and catalog rule can make integrity impossible to maintain. These four rules support significantly higher standards for data base administration and control (authorization and integrity control) than earlier DBMS supported. Users should remember that a

data base managed by a relational DBMS is likely to have both experienced and inexperienced users; it must be able to serve both.

#### Rule Zero not enough

Rules 1 and 4, the information and catalog rules, allow people with appropriate authorization (such as executives of the company) to find out easily via terminal what information is stored in a data base. I have encountered data base administra-

tors using nonrelational systems who were unable to determine if a specific kind of information was recorded in their data base.

Rule 3, which calls for the inclusion of systematic support for unknown and inapplicable information by means of null values that are independent of data type, should help users to avoid foolish and possibly costly mistakes. The treatment of nulls, when aggregate functions such as total and aver-

age are applied, holds considerable interest for users. The Oracle DBMS in particular has an outstanding approach to null values. The user may specify whether the aggregate function is to ignore null values or yield a null result if any null value is encountered.

In general, controversy still surrounds the problem of missing and inapplicable information in data bases. It seems to me that those who complain loudly about the complexities of manipulating nulls are overlooking the fact that handling missing and inapplicable information is inherently complicated. Going back to programmer-specified default values does not solve the problem.

Rule 5, the comprehensive

"

*The ANSI standard as now proposed is quite weak. It fails to support numerous features users need to reap the advantages of the relational approach.*

# You may ask why Elgar's new UPS looks just like an IBM System/36? ...Why indeed!

- Our new UPS is indeed a perfect match for IBM. It provides the kind of power protection necessary for IBM systems. And only Elgar is plug-in compatible with both the CPU and peripherals for fast, easy installation.
- Our new UPS is packed with quality features for trouble-free performance. Only Elgar has  $\pm 4\%$  dynamic power regulation. We deliver more precise power to your computer than any other UPS you can buy. Our new UPS costs less to install . . . to run . . . and maintain.
- Want more? It fits in your office. It's so quiet you'll hardly know it's running. And it's backed by a no-nonsense TWO YEAR WARRANTY from Elgar . . . the leader in quality power protection systems.
- Get the full story on our new T-Series UPS for mini computers.

Call toll free: 800-854-2213, Dept. 8.



**FREE BOOKLET OFFER**  
A reader for what every manager needs to know about power and company profits. Call today for your free copy.

# ELGAR

An Onan Operating Unit

9250 Brown Deer Road  
San Diego, California 92121  
(619) 450-0085 Telex 211063

IBM is a registered trademark of The International Business Machines Corporation.


data sublanguage rule, is important for several reasons. First, it allows programmers to debug their data base statements interactively, treating them separately from whatever nondata base statements occur in their programs — a significant contributor to productivity. Second, it means that a single tool can be used for defining relations derived from the data base, whatever the purpose. The view updating rule, Rule 6, is vital for the system to support logical data independence.

Rule 7, which requires a multiple-record-at-a-time attack on insertion, update and deletion, can help save a good portion of the total cost of intersite communication in a distributed data base. If the system includes a good optimizer (an important component in relational DBMS performance), this rule can also result in substantial saving of CPU and I/O time, whether the data base is distributed or not.

Failure to support independence (Rules 8 through 11) can, and very likely will, result in skyrocketing costs in both money and time. Developing and maintaining applications programs and terminal activities will be more expensive. Managers may even be unwilling to consider changing certain business policies simply because of the anticipated program maintenance costs.

Rule 12, the nonsubversion rule, is crucial in protecting the integrity of

relational data bases. All too frequently, I have seen situations in which data base administrators with nonrelational DBMS failed to control their data bases adequately; consequently, they could not maintain a state of integrity.

#### Domains

Many users confuse the domain concept with the concept of attribute of a relation or column of a table. Other people (often the vendors themselves) dismiss the domain concept as "academic." My reply to them is: The atom bomb was also academic!

In fact, the domain concept is very important, practical and simple. A domain consists of the whole set of legal values that can occur in a column. The column draws its values from the domain. Each column of a relational data base has precisely one domain, but any number of columns may share a domain. There are several reasons why domains should be supported.

For example, in a financial data base, there may be as many as 50 distinct columns (possibly, but not necessarily, in distinct tables) defined around the U.S. currency domain. Why repeat the definition of currency 50 times? In data bases supported by nonrelational systems, I frequently observe many inconsistent declarations of value type for fields that were intended to have the same type.

It is unreasonable to expect a DBMS to store all the legal values in a domain, unless there happen to be very few. However, it is entirely reasonable — and very worthwhile — to insist that a DBMS should store certain values:

- For each domain, a description of the type of values in that domain. This information is global since it applies to the entire database, and it should of course be recorded in the catalog.
- For each column, the name of the domain from which that column draws its values. This domain name is a reference to the global definition.

Of course, the domain description can include range restrictions. For example, it could specify that quantities of parts in an inventory must not only be integers, they must also be non-negative.

Furthermore, individual columns may include additional range restrictions where these are semantically justifiable. In this example, the quantities of very expensive parts held in the inventory may be limited to some specified maximum.

One of the benefits of supporting the domain concept is that, in cases where several columns share a common domain, the declaration of the description of the legal

values is largely or even completely factored out. For example, when there are 50 distinct columns defined on U.S. currency, the data base is much easier to manage and manipulate if one avoids making 50 distinct declarations for U.S. currency.

Before the relational discipline arrived, users had to make separate declarations, and as a result, many of the 50 in the example would turn out to be incompatible with one another by accident.

The factoring of declaration that prevents these errors is achieved in Digital Equipment Corp.'s RDB, which has a concept of "global field definition." But RDB fails to support domain constraints on certain operations, such as join.

Another benefit of supporting the domain concept is that relational operators, such as joins and divides, that involve comparison of values between different columns can be constrained by

the system. A DBMS can allow data base values to be compared only when they come from the same domain and are therefore comparable from the semantic viewpoint.

Such a constraint inhibits errors caused by interactive users of terminals who choose columns to be compared in such operations as joins. The wrong answers they obtain from these errors rarely uncover the errors themselves; meanwhile,

unwise business decisions may be made based on these wrong answers.

