

资料(数据)库系统

概 论

陳榮源·蔡易章 編譯

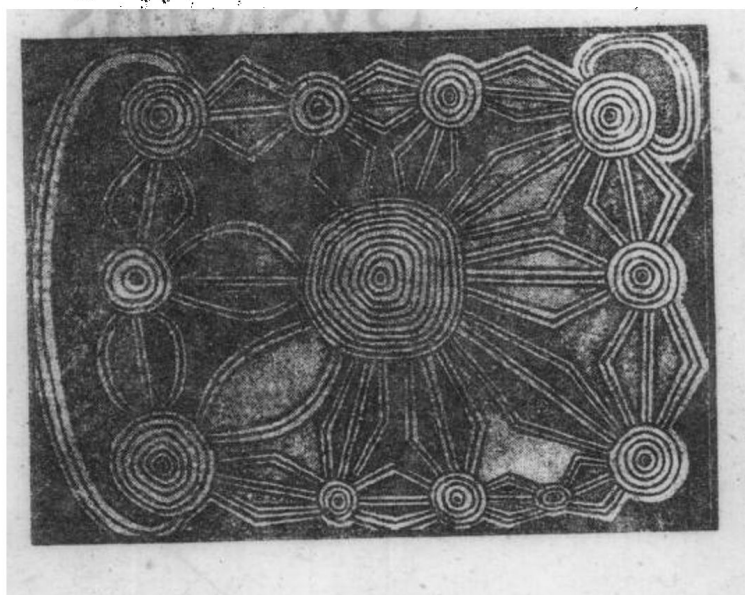
全華科技圖書股份有限公司

世界图书出版公司重印

大 專 用 書

資料(數據)庫系統 概 論

陳榮源·蔡易章 編譯



全華科技圖書股份有限公司 印行

世界圖書出版公司重印

资料(数据)库系统概论

陈荣源·蔡易章 编译

全华科技图书股份有限公司出版

世界图书出版公司 重印

(北京朝内大街137号)

北京中西印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

1990年6月第1版 711×1245_{1/16}

1990年6月第1次印刷 印张29.67

印数: 0001—1200 字数: 70.3万字

ISBN 7-5062-0724-9/TP·5

定价: J4, 30元

本书经全华科技图书股份有限公司香港和中国大陆总代理鑫港
出版社有限公司特许世界图书出版公司独家重印, 限国内发行

An Introduction to Database Systems

VOLUME I

Fourth Edition

C. J. DATE

第四版序言

本書的第四版異於先前各版本，實際上是一本新書。如下面所解釋的，甚至連書中之結構也改變了。但是總目標仍依然如舊——即，為根本的資料庫技術教育提供堅實的基礎，其一般的風格及解說之層次仍是相同的；本書主要是用於教科書，而非工作的參考書，且仍然保留其介紹性及指導性的韻味。

此處假設讀者專業地對某些資料處理模式有興趣，例如，他們可能是系統分析師或設計師、應用程式師、系統程式師、大學中電腦系的學生，或教此課程的老師。讀者最好對現代電腦系統之能力有適當的認識，且對該系統之檔案處理特色有特別的印象；他們也至少應懂得一種高階語言。由於這些事前的要求並不是特別的需要，然而，我希望本書將適合任何關心使用或製作資料庫的人，或任何想擴展電腦科學領域之一般知識的人，使之成為一種介紹性的教科書。

本書共分成六個主要部分：

1. 基本概念
2. 一個關連性系統範例
3. 關連性資料庫管理
4. 資料庫環境
5. 非關連性系統
6. 一些目前的發展

每一部分中是由很多章所組成：

■ 第一部分（共三章）廣義的介紹一般資料庫系統之概念，且特別地說明關連性系統與非關連性系統間之差異。在撰寫本書之

時所裝置的系統，仍以非關連性系統佔大多數。然而，這種情勢將不會持久——關連性產品以漸增的速率被發展出來及被安裝；確實地，幾乎所有現代的產品都是關連性的。更者，目前幾乎所有研究發展都基於關連性的觀念。因此，本書將比較強調關連性的方向。

■ 第二部分（共六章）以 **IBM系統DB2** 為工具，來舉例說明關連性系統的主要概念。

■ 第三部分（共八章）是本書最長的部分（確實地，可被看成書中之書），其對於理論及實際的，基礎的關連性系統（如**DB2**）之觀念有深入的探討。本部分也詳盡討論另一個關連性系統**INGRES**，其在很多方面和**DB2**有顯著的不同。

■ 第四部分（共三章）則討論資料庫系統（關連性的或其他）的一些其他模式——復元，一致性，安全，整合等等。

■ 其次，第五部分（共三章）介紹主要非關連性的作法（反轉串列，階層的，及網狀），及描述一些非關連性系統之代表：**DATACOM/DB**，**IMS**，及**IDMS**。

