

Thèse Docteur Ingénieur

Par **Michel Léonard**,

« Aide algorithmiques à la conception de bases de données »

Université scientifique et médicale de Grenoble, Institut Polytechnique de Grenoble

Juin – 1976.

Ce document est disponible dans la collection de l'ACONIT, Association pour un conservatoire de l'informatique et de la télématique, Grenoble, **fiche n°24639-05**

Auteur : Michel Léonard

Document de 5 + 4 pages, photocopie en 5 feuillets.

Edition : IMAG – Institut de Mathématiques Appliquées de Grenoble,

Dans les logiciels de **Système de Gestion de Base de Données (SGBD)**, l'IMAG-Grenoble s'est illustré avec le logiciel SOCRATE (1970) devenu ensuite le logiciel CLIO, qui ont été largement utilisés avec succès (voir AC_24639-04)

La présente thèse de docteur-ingénieur témoigne de l'importance de ce logiciel.

Le présent document comprend des extraits de cette thèse : table des matières ; introduction ; bibliographie.

Pour utiliser ce document, mentionnez l'Institut de Mathématiques Appliquées de Grenoble, et ACONIT n° **24639-05**.

Il est rappelé que la loi sur les droits d'auteur n'autorise que les reproductions partielles de documents à titre individuel et pour des motifs de recherche.

Fiche rédigée le 26/08/2020 par Jean Ricodeau.

INAG - 1976 - LEO

AC_24633-05

117023
012382

THESE

présentée à 42.4

UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE
INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

POUR OBTENIR LE GRADE DE
DOCTEUR-INGENIEUR

INSTITUT IMAG

Informatique, Mathématiques Appliquées de Grenoble

CNRS - INPG - USMG

MÉDIATHÈQUE

P.P. 68

38402 ST-MARTIN-D'HÈRES CEDEX

1 A 508

Tél. (73) 51 46.36

Michel LEONARD

AIDES ALGORITHMIQUES

A LA

CONCEPTION DE BASES DE DONNÉES.

Thèse soutenue le 28 Juin 1976 devant la Commission d'Examen :

Président : L. BOLLIET

Examineurs : C. DELOBEL

F. PECCOUD

E. PICHAT

Je tiens à remercier :

Monsieur Le Professeur BOLLIET : il m'a fait l'honneur de présider le Jury de cette thèse,

Monsieur DELOBEL, Maître de Conférences : il m'a initié à la recherche et ne m'a ménagé ni son temps ni ses conseils,

Monsieur PECCOUD, Maître de Conférences, et Monsieur FICHAT, Maître de Conférences : ils m'ont apporté les critiques et les conseils indispensables pour réaliser cette thèse,

Monsieur ADIBA : il m'a particulièrement aidé par ses encouragements et ses compétences,

Monsieur DE DONMARTIN : il m'a montré l'application des résultats de recherche dans le domaine industriel et l'intérêt d'une démarche rigoureuse dans la prise de décisions,

Madame REYNAUD-GAROCHE et Madame STIERS : elles ont toujours répondu avec toute leur compétence à mes nombreuses questions,

Monsieur PORTAL : il m'a fait bénéficier de son expérience dans le domaine des bases de données,

Madame DOREL : elle a su rendre cette thèse lisible de par son application et son goût du travail bien fait,

Monsieur IGLESIAS et les membres de son service : ils ont réalisé matériellement cet ouvrage.

Michel LEONARD

Le Laboratoire de Mathématiques Appliquées de Grenoble a offert le cadre intellectuel et matériel de cette thèse ; le C.N.R.S. en a fourni le support financier par l'intermédiaire de l'ATP n° 650.764.73.01.

Plan de thèse

INTRODUCTION

1. MODELE RELATIONNEL DE REPRESENTATION D'UNE BASE DE DONNEES :	
RAPPELS	I.1
1.1 Relations	I.1
1.2 Principales opérations définies sur les relations	I.6
1.3 Relation fonctionnelle	I.8
1.4 Décomposition d'une relation	I.12
2. TRANSFORMATION D'UNE STRUCTURE RESEAU DANS UN ENSEMBLE DE RE-	
LATION	II.1
2.1 Structure arborescente et relations	II.2
2.2 Aspects relationnels d'une structure réseau : cas parti-	
culier des structures SOCRATE et CODASYL	II.3
2.3 Définition d'une structure RESEAU et transformation des	
structures SOCRATE et CODASYL en une structure RESEAU	II.13
2.4 Transformation d'une structure RESEAU en un ensemble de	
relations	II.16
2.5 Exemple de transformation d'une structure SOCRATE ou CO-	
DASYL en un ensemble de relations	II.18
2.6 Conclusions	II.25
3. ELIMINATION DE REDONDANCE ET D'INCOHERENCES D'INFORMATIONS	III.1
3.1 Imperfection d'une base et règles d'intégrité	III.3
3.2 Formes d'une relation	III.11
3.3 Décomposition d'une relation en sous-relations directes	III.16
3.4 Construction d'un algorithme de décomposition d'une rela-	
tion en sous-relations directes	III.26
3.5 Conclusions	III.28