For various reasons, it is important to support as a qualifier in a command what I call "semantic override" — the ability to have the system ignore the usual comparison constraints. Users should be able to authorize this override qualifier separately from the operator involved and should authorize it rarely, reserving it chiefly for detective work.

## TO INSURE THAT EVERYONE CAN ACCESS DATA, TRANSAMERICA USES INTELLECT AT EVERY LEVEL OF THE PYRAMID.

**"We use INTELLECT because we want to give our users a better way to do business."**

**—Mr. Carl Rahmqvist, Senior Systems Manager, Information Systems, Transamerica Insurance Group**

Transamerica Insurance Group is recognized for the pyramid-shaped building of its parent, Transamerica Corporation, and for being one of the country's leading commercial and personal insurance companies. In some quarters, Transamerica is also recognized for its variety of innovative INTELLECT applications.



At Transamerica they see INTELLECT as more than a state-of-the-art natural language information retrieval system. They see it as a new way to do business: Giving all their end users—even those with no computer skills—instant access to more information than they'd ever had before.

**"When they saw how easy, fast, and resource-effective it was to use, INTELLECT became a very popular tool."**

**—Ms. Sandra Dahlgren, Information Center System Supervisor**

One Vice President uses INTELLECT to get the most current information on premiums and losses, and for longer term strategic planning with requests such as, "Give me the total June premiums and losses for each region." Regional offices use INTELLECT for a variety of tasks including asking INTELLECT to: "Tell me all about policy number 98579897." Personnel keeps track of employee records, EEO compliance, human resource utilization, and more, by questioning INTELLECT in plain English. Payroll, Claims, Underwriting, and Services also use a variety of INTELLECT applications. So from the top of the pyramid right down to the mailroom, Transamerica is using INTELLECT to work faster and smarter.

**"After just minimal training, people began using INTELLECT frequently and effectively."**

**—Mr. Carl Rahmqvist**

After some initial implementation assistance from AIC,

Transamerica's Information Center took over, developing custom INTELLECT applications for each department. Mr. Rahmqvist and Ms. Dahlgren trained a group of Transamerica's "veteran" INTELLECT users to go into the field to train the company's 15 regional office personnel in using the company's many applications. Transamerica's INTELLECT Support Staff helps company personnel use INTELLECT more effectively. They've also devised several methods for enabling users to access their many INTELLECT applications. Easily understood menu screens help beginning users work faster, while experienced people can use an express mode for their application needs.

**"We've found that we've saved both time and money by using INTELLECT to obtain pertinent information."**

**—Ms. Sandra Dahlgren**

Anyone who can ask a question in everyday conversational English can get the information he or she needs. Immediately. You can imagine how much time a system like that can save an information-dependent organization like Transamerica. Combine that with the ability to get more people into the system and you can see how INTELLECT has increased the pyramid's power.

Find out how INTELLECT can get your organization into better shape. Write for a free demo diskette. Or for fast action, call AIC at (617) 890-8400.

I want to know more about INTELLECT!

Name \_\_\_\_\_

Title \_\_\_\_\_

Company \_\_\_\_\_

Address \_\_\_\_\_

State \_\_\_\_\_ Zip \_\_\_\_\_

Telephone (\_\_\_\_) \_\_\_\_\_

O/S: VM/CMS MVS DBMS \_\_\_\_\_

Please send me information about INTELLECT.

Please have an AIC representative contact me.

CW1021

**INTELLECT<sup>TM</sup> ARTIFICIAL INTELLIGENCE CORP.**  
100 Fifth Avenue, Waltham, MA 02254 617-890-8400

INTELLECT<sup>TM</sup> is a trademark of Artificial Intelligence Corporation.

## IN DEPTH/RELATIONAL DBMS

Even when the domain concept is restricted to assigning types to data, it should not be confused with the hardware-supported data type. Consider the example of a data base listing suppliers, parts and projects. Suppose the hardware-supported data types of supplier serial numbers and part serial numbers are identical: each type consists of fixed-length strings of 12 characters. The system still needs to keep these two data types distinct and remember which columns are defined on one and which columns are defined on the other.

If it can make these distinctions, then when a request comes in to delete or archive all records containing X3 as a supplier serial number, the system can handle such a transaction correctly. The system will not delete or archive any record that contains X3 as a part serial number and that also does not contain X3 as a supplier serial number.

Today, such a data type is often called an application data type. The concept is supported in Pascal but in very few other languages that enjoy current use. The Pascal support does not, of course, include constraints on selects, unions, joins and divides.

The domain concept is basically what makes all the meaningful selects, unions, joins and divides known to the DBMS. Thus, the domain makes the data base meaningfully integrated, and it does so without prejudicing distributability.

Contrast this with CODASYL links and IMS hierarchic links. They represent the CODASYL and IMS con-

Fidelity to the 12 rules (by DBMS)			
Rule	DB2	IDMS/R	Datacom/DB
1 Information rule	Yes	No	No
2 Guaranteed access rule	Partial	No	No
3 Systematic treatment of nulls	Partial	No	No
4 Active catalog based on resource management	Yes	No	No
5 Comprehensive data sublanguage	Yes	No	No
6 View-updating rule	No	No	No
7 High-level insert, update, delete	Yes	No	No
8 Physical data independence	Yes	Partial	Partial
9 Logical data independence	Partial	No	No
10 Integrity Independence	No	No	No
11 Distribution Independence	Yes	No	No
12 Nonsubversion rule	Yes	No	No
Score (1 for yes, 0 otherwise)	7	0	0

Figure 1

cept that a link "integrates an otherwise unintegrated data base," but they have several unfortunate restrictions. Most importantly, they obstruct data base distribution because of the constraints and complexity their data structures introduce into decisions regarding how the data should be deployed.

A second serious drawback of links is that they are only paths. Generation of a result such as a join requires traversal of these paths by the application program. It seems superfluous to cite other difficulties with this concept.

Many relational DBMS and languages including SQL do not support

the concepts of primary key and foreign key. I fail to see how these products can support the guaranteed access or the view updating rules without making the system aware of which column(s) constitute the primary key of each base table.

