

1970
The Interactive Generation of Functional Dependencies

by

WILLIAM CIEN HUNT

E. G. S., University of Nebraska at Omaha
Omaha, Nebraska, 1972

A MASTER'S REPORT

submitted in partial fulfillment of the
requirements for the degree

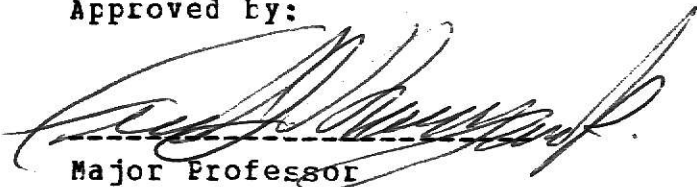
MASTER OF SCIENCE

Department of Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1978

Approved by:



Major Professor

Document

LD

2668

R4

1978

H85

C.2

Acknowledgement

A special thanks to Dr. Fred J. Maryanski for allowing me to participate in the development and implementation of the system for the automatic generation of third normal form relations, and for his professional attitude during its development. Thanks to Dr. Paul S. Fisher and Edwin W. Basham for serving as members of my committee. The implementation was made possible through a grant from NCR, Inc.

Table of Content

1.0	Introduction.....	1
2.0	Background.....	4
2.1	The Problem of Third Normal Form.....	5
2.2	Project.....	6
2.3	FDGEN.....	7
3.0	Functional Dependency Generator.....	9
3.1	The FDGEN Algorithm.....	10
3.2	Amplification of the FDGEN Algorithm.....	17
4.0	Implementation.....	41
4.1	Walk-Through.....	41
4.2	Testing.....	76
5.0	Conclusion.....	94
5.1	Enhancement Considerations.....	94

Reference and Bibliography.....	97
---------------------------------	----

Appendices

Abstract Code	A-1
Implementation - FDGEN Code - NCR Level II COECL	B-1

List of Figures

System Diagram.....	8
Input File Function and Format.....	11
Data Name Table Function and Format.....	12
Functional Dependency Table Function and Format.....	16
Output Format of Functional Dependencies.....	17
Hierarchical Structure of Data-Name-Items.....	19
Input File Example.....	20
Data Name Table Example.....	22
Sample Extracts of an Input File and Data Name Table with status 'S' and 'N'.....	25
Link Stack (Concatenation Stack) Format and Function.....	28
Storage Stack Format and Function.....	30
Save Stack Format and Function.....	31
Functional Dependency Output File Format and Function.....	39
Input File.....	44
Concatenation Code Table.....	46
Logical Internal View of the Data Name Table.....	47
Update Section of the Data Name Table.....	52
Update Logical View of the Data Name Table.....	55
Logical View of the Concatenation and Storage Stack.....	62
Logical View of the F-D-Table.....	62
Current View of the Concatenation and Storage Stacks.....	64
Current View of the Save Stack and F-D-Table.....	68
Current View of the Concatenation and Storage Stacks.....	70
Current Logical View of Internal Tables.....	72
Final View of Program Generated and Updated Tables.....	74
Functional Dependency Output Format.....	75
Walk-Through Generated Functional Dependencies.....	75

Input Test File 1.....	77
Input Test File 2.....	78
Input Test File 3.....	80
Test Schema.....	81
FDGEN Test Path Diagram.....	82
User's Cutput of Test File 1.....	86
User's Cutput of Test File 2.....	87
Test Run 1 -- Input Test File 3.....	88
Test Run 2 -- Input Test File 3.....	89
Test Run 3 -- Input Test File 3.....	90
Cutput File Generated from F-C-Table.....	91

1.0 INTRODUCTION

Information is a prime commodity in today's hurried existence. Managers make decisions based on the best information available to them at the moment. Strategic planning is based on information compiled from many different sources. Every aspect of modern day life is influenced by information. Information is often related to power.

What is information? It is the meaningful interpretation of data which has been collected and integrated. (1) To meet the needs of today's society data must be collected in great quantities. In order for massive amounts of data to be useful to a user, he must have real-time access to it and have a means of synthesizing the available data into meaningful collections so that it might be interpreted into useful information.

Systems have been developed which provide the user with mass storage of data, quick access to the data, and a means of synthesizing the accessed data. The systems, known as Data Base Management Systems (DBMS), allow a user to retrieve, manipulate, and store data without requiring him to know how the data is organized or stored within the storage device.

The user views the data in prespecified logical groups linked together in some organized structure. The more common DBMS structures are known as network and hierarchy. In such views, data may be represented in three ways: (2)