4. RESTRUCTURATION D'UNE BASE DE DONNEES	IV.1
4.1 "Trois niveaux d'abstraction de la resturation d'une base de données"	IV.3
4.2 Formalisation du processus de restructuration	IV.12
4.3 Restructurations possibles d'une base de données	IV.21
4.4 Conclusions	IV.27
5. TRANSFORMATIONS D'UN ENSEMBLE DE RELATIONS DANS UNE STRUCTURE DE DONNEES DE TYPE RESEAU	V.1
5.1 Transformation d'un ensemble de relations en une structure RELATIONNELLE	V.3
5.2 Transformation d'une structure RELATIONNELLE en une structure RESEAU	V.10
5.3 Transformation d'une structure RESEAU en une structure SOCRATE ou CODASYL	V.18
5.4 Conclusions	V.24
6. REALISATIONS TECHNIQUES	VI.1
6.1 Transformation d'une structure SOCRATE en un ensemble de relations	VI.3
6.2 Analyse d'un ensemble de relations	VI.14
6.3 Saisie d'un ensemble de relations	VI.18

CONCLUSIONS

ANNEXE A Algorithme de fermeture et de couverture minimale

Annexe B Propriétés des décompositions arborescentes et de leurs combinaisons avec des relations fonctionnelles

Liste des principales définitions

Bibliographie

EVOLUTION D'UNE BASE DE DONNEES

L'évolution d'une base de données peut être rendue nécessaire pour plusieurs raisons :

1) manque de pertinence des informations fournies qui, comme l'a montré CODD [9], peut résulter de la structure elle-même : celle-ci de par sa définition, contient des sources d'incohérences logiques lors des mises à jour des informations (cf. Chapitre 3) ;

2) manque d'efficacité des traitements qui peut être causé par l'exécution des procédures liées à la cohérence d'informations ou de relations redondantes. DELOBEL et CASEY [13] ont montré le lien étroit qui existe entre ce point et le précédent ;

3) modification des règles de fonctionnement de l'organisation pour laquelle la base a été créée ; comme à ces règles correspondent généralement des règles d'intégrité de la base, c'est une restructuration qu'il est nécessaire d'envisager ;

4) évolution des besoins en informations et manque d'efficacité des traitements qui réclament la création de nouveaux chemins d'accès logiques aux données et ainsi une restructuration

5) évolution des logiciels et des configurations de matériel ; "pour permettre aux utilisateurs de bénéficier des avantages et des possibilités de nouveaux logiciels et (ou) de configurations de matériel plus économiques, il devient de plus en plus important de développer des techniques qui permettent la conversion entre le nouveau et l'ancien système" (Fry et Jeris [18]).

Toutes ces raisons obligent à reconsidérer l'organisation d'une base de données. Fry et Jeris [18] séparent les problèmes posés en deux :

- restructuration
- et - réorganisation de l'espace physique,

tout en faisant remarquer qu'une réorganisation de base de données dissocie rarement dans la pratique ces deux aspects. Ils définissent la restructuration plus comme un changement du schéma de structure des données que de celui de sa représentation physique.

"Ce changement est susceptible de modifier les relations et la composition de blocs. Certes, la représentation physique change également, mais sa modification est commandée par le changement du schéma" [18].

Comme Fry et Jeris [18] et Navathe et Fry [22], nous nous sommes intéressés seulement à la restructuration et nous avons cherché à répondre aux questions suivantes :

D : Comment peut-on transformer une structure pour éliminer les incohérences d'informations qu'elle cause ? (Chapitre 3).

E : Comment diminuer la redondance de relations et d'informations que contient une base de données ? (Chapitre 3).

La réponse à ces questions permettra d'améliorer la qualité des services rendus par une base de données.

F : Comment peut-on transformer une structure pour tenir compte de l'évolution des règles d'intégrité ? (Chapitre 4).

Il s'agit de déterminer une nouvelle structure cohérente avec les nouvelles règles et les opérations à mettre en oeuvre pour atteindre ce but.

G : Quelles sont toutes les relations qui sont contenues implicitement dans une base de données ? (Chapitre 2).