■ 最後，第六部分（共二章）討論一些目前在資料庫領域的研究方向。

注意：此六個部分被刻意的寫成略為相互獨立；假如讀者喜歡，第三部分（至少第11章至第14章）可在第二部分之前研讀，第五部分可先於第二部分及第三部分研讀，或在第六部分之後研讀。然而，依拙見則是依本書所安排之次序來閱讀。

大部分每章之後都附有習題及（通常有）解答；這些解答經常針對習題的主題附有額外的資訊。大部分每章之後跟隨很多參考書目之名稱，且這些參考書目大部分都附有註解說明。文章中以中括號包含數字來表示參考書。例如，[11.1]代表第11章後面之參考書目中的第一項參考書，即E.F.Codd於**Communications of the ACM (CACM)**，Vol.13, No.6 (June 1970)中所發表的論文。

對於熟悉本書第三版的讀者，本版和第三版主要不同之處歸

納如下。

第一部分：舊版介紹的章（第1章）被擴展成二章（第1章及第2章），且內容比較溫和。舊版第2章談到的儲存體結構（新版第3章），則被完全重寫且擴大談論範圍。舊版第3章談到不同之資料模型，已被刪除；然而，該章所有的材料仍然存於（確實的以擴充形式）本書其他重點之中。

第二部分：新版4～9章取代舊版5～10章。所有談到以系統R為標準的材料，均被以DB2產品的材料所取代。

第三部分：此部分大多數是新的，或至少由第三版大大改進而來。第10章（新的）稍深入描述INGRES系統。第11至14章提出重要的關於關連性模型之擴充及改進處理（取代舊版第4及11～13章）；特別地，深入含蓋了該模型之整合性模式。第15章（新的）討論一些關連性系統的實用模式。第16章（新的）介紹關連性查詢最佳化。最後，第17章是舊版第14章所談到的進一步正規化之修訂及擴充。

第四部分：此部分亦是新的。它包含新的三章。第18章討論到異動處理（特別是，復元及一致性之考慮）；第19章描述資料庫系統中之安全及整合問題；及第20章討論一些輔助（前端）系統——報告撰寫器，商用繪圖系統，自然語言介面，應用產生器等等。

第五部分：此部分取代舊版的第3～5部分。其包括3個冗長的章，各說明一種主要的非關連性方向（反轉串列，階層的，及網狀）。於各章中，首先討論相關的資料模型，然後再描述一個提供該模型的典型系統。第21章（新的）談到DATACOM/DB；第22章（舊版第15～22章之修訂及濃縮）談到IMS；及第23章（舊版第23～26章之修訂及濃縮）談到IDMS。注意，第23章較特別，其描述一個（重要的）具體的系統，而非從未完全付諸實施的“正式的”CODASYL規格。第21及23章亦包含DATACOM/DB及IDMS之新“關連性”延伸之題材。

第六部分：此部分亦是新的，第 24 章描述分散式資料庫系統並簡單討論資料庫機器。第 25 章介紹語意模型製作；特別地，其描述實體／關係模型及擴充關連性模型 RM/T。

另外，雖然舊版第 27 及 28 章已被刪除，但此兩章中之很多材料仍被融入本書各種要點之中。

以下是為何有這些改變的理由：

1. 首先，關連性系統之諾言——回溯本書第一版（1975）中之預言——至少已經開始成熟。如前所述，幾乎所有現代資料庫產品是關連性的，且幾乎所有關於資料庫的研究都是基於關連性的。更者，目前大部分老舊（非關連性的）產品都被擴充，以提供某種關連性；因此，花費在非關連性系統上的研究似乎已不再“等於”用在關連性系統上的時間。
2. 跟隨上述要點：由於關連性方向被廣泛地接受，對於研究及從業人員而言，具實了解此方向的基本原則也顯得比以往更重要——事實上，是對於任何涉及資料庫管理模式的人。因此，本書中之關連性部分被完全地重新編寫，同時也將之強化及大大的擴充，以提供該具實了解之基礎為目標。
3. 再者，“成熟的”關連性系統是可以利用現有的商業用品來舉例及討論，而不只是像以往版本中的研究規範。（當然，更不用說，商業的產品是會不斷的改進；因此，特殊系統的細節是不能保證會和最新的版本一樣精確。更者，本書引用某特別產品與否，並不表示對該產品之支持或其他——之所以包括某特殊系統，只是用來製作某特別要點或舉例說明某特殊觀念。）
4. 縮減篇幅之非關連性系統部分，包含了一些第三版中沒有提及的論題。特別地，至少本書現在包括以下所有項目之介紹：查詢最佳化，復元，一致性，安全，整合，以格式為基礎的系統，自然語言介面，報告撰寫器，繪圖介面，應用產生器，語言模型製作（包括實體／關係模型製作），分散式資料庫，及資料庫機器。注意：很多這些論題在本書第二冊中