Furthermore, I fail to see how these products can support referential integrity or the view updating rule without offering clear support for both primary keys and foreign keys. For example, in SQL, the CREATE TABLE command should be extended to permit the user to declare which column or columns constitute the primary key and which constitute foreign keys. In addition, there

should be a new CREATE DOMAIN command in SQL.

## Fidelity

Figure 1 shows fidelity to the 12 rules by IBM's DB2, Cullinet Software, Inc.'s IDMS/R and Applied Data Research, Inc.'s Datacom/DB — examples chosen for their wide differences. These scores represent counts of compliance with each rule (score one for "yes" and zero for either "partial" or "no").

Actually, the information rule is so fundamental to the relational approach that a system's compliance with this rule should receive a much higher score than one. Weighting it as high as 10 would not be excessive. However, I shall avoid assigning different points for different features, just as I avoided a fractional score for partial support of a feature: It is too easy to be subjective in these matters.

DB2 scores quite well on the fidelity evaluation. Very few other DBMS score higher on the 12 rules, although some others score equally well. Both IDMS/R and Datacom/DB allow information to be represented in the order of records in storage and in repeating groups — directly violating the information rule. In the case of IDMS/R, information may also be represented in links between record types (CODASYL calls them "owner-member sets") and also in "areas."

Some vendors of nonrelational DBMS have quickly added a few relational features — in some cases, very few features — in order to be

## TSO users DSM™(Data Set Manager) Improves TSO Performance

```
-----  
Data Set Manager  
-----  
CATALOG PREFIX: JONES  
PFN: Global Command Input Area  
-----  
--DSNAME--- VOLSER ALLOC USED EX QRG RECFM BLKSIZE LRECL LASTREF  
SPF AN2.CNTL TS0010 3 2 1 PS FB 3120 80 08/23/84  
B APR01.TEXT TS0012 8 5 1 PS FB 3120 80 07/08/84  
FSE B13.COBOL TS0008 10 8 1 PS FB 3120 80 08/13/84  
DEL CNTL.TEXT TS0005 6 6 1 PS VB 3120 255 08/23/84  
REN DENSFORT TS0006 4 4 3 1 PS FB 3120 80 08/20/84  
COPY FZASHA TS0002 3 3 1 PS FB 3120 80 10/02/84  
SUB FM001.CNTL TS0006 1 1 1 PS FB 3120 80 09/12/84  
PRINT G202.LIST TS0003 20 15 1 PS VB 3120 255 09/20/84  
COMP HOL.TEXT TS0008 30 30 3 PO VB 3120 255 10/15/84  
RLSE IM001.LOAD TS0010 50 15 1 PO U 0 19008 08/03/84  
LM JCL.CNTL TS0002 20 15 1 PO FB 3120 80 09/15/84  
FREE KAC.B001 TS0004 6 5 1 PS FB 3120 80 08/20/84  
---- TS0003 10 8 1 PS VB 3120 255 10/23/84
```

- DSM provides full-screen displays of cataloged files and supports TSO functions with simple, easy-to-remember command names.
- DSM supports SPF, FSE, CLISTS and Command Processors. DSM's flexible design allows users to define their own commands.
- DSM file commands include: DELETE, RENAME, COPY, SUBMIT, PRINT, COMPRESS, Release unused space, List PDS members and deallocate.
- DSM supports PDS member displays, HSM and MSS.
- Security? Our user exit lets you control all functions.

Call or Write for details:

Applied Software, Inc.  
840 U.S. Highway 1, Suite 250,  
P.O. Box 13027, North Palm Beach, FL 33408  
(305) 626-4818

## Need to recruit people for your IBM systems? You'll find them reading the Computerworld Extra! on IBM.

Published December 4th and closing October 25th, this special edition of Computerworld Extra will take a hard look at IBM's products and strategies. Anyone working in the IBM arena will certainly review this issue. So if you're looking for pros in IBM systems, get your ad in this special issue and be surrounded by in-depth editorial on IBM.

Computerworld Extra! will discuss IBM's strengths and weaknesses. We'll look at how SNA evolved, and how it will continue to evolve. And we'll discover whether IBM plans to provide a universal interconnect to SNA. Finally, we'll cover the alternatives, from PCs to mainframes, LANs, Communications. And, of course, the compatibility issue.

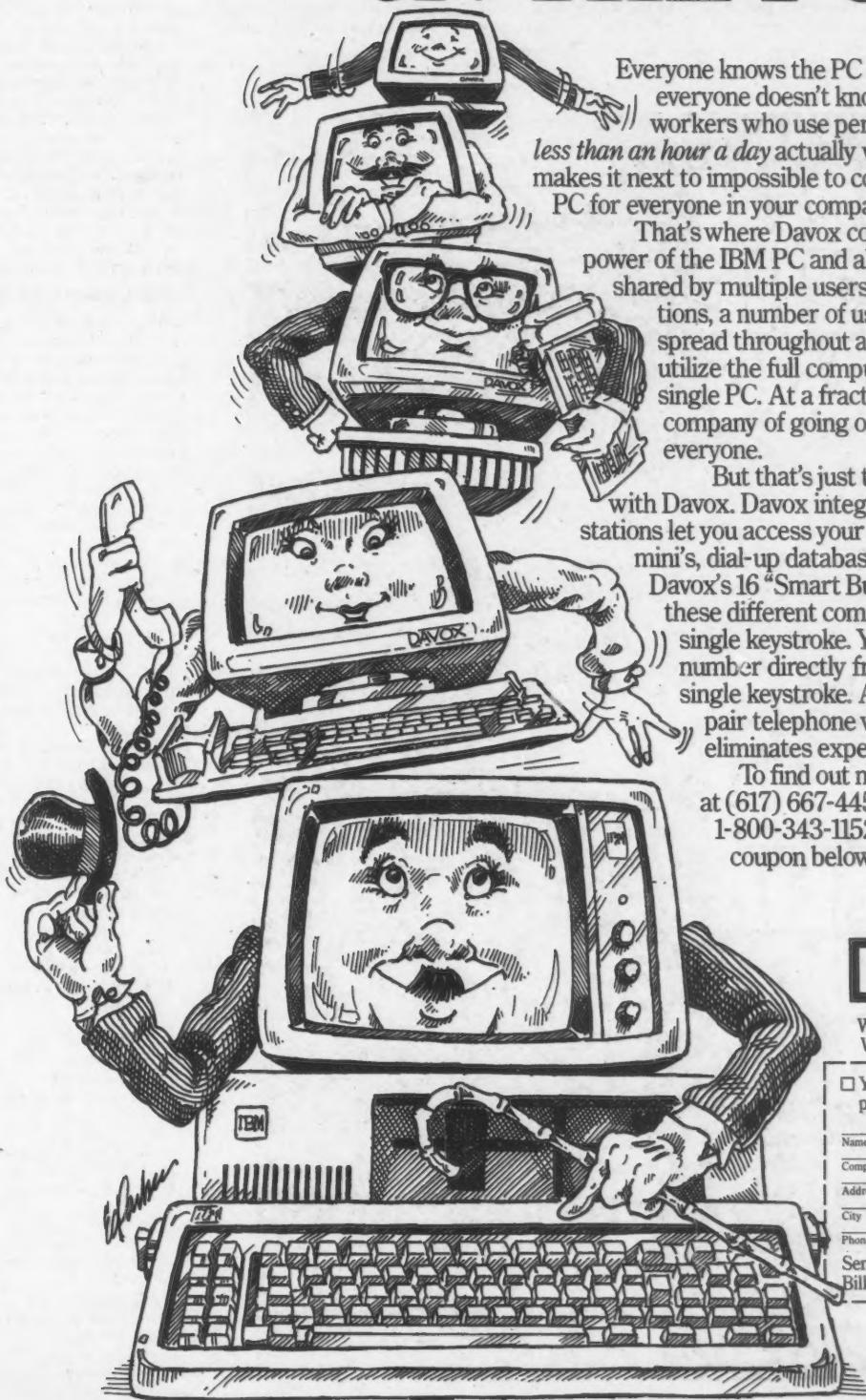
As you can see, this issue will have complete appeal for computer professionals working at IBM installed sites.

