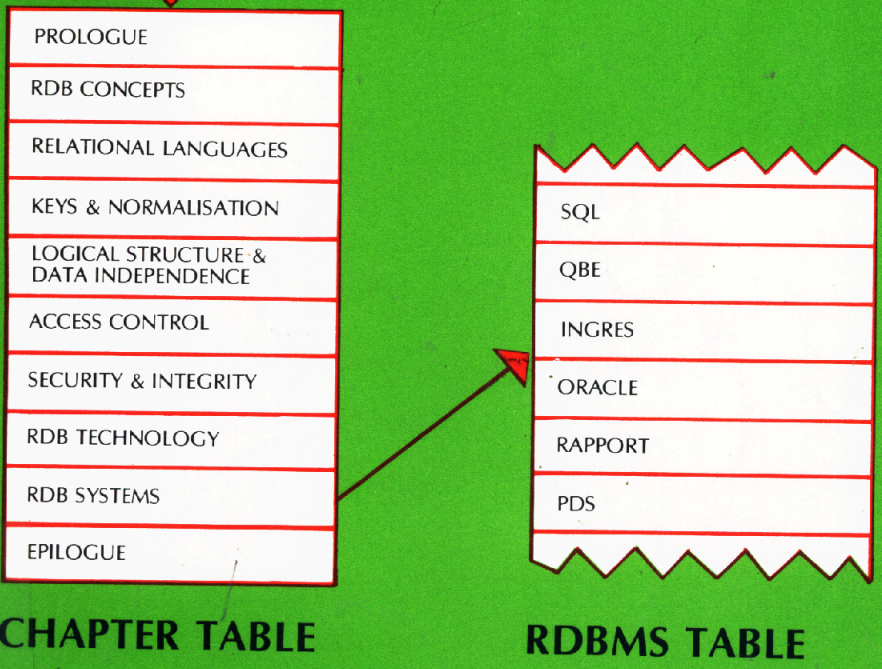
NCC
The National Computing Centre

ALAN MAYNE
MICHAEL B. WOOD

# INTRODUCING RELATIONAL DATABASE





# Introducing Relational Database

Alan Mayne
Michael B Wood



PUBLISHED BY NCC PUBLICATIONS

British Library Cataloguing in Publication Data

Mayne, Alan
Introducing relational database.
1. Electronic data processing – Relational processing
I. Title II. Wood, Michael B.
001.64'42 QA76.9.R/

ISBN 0-85012-397-6



First published in 1983 by:

NCC Publications, The National Computing Centre Limited, Oxford Road, Manchester M1 7ED, England.



Typeset by UPS Blackburn Limited, 76-80 Northgate, Blackburn, Lancashire, and printed by Hobbs the Printers of Southampton.

ISBN 0-85012-397-6

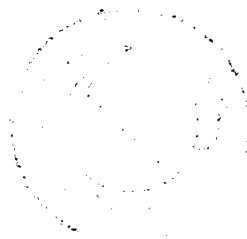
# Acknowledgements

Relational database depends upon the sound foundation of the pioneering work of Dr E F Codd of IBM. The present authors acknowledge their debt to him and to other workers, particularly C J Date. The literature on relational systems is extensive. A short reading list appears at the end of the book. Much of the literature is difficult for the non-mathematical reader.

The authors also thank the various suppliers for their help with information, and for their valuable comments on certain parts of the draft. In many cases we have simplified explanations and omitted details in the interest of clarity. Again the reader should consult suppliers for current product information.

Our thanks also to Rosemary Wood who read and criticised the MS and to Geoff Simons who gave much help as Editor.

The Centre acknowledges with thanks the support of the Electronics and Avionics Requirements Board for the project from which this book derives.



# Prologue

A dynamic young sales executive decided to give up his expense account and become an entrepreneur. It was fashionable, and politicians encouraged the notion that small business start-ups would save the country.

At first our entrepreneur was in total control of his new business. All the information was at his finger tips or in his head. He had a compact and efficient filing system for his data; in fact he had a database: *A database is a collection of stored data organised in such a way that all user data requirements are satisfied by the database. In general there is only one logical copy of each item of data.*

Our entrepreneur's business is small. He only keeps one record of each item of data, because he has no time to do more.

The business expands rapidly until there are several departments, and branch offices. Our entrepreneur can no longer exercise complete control, but must rely upon his managers.

The various managers need data to control their own areas, so each builds his own set of files and ledgers. Ideally they would have shared a common set of data, and used similar systems, but this is not an ideal world. Some managers did know that others were collecting and using the same, or similar data. Some did not care even when they knew. Some managers did not trust the data from others, and would not risk becoming dependent for political reasons. No manager wanted to share his hard-won data with others, because the data might have been used against him in the struggle to the top.

How had success affected the company? When it was a one-man show the data needed to run it was available from one source. Duplication of

data was rare. The company could indeed claim to have a database. However, the company's expansion resulted in data being spread around and there was considerable duplication of data.

No one knew what data was actually being maintained across the company. Managers were blissfully ignorant of the problems created by duplication of data. More correctly, they were ignorant of the cause and attributed the problems to the incompetence of others.

If the address of a customer changed, the information was usually only sent to one department in the company. The department lucky enough to receive the notification could ensure that goods and correspondence were sent to the new correct address. However, other departments would continue to use the old address, being unaware of the change.

