A STORAGE AND ACCESS METHOD FOR MULTIDIMENSIONAL DATABASES

by

Rekha Rani

Department of Computer Science and Engineering

Thesis submitted
in fulfilment of the requirements of
the degree of
Doctor of Philosophy
to the

INDIAN INSTITUTE OF TECHNOLOGY, DELHI
March, 1984

CERTIFICATE

Certified that the thesis entitled 'A Storage and Access Method for Multidimensional Databases' being submitted by Ms. Rekha Rani is a record of work carried out by her under my supervision and guidance. The matter embodied in this thesis has not been submitted for the award of any other degree.

March 1984.

(Prof. K.D. Sharma)
Dept. of Computer
Science and Engg.
IIT, Delhi

ACKNOWLEDGEMENTS

I wish to express my thanks to Dr. K.D. Sharma, Professor, Department of Computer Science and Engineering for his constant guidance and active interest in the completion of my thesis. He spared a lot of his valuable time for very fruitful discussion and meaningful suggestions. His encouragement was a source of great inspiration and driving force for me in pursuing my work.

I also take this opportunity to express my thanks to Dr. R.C. Malhotra, Head, Department of Computer Science and Engineering and Dr. M.K. Jain, Head, Computer Services Centre, who were kind enough to provide all the facilities to me at the Department and Computer Centre respectively.

Without the constant encouragement and help of my family members, I could not have pursued my work.

Mr. Ranjit Kumar who typed this thesis has been very cooperative and helpful and I thank him for it.

REKHA RANI

ABSTRACT

The efficient physical design of multidimensional databases is crucial to the optimal performance of any computer system. Since a large percentage of the databases are used on-line, it is necessary to evaluate the performance characteristic of the physical model before implementing the database. There are only a few methods available which incorporate various system parameter alternatives into the physical design of the multiattribute database. The aim of this thesis is towards bridging this gap. With this end in view, a new data structure - the modified k-d-B tree ... - which may be efficiently used for structuring a dynamic database subject to multiattribute queries is proposed.

The thesis has been divided into eight chapters. The first chapter traces the development of database management systems and the various interfaces which exist between the user and the physical database. The evolution of multiattribute database system is also sketched.

Chapter two is an overview of the different multiattribute database organisations and the advantages and drawbacks of the various organisations is also considered.

Chapter three describes the modified k-d-B tree. The notation and implementation details are described. The algorithms for insertion, searching and deletion have been

developed, implemented and analysed. The modified k-d-B tree has been observed to perform favourably with the existing data structures.

Chapter four is another major contribution of this thesis. It provides a numerical method for finding out integer optimal values of the branching factors for the modified k-d-B tree subject to partial match queries. If probabilities of access to various attributes are known then this knowledge may be incorporated into the method to obtain optimal branching factors under a specified usage pattern.

In chapter five a scheme for choosing a sequence of discriminators for the modified k-d-B tree is presented. Since no formal results are available about the methods for choosing the sequence of discriminators, the results of the proposed scheme are compared with a cyclic sequence (based on decreasing probabilities of access). The method is then modified to take various system parameters like page placement scheme and future use of the database to arrive at a cost matrix. The choice of sequence of discriminators for the different levels is then formulated as a 0-1 integer programming problem.

Chapter six considers the storage requirements of the various database organisation considered in chapter two. The requirements have been formulated within a parameteric framework to enable the user to make a comparative

evaluation of the various storage strategies for different values of the system parameters.

Chapter seven considers the implementation aspects of the modifed k-d-B tree on the magnetic bubble memory (MBM) chip. Since MBMs are likely to become an important and integral part of mini and micro computer system, the issues considered in this chapter help the user to arrive at the optimal values of the parameters to be used when implementing the modified k-d-B tree on the MBM chip.

Chapter eight is a summary of the thesis and provides areas for further research.

TABLE OF CONTENTS

Page

CHAPTER III -- THE MODIFIED k-d-B TREE

Page

CHAPTER VII -- THE MAGNETIC BUBBLE MEMORY CHIP AND THE MODIFIED k-D-B TREE

CHAPTER VIII - CONCLUSIONS AND AREAS OF FURTHER RESEARCH

~~~~~
~~~~~