To reserve space call Al DeMille, National Recruitment Sales Manager, at (800) 343-6474, or (617) 879-0700 in Massachusetts.

There is no special classified section; all recruitment ads are considered display advertising for this issue.

**COMPUTERWORLD**  
E X T R A

# THE WORKSTATION THAT PIGGYBACKS ON THE PC.



Everyone knows the PC is a powerful tool. What everyone doesn't know is that many office workers who use personal computers spend *less than an hour a day* actually working on them. Which makes it next to impossible to cost-justify an individual PC for everyone in your company.

That's where Davox comes in. Davox takes the power of the IBM PC and allows that power to be shared by multiple users. With Davox workstations, a number of users (typically 4 or 5) spread throughout a network can access and utilize the full computing capabilities of a single PC. At a fraction of the cost to your company of going out and buying PCs for everyone.

But that's just the first reason for going with Davox. Davox integrated voice/data workstations let you access your IBM mainframe, DEC mini's, dial-up databases or IBM PC's. And Davox's 16 "Smart Buttons" let you access these different computer resources with a single keystroke. You can even dial a phone number directly from the screen with a single keystroke. And DavoxNet's twisted-pair telephone wire networking design eliminates expensive coaxial wiring.

To find out more, call Mike Giltner at (617) 667-4455 or toll-free at 1-800-343-1152. Or mail in the coupon below.

## DAVOX™

We *are* data and voice systems.  
We do it all. And we do it now.

Yes! I'd like to find out how to piggyback on a PC.

Name \_\_\_\_\_  
Company \_\_\_\_\_  
Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip \_\_\_\_\_

Phone \_\_\_\_\_

Send to: Davox, 4 Federal Street,  
Billerica, MA 01821.

10218SCW

## IN DEPTH/RELATIONAL DBMS

able to claim their systems are relational, even though they may not meet simple requirements for being rated minimally relational.

These "born-again" systems keep their lower level languages (single record-at-a-time) open to users either to support compatibility with previously developed application programs or because the vendor takes the position that relational operators are applicable to query only.

In view of this, such systems fail to support the non-subversion rule — a heavy penalty to pay for compatibility. IDMS/R and Datacom/DB are both born-again systems, and both fail to support the non-subversion rule for integrity.

## Features of the model

For a more detailed evaluation of DBMS, users can compare a system to the nine structural features, the 18 manipulative features and

the three integrity features of the relational model. Each feature is defensible on practical as well as theoretical grounds.

The nine structural features are as follows:

S.1 Relations of assorted degrees — or equivalently tables with unnumbered rows, named columns, no positional concepts and no repeating groups.

S.2 Base tables representing the stored data.

S.3 Query tables — the

result of any query is another table, which may be saved and later operated upon.

S.4 View tables — virtual tables that are represented internally by one or more relational commands, not by stored data. The defining commands are executed to the extent necessary when the view is invoked.

S.5 Snapshot tables — tables that are evaluated and stored in the data base, together with an entry in the catalog specifying the date

and time of their creation plus a description.

S.6 Attributes — each column of each relational table is an attribute.

S.7 Domain — the set of values from which one or more columns obtain their values.

S.8 Primary key — each base table has one or more columns whose values identify each row of that table uniquely. The primary key provides the unique associative addressing property of the relational model that is implementation, software and hardware independent.

S.9 Foreign key — any column in the data base that is on the same domain as the primary key of some base relation. The foreign key serves as an important part of the support for referential integrity without introducing links into the programmer's or user's perspective.

## Manipulative features

It is important to keep in mind that the relational model does not dictate the syntax of any DBMS language. Instead, it specifies the manipulative capability (that is, power) that a relational language should possess. At the same time, the model does not require the user to request the data base administrator to set up any special access paths, nor does it require the user to resort to iterative looping or recursion or Cartesian products.

The model also does not require the system to generate a Cartesian product as an intermediate result. In early papers this manipulative capability was expressed in two ways: algebraic and logic-based. The two ways were then shown to be of equal power.

This article uses the algebraic method of expressing the manipulative power, for explicative reasons.