The company demonstrated to its customers that its left hand did not know what its right hand was doing. The problem arose because the company held multiple copies of essentially the same items of data which became inconsistent as individual copies were modified. Departmental managers were using out-of-date and inaccurate data to make their decisions. Clearly, this was not the right way to run the business!

Our entrepreneur was now a worried man. He received inconsistent and conflicting reports from his departmental managers, and could not plan the company's future.

One day a computer salesman walked into his office. The computer salesman listened patiently to our entrepreneur's problems as they enjoyed an expense-account lunch. (The salesman had heard it all before, many many times.) The salesman explained how all the problems would be solved by the magic of computers. After making the sale he rushed off to order a new car.

The company created a data processing department, hiring teams of expensive programmers and analysts from outside. Analysts designed elaborate computerised systems to meet the perceived requirements of the various departments. In practice, the managers had found it very difficult to specify precisely what they wanted. Eventually, after many months of development, the data processing department provided an acceptable, though imperfect, system.

The computer processed large volumes of data faster and more reliably than had ever been possible with clerical methods. However, our entre-

preneur still encountered difficulty in getting consistent and accurate reports from his departmental managers. There were still some problems from inconsistent data. (One department refused a customer's business because his credit limit had been exceeded. At the same time another department sent letters to the same customer saying how glad they were to have his business.)

The black box in the computer room failed to solve the original problems. Worse than this, it introduced many new ones.

Departmental managers now complained about the cost and time it took to make apparently trivial changes to these computerised systems. The data processing department told them that they would have to wait months to get an extra piece of data printed on one small report. Moreover they told them that it would cost them several thousand pounds for program modifications.

The data processing manager explained to our entrepreneur why maintenance took so long and was so very costly. In order to print the extra data on the report it was necessary to change the format of the computer files to accommodate the new data. In consequence, every one of the computer programs that used the modified computer file also had to be altered, not only the single report program. This meant that 37 programs were affected by the change, not just one. The programs were unfortunately dependent upon the data, even upon data that they did not actually use.

Our entrepreneur called in the computer salesman once more. He told him of the new problems created by the computer. He also pointed out that it had done little to resolve the original problems. The salesman was unabashed by these remarks. He claimed to understand the problem completely. Indeed he had the perfect solution. The company should "go database".

The salesman explained the benefits of going database. The company's data would be stored on the computer in a way which would allow everyone's requirements to be satisfied. Everyone would share the same data. There would be only one copy of each item of data so all those headaches caused by duplicate data would be solved.

The salesman also explained that to "go database" the company needed to buy a database management system (DBMS). In addition to supporting a database the salesman pointed out that the DBMS software

also provided data independence. This allows new data to be added to support a new program without requiring any existing programs to be modified. It also supported back-up and recovery facilities as well as access control facilities. A Database Management System is formally defined as:

> a general-purpose set of programs that aid and control each user's access to and use of the database for adding, modifying and retrieving data, and that includes facilities for giving data independence, integrity and security.

The sales talk was convincing and the entrepreneur signed the contract and instructed the data processing manager to start using the DBMS to implement new systems. He then waited for the promised benefits of "going database".

Alas not all the benefits were forthcoming. The old problems refused to disappear. Managers still provided him with conflicting reports. The difficulties caused by data duplication persisted. On the positive side though, the DBMS had eased the problems of program maintenance. The integrity and security facilities also proved useful.

Our entrepreneur then called in a Data Management Consultant to find out what had gone wrong. The consultant listened sympathetically. He had heard this story many times before; to him it was like being a marriage guidance counsellor listening to a wife telling of her husband's affairs. After hearing the story and studying the organisation of the company, he explained what had gone wrong and suggested what needed to be done to put things right.

When the company was a one-man show it could claim to have a database, although not a computer-based one. The growth of the company destroyed the database. It was replaced by a collection of filing systems that were not integrated, and had considerable duplication of data. Introducing the computer in no way changed the company's attitude to data. The computerised systems simply replaced manual files with computerised files. Each department still maintained its own set of files for its own computer systems. There was no integration of data among the various departments.

When the company decided to "go database" there was hope that the problems might be resolved. The company did not know what "going database" really meant, and it merely introduced a database manage-

ment system. The DBMS continued to model the fragmentation that actually existed. Each department continued to have its own copy of the data it needed. The DBMS was simply being used as an access method. Use of a DBMS does not automatically mean that a database exists.

*The root of the company's problems was that technical tools were being introduced, first the computer, then the DBMS, in an attempt to solve managerial problems*. The real problem was the company's attitude to its own data. Irrespective of whether the company's data is stored in filing cabinets, conventional computer files, or under the control of a DBMS, it is essential to have a philosophy of sharing data and eliminating, or at least controlling, duplication of data.

The technical problems of sharing data and controlling duplication can be solved relatively easily. The political problems are often much more intractable. Departmental managers have to understand the need for data sharing. They must appreciate that the data they collect and use belongs to the company as a whole. Senior management must define who shall be responsible for updating specific items of data and also who shall be allowed to access which items of data.

Having a database means having a philosophy of sharing data for the benefit of the entire company. An on-going commitment of senior management is necessary. It is all too easy for a database to degenerate into a collection of unintegrated data files. The state of well-organised data is intrinsically unstable. It needs to be continually policed. For this reason it is usual to establish a Data or Database Administration function to act as the police force.

At this point in our story another salesman arrived at the company. He too listened sympathetically to our entrepreneur's tale of woe. Naturally enough he also had the perfect solution. It was a *relational database system*.

What could he have meant?

# Contents