H : Quand peut-on affirmer que deux structures sont équivalentes ? (Chapitre 4).

Mac Gee [21] a montré comment ce problème est lié à celui de la transposition d'une base de données d'un logiciel vers un autre logiciel quand on cherche à préserver l'information ; il a cherché à donner des éléments intuitifs de solution.

Bibliographie

[1] J.R. ABRIAL

"Projet SOCRATE"

Cours avancé sur l'Architecture des Systèmes.

Alpe d'Huez. Décembre 1972.

[2] J.R. ABRIAL

"Data Semantics"

IFIP TC-2 Working Conference on Data Base Management Systems,

Campese Avril 1974.

[3] M. ADIBA, C. DELOBEL, M. LEONARD

"An Unified Approach for Modelling Data in Logical Data base Design".

IFIP Workshop TC2, Freudenstadt, Janvier 1976.

[4] M. ADIBA, M. LEONARD

"Approche algorithmique de la conception d'un système de bases de données".

Rapports TS0, TS1, TS2.

Séminaire de Programmation de Grenoble, Novembre 1975.

[5] P.A. BERNSTEIN, J.R. SWENSON, D.C. TSICHRITZIS

"A unified Approach to Functional Dependencies and Relations".

Proc. 1975 ACM-SIGMOD Workshop on Management of Data, San Jose,

May 1975.

[6] P.A. BERNSTEIN

Normalization and Functional Dependancies in the Relational Data Base Model.

Thèse, Université de Toronto, Octobre 1975.

[7] J. BOITTIEUX - ZIDANI

"Etude mathématique des relations d'un ensemble de notions"

Contrat DGRST 65 FR 201 (juillet 1967).

[8] CODASYL

Data Base tack group report,

ACM, New York, 1971.

[9] E. CODD

"Further normalization of the relational data base model"

IBM Research report, RJ 909 (august 1971).

[10] E. CODD

"Normalized data base structure : a brief tutorial"

ACM-SIGFIDET Workshop on data description access and control,

San Diego, novembre 1971.

[11] E. CODD

"Recent Investigations in Relational Data Base Systems",

Proc. of IFIP 1974, North-Holland.

[12] C.J. DATE

"An Introduction to Database Systems".

ADDISON-WESLEY PUBLISHING COMPANY, 1975.

[13] C. DELOBEL, R.G. CASEY

"Decomposition of a data base and the theory of boolean switching functions"

IBM RJ. 1016 (avril 1972).

- [14] C. DELOBEL
"Contributions théoriques à la conception et l'évaluation
d'un système d'informations appliqué à la gestion".
Thèse, Université de Grenoble, Octobre 1973.
- [15] C. DELOBEL, M. LEONARD
"The Decomposition Process in a relational model".
IRIA Workshop on Data Structures, Namur, Mai 1974.
- [16] C. DELOBEL, D. PORTAL
"Aide algorithmique à l'organisation logique d'un fichier".
Colloque INFORSID-IRIA, Aix-en-Provence, Mai 1975.
- [17] C. DELOBEL
"Les Systèmes de Bases de Données.
Ecole d'Eté de l'AFCEP 1975.
- [18] J.P. FRY, D.W. JERIS
"Towards a Formulation and Definition of data reorganization".
Data Translation Working Paper 801.
Université du Michigan, Octobre 1973.
- [19] D. KNUTH
"The Art of Computer Programming".
Vol. 3. Addison Wesley Publishing Company (1973).
- [20] J. KUNTZMANN
"Algèbre de Boole".
Dunod, 1965.
- [21] W.C. MCGEE
"A Contribution to the study of Data Equivalence".
IFIP TC-2 Working Conference on Data Base Management Systems,
Cargèse, Corse, Avril 1974.
- [22] S.B. NAVATHE, J.P. FRY
"Restructuring for large data bases : Three levels of abstraction".
Data Translation Project : Technical Report 8.1, Septembre 1975.

- [23] G.M. NIJSSEN
"Set and Codasyl set or coset".
Ecole d'Eté de Freudenberg, Août 1975.

- [24] D. PORTAL
"Conception d'un système automatisé de gestion de la
scolarité de l'enseignement supérieur".
Thèse, Université de Grenoble, Mars 1975.

- [25] M. TREHEL
"Catalogues, treillis de catalogues, génération et
décomposition".
Mathématiques et Sciences Humaines n° 40, 1972.

- [26] C.P. WANG, H.H. WEDEKIND
"An Approach for Segment Synthesis in Logical Data Base
Design".
IBM Journal of Research and Development, Vol. 19, n° 1,
Janvier 1975.