The manipulative features are as follows:

- M.1 theta select
- M.2 project
- M.3 theta join
- M.4 outer theta join
- M.5 divide
- M.6 union
- M.7 intersection
- M.8 set difference
- M.9 outer union
- M.10 relational assignment

- M.11 theta select maybe
- M.12 theta join maybe
- M.13 outer theta join maybe

- M.14 divide maybe

- M.15 theta select semantic override (s/o)
- M.16 theta join s/o
- M.17 outer theta join s/o
- M.18 divide s/o

In the list above, "theta" stands for any one of the comparators: equal, not equal, greater than, less than, greater than or equal to, less than or equal to.

# TORCH THE BACKLOG

## WITH REALIA™ COBOL ON A PC

The fastest micro COBOL.  
Now, the fastest SORT.

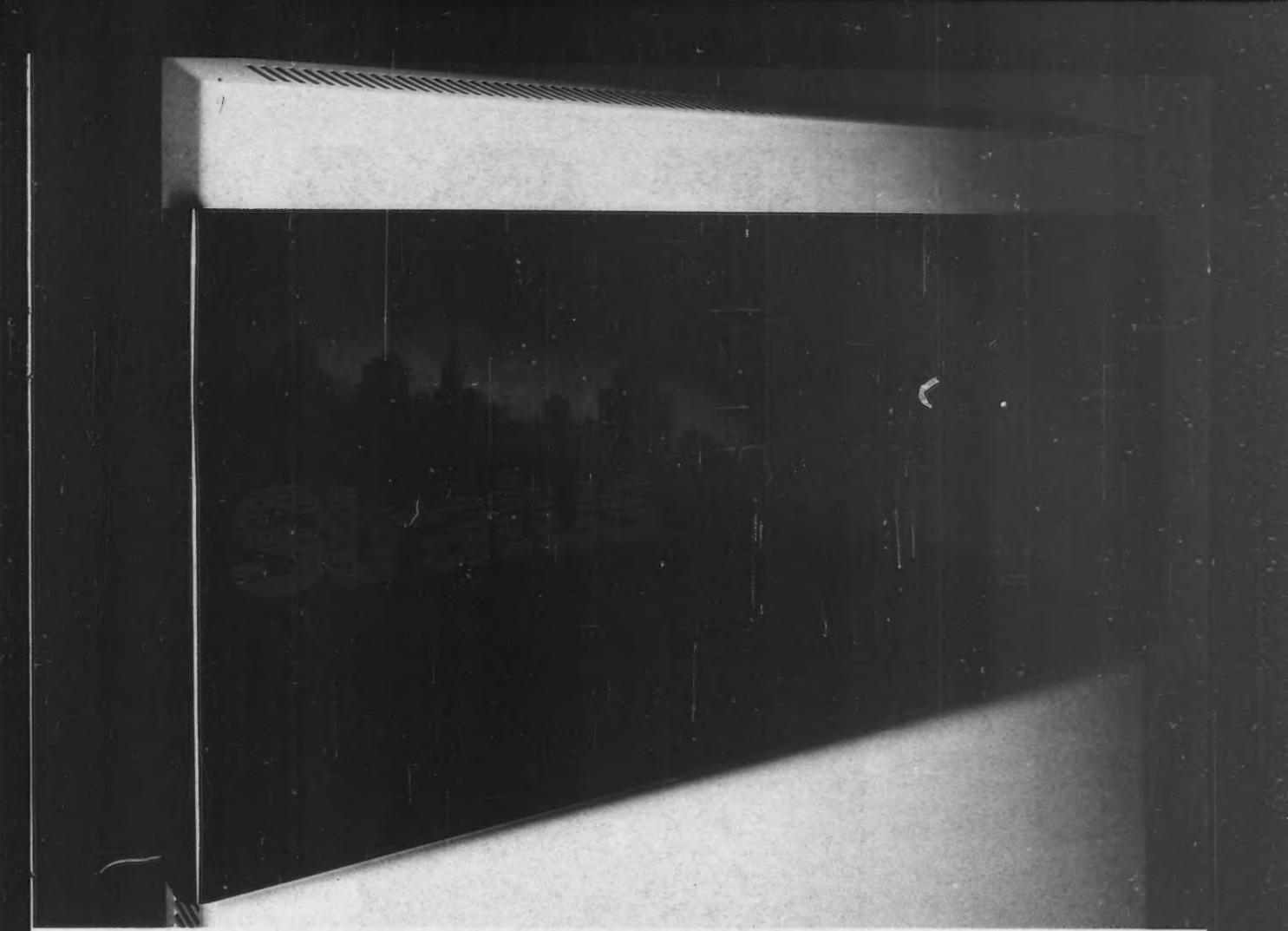
IBM mainframe COBOL compatibility.

Superb support.

Try it for free, if you qualify.

10 South Riverside Plaza  
Chicago, IL 60606  
Phone: 312-346-0642  
Telex: 332979

**REALIA**  
**inc.**



## **WHY THE #1 COMPANIES HAVE CHOSEN THE #2 FAULT TOLERANT COMPUTER COMPANY.**

America's largest car makers, banks, brokerages, and some of the U.S. government's key installations have chosen to depend on Stratus Computer, Inc. for their newest, most critical on-line applications. Even the world's largest computer companies have put their names on Stratus fault tolerant computers.

Why was Stratus picked over its larger fault tolerant competitor? It is because Stratus has quickly proven that its hardware-based fault tolerance is the superior approach to reliability. It is because Stratus computers have demonstrated outstanding throughput and expandability for large-scale networks. It is because of Stratus' ease of programmability, its range of software, its innovative service technology, its customer education.

It is for many reasons that all center around one word—reliability. Because after all, when you want to be known for reliability, you have to excel in everything you do.

As we move into an age of greater and greater computer dependency, many of the giants of business and industry are lining up with the one they have decided is better, Stratus. Apparently they prefer to do their thinking for themselves. Perhaps that's how you get to be a giant. For information call Peter Kastner in Massachusetts at (617) 460-2192, or toll-free at 1-800-752-4826. For a copy of our annual report, write: Marketing Communications, Stratus Computer, Inc., 55 Fairbanks Blvd., Marlboro, MA 01752.

**Stratus**  
CONTINUOUS PROCESSING™

**Now that the world relies on computers it needs a computer it can rely on.**

# THE SAS®

## Fourth Generation Software

**Now there's one software solution for all your Information Center needs. One solution for all your applications, for all your mainframes, minicomputers, and microcomputers. One solution—the SAS® System.**

### One Solution to Integrate All Your Computing Tasks.

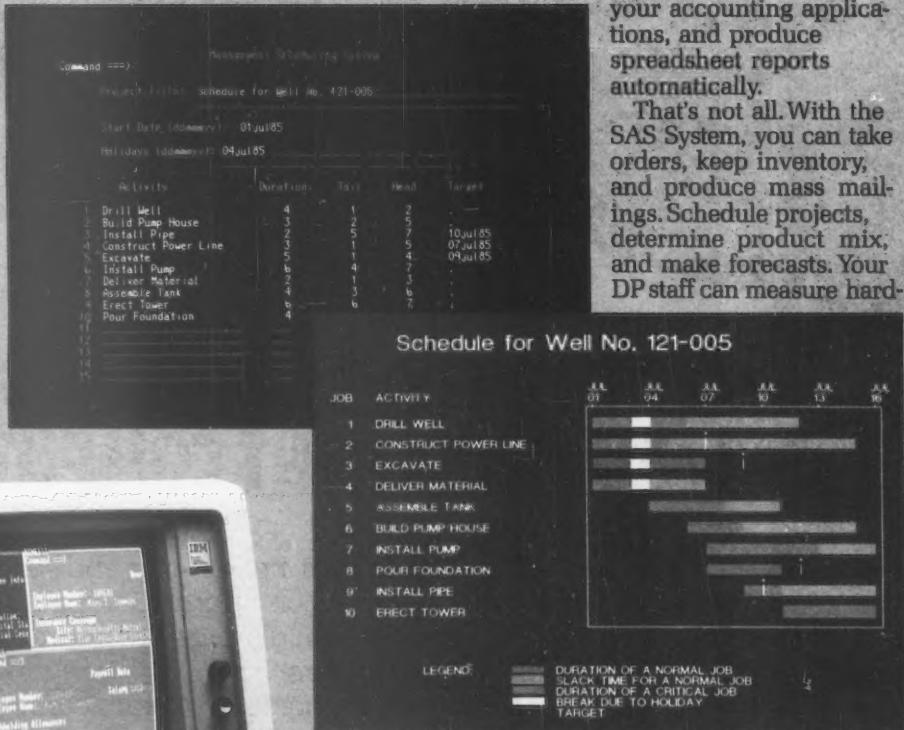
The SAS System gives you efficient data management, superior statistical tools, an easy report generator, customized presentation graphics, and more. Choose between the simple English-like command language or a front-end menuing system with

fill-in-the-blank screens. On-line help facilities make it easy to handle every application, quickly and accurately.

You can track sales leads, manage prospect files, determine market

share, and present results with the SAS System. Plus you can file employee and applicant records, analyze benefit programs, and manage the payroll. The SAS System can handle all your accounting applications, and produce spreadsheet reports automatically.

That's not all. With the SAS System, you can take orders, keep inventory, and produce mass mailings. Schedule projects, determine product mix, and make forecasts. Your DP staff can measure hard-



# SYSTEM

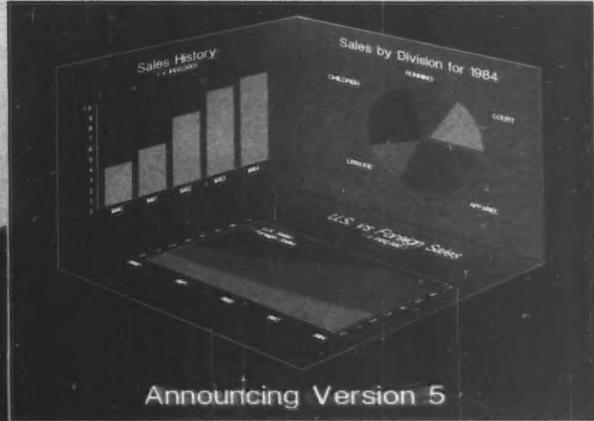
## for Your Information Center.

ware resources or system usage, test data bases, and run production programs.

### One Solution that's Friendly.

It's simple with the SAS System. You can write front-ends for all your SAS applications. With just a few keystrokes, you can modify the applications as your information needs change. One language handles all your tasks. And if you need to move between several operating systems, you'll find the language, syntax, and commands the same for the mainframe, minicomputer, and PC SAS System.

computer-based training. Technical support is provided for our mainframe, minicomputer, and microcomputer users, and documentation comes with your system.



Call us today.  
International customers,  
call the International  
Marketing Department  
for information on your  
local distributor.

Whatever your  
application, the  
**SAS System**  
is your solution.

SAS Institute Inc.  
SAS Circle, Box 8000  
Cary, North Carolina  
27511-8000, USA.

(919) 467-8000, x280  
Telex 802505



The SAS System runs on IBM 370/30xx/43xx and compatible machines under OS, TSO, CMS, DOS/VSE, SSX, and ICCF; on Digital Equipment Corp. VAX™ 8600 and 11/7xx series under VMS™; on Prime Computer, Inc. Prime 50 series under PRIMOS®; on Data General Corp. ECLIPSE® MV series under AOS/VS; on IBM XT/370 and AT/370 under VM/PC; and on IBM PC XT and PC AT under PC DOS. Not all products are available on all operating systems.

SAS is the registered trademark of SAS Institute Inc., Cary, NC, USA. VAX and VMS are trademarks of Digital Equipment Corp., Maynard, MA. PRIMOS is the registered trademark of Prime Computer, Inc., Natick, MA. ECLIPSE is the registered trademark of Data General Corp., Westboro, MA.

Copyright © 1985 by SAS Institute Inc. Printed in the USA.

The integrity features of the relational model must also be followed closely:

- I.1 Entity integrity.
- I.2 Referential integrity.
- I.3 User-defined integrity.

The integrity features cited in I.3 are part of the comprehensive data sublanguage. They support the trigger and assertion approach to defining those integrity constraints that are specific to the particular database. By contrast, I.1 and I.2 apply to all relational data bases. Examples of these extensions have been published, although not fully implemented, for both SQL and QBE.

#### A simple rating technique

To be mid-80s fully relational, a DBMS must fully support all 12 of the basic rules, as well as all nine structural, all 18 manipulative and

all three integrity rules of the relational model — a total of 42 features. I use the term "mid-80s" because it is likely that there will be a few more requirements by the nineties.

To provide a simple method of rating any DBMS on its fidelity to the relational model, treat each rule or feature fully supported by that DBMS as contributing one to the overall score (otherwise the contribution is zero). Then double the total score to obtain a percentage fidelity

rating for the system.

If a DBMS were to achieve a total score of 42 out of 42 (and I believe no such DBMS presently exists), add 8 points to that score before doubling it — as a reward for true fidelity. Thus its fidelity percentage would be calculated to be 100.

The resulting fidelity percentage is not highly accurate. In fact, if it falls between 10% and 90%, I would recommend rounding it to the nearest multiple of 10% in order to avoid misrepresenting the accuracy by dis-

playing more than one significant digit.

#### Evaluation against the model

By today's standards, 46% is a good, but improvable, fidelity percentage. Figure 2 shows the systems DB2, IDMS/R and Datacom/DB evaluated against all 30 features of the relational model. Often the 12 rules are by themselves adequate for comparison purposes. But this more detailed evaluation of the three systems primarily serves expository purposes.

Sometimes users say of a DBMS: "Why should I worry about the degree of its fidelity to the relational model? Surely it is enough for me to know about its fidelity to the ANSI SQL standard."

Unfortunately, the ANSI standard as now proposed is quite weak. It fails to support numerous features that users really need if they are to reap all the advantages of the relational approach.

ANSI's proposed standard for relational systems functions like a convoy, which can proceed only as fast as the slowest ship. The standard is based heavily on that portion of SQL supported by several vendors.

Listing the major differences between ANSI's SQL and, as an example, the SQL implemented in IBM's DB2 shows that ANSI's SQL is even less faithful to the relational model than the vendor's SQL:

- The draft ANSI SQL does not specify catalog tables and does not allow CREATE or GRANT statements to be included in application programs. Instead, it requires a "schema" that specifies an authorization ID and a list of definitions of tables, views and privileges.

- ANSI does not support "dynamic SQL" — SQL statements that are computed at execute time.

- The set of reserved words in ANSI is significantly smaller than that in DB2.

- In ANSI, the "Unique" attribute applies to a column or combination of columns as it should, whereas in DB2 it applies to an index (which it should not).

The ANSI version, therefore, is inadequate as a tool for evaluating DBMS products. The remarks about DB2 apply to certain other vendors' products also.

My view of these ANSI items is as follows.

- Omitting catalog tables was a poor judgment; the catalog needs to be standardized. The ANSI version looks like a survivor of non-dynamic CODASYL.

- Failure to support dynamic SQL was another poor choice. This feature is needed and is used.

- The smaller set of reserved words places vendors with relational DBMS products that go beyond the proposed ANSI standard in a potentially difficult situation. Several vendors find themselves in this category.

- The ANSI treatment of the "Unique" attribute is good in my opinion. An index is treated by ANSI as a purely performance-oriented tool, so there are no semantic consequences of dropping one.

My main criticism of the ANSI Level 1 and Level 2 proposed standard for relational data bases is that inadequate attention is given to some very important areas. For example, the comprehensive, dual-mode data sublanguage capability that SQL



## WHAT TO WARE TO SURVIVE THE HIGH SEAS OF BANKING



Ware something extraordinary.  
Ware INFOPOINT™ banking software by UCCEL. The most complete, most proven line of integrated financial software on the market. The software that keeps you cool in calm waters no matter how turbulent the market gets.

With INFOPOINT's 21 major applications, you'll not only survive — you'll survive in high style. And continue to grow and prosper — even while others are going under. What's more, INFOPOINT costs less than ordinary software and can be implemented in less time.

Survive the high seas of banking in high style. Call UCCEL now at 1-800-UCC-1234.

INFOPOINT BY  
**UCCEL**

**Banking software that makes you look good.**

UCCEL CORPORATION, UCCEL TOWER, EXCHANGE PARK, DALLAS, TEXAS 75235. Formerly University Computing Company. UCCEL is a trademark of UCCEL Corporation.

## NOW WE CAN SELL YOU THE WHOLE LINE OF EQUIPMENT YOU NEED.

### INCLUDING THE LINE.

The shortest distance between you and the best business communications system is to head straight to us. NYNEX Business Information Systems.

Because we have everything it takes to build the best customized communications system for your business. The experience, the technical know-how, the products and the people.

As authorized agents of New York Telephone and New England Telephone, we'll provide you with a network that delivers the most reliable, practical and cost-effective voice and data transmission services available today. We're also in the perfect position to build your company a custom network utilizing the latest fiber optic and microwave technologies.

What's more, our equipment solutions

go beyond those of any single manufacturer. Because we've selected all the best products from all the best companies. In fact, we offer a full range of networking, voice and data communications, and computer products that work together in one neatly integrated system.

The result? An efficient system that can meet all your needs today and still grow tomorrow. A system designed, installed and maintained by a team of professionals. And backed by a company with over 100 years of experience that will continue to lead the industry far into the future.

To get our team working on all your communication needs, all it takes is one phone call.

Then compare us with the competition. You'll find we're ahead by a wide margin. Not just a fine line.

For more information, contact  
NYNEX Business Information Systems  
P.O. Box 655, White Plains, NY 10604-0655.  
Or call:

**1-800-346-9X9X** ext. 806

Please have a NYNEX Account Executive  
call me.

Name \_\_\_\_\_

Company \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip \_\_\_\_\_

( ) \_\_\_\_\_

Telephone \_\_\_\_\_

**BUILDING BUSINESS COMMUNICATIONS FROM STRENGTH.<sup>SM</sup>**

**NYNEX**  
Business Information Systems



<sup>SM</sup>Service Mark of NYNEX Corporation. ©1985 NYNEX Business Information Systems.

## IN DEPTH/RELATIONAL DBMS

(as implemented) already possesses is underemphasized. The entire range of SQL implementations from the large mainframes down to the micro is not adequately addressed.

Finally, ANSI ought to extend presently supported SQL to a version that fully supports the relational model, including distributed data base. At the very least, ANSI should generate a statement of direction adequate to permit vendors to extend the fidelity of their products without risking incompatibility with some future standard.

Extensions of SQL that provide this support now can be forecast in detail and with some reliability. Any standard adopted now should not make these extensions impossible or even difficult in the future.

**Three buying factors**

Any buyer confronted with the decision of which DBMS to acquire should weigh three factors heavily. The first factor is the buyer's performance requirements, often expressed in terms of the number of transactions that must be executed per second. The average complexity of each transaction is also an important consideration. Only if the performance requirements are extremely severe should buyers rule out present relational DBMS products on this basis. Even then buyers should design performance tests of their own, rather than rely on vendor-designed tests or vendor-declared strategies.

The second factor is reduced costs for developing new data bases and new application programs. Relational DBMS provide significant reduction in these costs, when compared with either the CODASYL or hierachic approaches. Fourth-generation languages are no substitute, although they may provide some additional productivity.

The third factor is protecting future investments in application programs by acquiring a DBMS with a solid theoretical foundation and reliable support for high productivity and distributability. In every case, a relational DBMS wins on factors two and three. In many cases, it can win on factor one also — in spite of all the myths about performance.

Then the question arises: Which relational DBMS? The system chosen should not only be a DBMS with a good percentage of fidelity to the

“  
At the very least,  
ANSI should  
generate a  
statement of  
direction to permit  
vendors to extend  
their products  
without risking  
incompatibility  
with some future  
standard.”

relational model, but should be extensible at some future time. Ideally a good DBMS will be extended soon to provide 100% support without logically impairing the customer's investment in application programs.

Buyers should be cautious with vendors that make strong claims — claiming the system is "post relational" (especially when no definition for this term is supplied), or claiming that the DBMS choice has no importance. In fact, your choice of DBMS now may well determine how readily your organization adapts to changes in the future.

It is time vendors realize that all the features of the relational model are interrelated and interdependent. Missing features leave large gaps in the integrity control and usability of a DBMS implementation.

There is nothing on the horizon right now that looks strong enough and practical enough to replace the relational approach. Moreover, because the relational approach relies on such a solid theoretical foundation, its lifetime will last much longer than the CODASYL, hierachic or tabular approaches.

I also believe that it will be much easier for relational DBMS users to convert to whatever future approach appears to be superior, for two reasons. The relational approach insists on all information being recorded explicitly. Moreover, the approach has a close tie to first-order predicate logic — a logic on which most of mathematics is based, hence a logic which can be expected to have strength, endurance and many applications.

**Evaluation against the relational model  
(by DBMS)**

	DB2	IDMS/R	Datacom/DB
Relations	Yes	No	No
Base tables	Yes	Yes	Yes
View tables	Yes	No	No
Query tables	Yes	No	Yes
Snapshot tables	No	No	No
Attributes	Yes	Yes	Yes
Domains	No	No	No
Primary keys	Partial	No	No
Foreign keys	No	No	No
Theta select	Yes	Yes	Yes
Project	Yes	Yes	Yes
Theta join	Yes	Partial	No
Outer theta join	No	No	No
Union	Yes	No	No
Outer union	No	No	No
Intersection	Yes	No	No
Set difference	Yes	No	No
Division	Yes	No	No
Relational assignment	Yes	No	No
Maybe theta select	Yes	No	No
Maybe theta join	Yes	No	No
Maybe outer theta join	No	No	No
Maybe divide	Yes	No	No
S/O theta select	No	No	No
S/O theta join	No	No	No
S/O outer theta join	No	No	No
S/O divide	No	No	No
Entity integrity	Partial	No	No
Referential integrity	No	Partial	No
User-defined integrity	Partial	No	No
Total score against relational model (1 for full support, 0 otherwise)	16	4	5
Total score on 12 rules	7	0	0
Grand total score	23	4	5
% fidelity	46	8	10

Figure 2

**You can reach over  
50,000 computer  
professionals in  
Norway.**

Advertise in CW Communications' Norwegian publications and sell your products directly to Norway's thriving computer market. According to International Data Corporation (IDC), the world's leading information industry research firm, Norway's total DP expenditures for 1983 were \$1.6 billion. IDC forecasts that DP expenditures will reach \$3 billion by 1989. The personal computer market is also a lucrative one with expenditures in 1983 of \$58 million.

CW Communications has two publications covering the Norwegian market: *Computerworld Norge* and *PC Mikrodata*.

Twice a month, 20,000 MIS/DP executives read *Computerworld Norge*. Modeled after *Computerworld*, its sister publication in the United States, *Computerworld Norge* is written for the professional end-user and reports on the most timely news concerning new products, new software applications and market trends and opportunities.

And every month, 30,000 professional and home PC users read *PC Mikrodata*. The editorial focuses on a wide range of topics including the latest information on the PC market, new software, hardware, interviews, buyer's guides and instructions for the PC.

CW International Marketing Services makes advertising your products in Norway, and around the world, easy. We have over 50 publications in more than 25 countries. For more information on our wide range of services, call Diana La Muraglia, General Manager, CW International Marketing Services, at (800) 343-6474. On the West Coast, call Isabella Barbagallo at (415) 328-4602 or (800) 277-8365. Or fill out the coupon below and mail today.

**Please send me more information on:**

- Computerworld Norge  
 PC Mikrodata  
 Your other foreign publications

Name \_\_\_\_\_  
Title \_\_\_\_\_  
Company \_\_\_\_\_  
Address \_\_\_\_\_  
City \_\_\_\_\_ State \_\_\_\_\_ Zip \_\_\_\_\_

**Improve Performance!**

The *EDP Performance Review* is a monthly report for EDP managers who want to discover new technologies and ideas for improving performance. Every issue brings you articles and reports from authorities in the field, case studies, user experiences, products and services, CPE calendar and more. The *EDP Performance Review* is like putting a performance/capacity management expert on your staff. The EDP/PR is a management tool no mid or large size DP shop should be without. Just one idea could pay the subscription price many times over.



**ACR**

To receive a FREE review issue of the *EDP Performance Review*, return this coupon with your business card or name and address to: ACR, P.O. Box 9280, Phoenix, AZ 85068-9280.

- 12 Monthly Issues
- 12 Pages
- Annual Product and Literature Reference Issue
- \$95.00/year