有更深入的說明（見下面）。

附帶一提的是，這些改變所產生的結果之一，即是和前面各版本之序言不再有所關連。因此，在此版中將它們刪除。

如上所述，本書有一續篇，即 *An Introduction to Database System: Volume II*（第一版，Addison-Wesley, 1982）——本書中簡稱第二冊（Volume II）。然而，我希望以上的討論是要使讀者確認第一冊之第四版並非只是改寫前（第三）版及第二冊之材料；相反的，其包容許多新題材，及對於現有材料之顯著不同的強調。同樣地，也不企圖取代針對各論題做更深入探討的第二冊。

致謝：我很樂意將本書之成就直接或間接和對本書出版有所貢獻的人士一起來分享。首先，本書之教材從我以往指導過的討論會中，綜合許多學生的寶貴意見而得到很大的助益，也來自很多對本書之第三版有所心得的人士：Randell Flint, Alan Hevner, Larry Rowe, Sharon Salveter, Stewart Shen, Paul Spirakis, Philip Stone, Richard Walters, 特別是 Dave Dewitt。我也感激以下各位審閱本版某部分原稿，或提供技術性材料，或幫忙找尋某技術性問題答案的人：Paul Butterworth, Roger Buchanan, Don Casey, Ted Codd, Dave Peter, Kreps, Jack Nevison, Larry Rowe, Phil Shaw, Mike Stonebraker, 及 Sharon Weinberg。我也感激 Relational Technology 公司能允許以我曾為該公司所寫的文章“INGRES之虛擬程式撰寫工具”其中第20章之某部分為依據。最後，本書之出版若沒有我家人的鼓舞或我的編輯 Elydia Siegel 及 Addison-Wesley 的全體人員之專門性的技術指導，將無法順利地出書。再次由衷的感謝他們。

C.J.Date

作者簡介

C.J.Date 是一位自立的作者，講師，及顧問。他尤其專精於關連性資料庫系統，目前在加州的 Saratoga。

Date 先生於 1967 年在英國加入 IBM，當時從事的工作是整合資料庫的功能，將之納入 PL/I 之中。後來，於 1974 年他又轉移到加州的 IBM 發展中心，負責有關資料庫語言 UDL (unified database language) 之設計。近年來，他參加 IBM 關連性產品 SQL/DS (給 DOS 用的發表於 1981 年，給 VM 用的發表於 1983 年) 及 DB2 (發表於 1983 年，是用於 MVS) 之技術性策劃及外貌之設計。1983 年五月離開了 IBM。

Date 先生活躍於資料庫領域超過 15 年，他是首先認知 Codd 首倡之關連性模型之重要性的人之一，他之廣泛地有關技術性主題之講演——主要是關於資料庫的題目，尤其是在於關連性資料庫——普及全美國、歐洲、澳洲、拉丁美洲、及遠東。除了本書以外，他的著作尚包括 An Introduction to Database Systems : Volume II (1982)，其中含蓋更進一步的資料庫系統模式；Database : A Primer (1983)；其將資料庫系統 (由微電腦至大型電腦) 以非專家的觀點看待；及 A Guide to DB2 (1984)，是描述 IBM 關連性產品 DB2 及其伙伴 QMF 及 DXT。他也發表了許多技術性的論文及文章，也對資料庫理論做了很多貢獻。目前，他正於美國及其他地方指導一系列有關專業資料庫技術的討論會。

前 言

由於具有創造力的程式設計師及管理人員之努力，以產生快速成長之電腦工業所需之實用系統程式，因而使系統程式之領域有所擴充。當時，程式之撰寫被視為一種藝術來從事，經由一些近鄰同事之切磋，每位程式師針對問題都有獨到的解答。1968年IBM的已故Ascher Opler先生認為有必要將程式撰寫的知識組合成一種能被所有系統程式師所知曉的形式。觀察這種情形，他決定以足夠有用的材料來支持重要的編纂努力。由於他的推薦，IBM決定贊助系統程式系列(The System Programming Series)，將之視為長遠的計劃以收集，組織及出版這些對整個工業具有持續價值的原理及技術。

此系列是由開放式的原文參考書之蒐集所組成。每本書的內容代表個別作者對該主題範圍之觀點，而不是反應IBM公司的觀點。每本書被安排給課程使用，但是也足夠詳盡以用於參考。更者，此系統分成三個層次：廣泛介紹性材料之基礎書籍，較專門材料之軟體書籍，及很特殊理論之電腦科學書籍。就目前情況看來，該系列符合初學者，有經驗的程式師，及電腦科學家之需求。

總言之，此系列為系統程式之藝術性的寫實，其可做為系統程式訓練的技術基礎。

編輯部

譯者序

在探討資料庫領域書籍中，本書是一本銷路非常好的教科書，因此，編譯的方式強調介紹性和指導性。根據資料結構與運算符號而言，目前市面上流行的資料庫系統可分為四類：關連性（relational）、逆串列式（inverted list）、階層式（hierarchical）和網路式（network）。不過，近幾年來，資料庫方面的發展幾乎都是關連性的，即使是舊系統的供應商也企圖加入關連性的方法來增強其產品功能，本書第四版的發行就是反應了學術上的發展與市場需要的趨勢。

本書分為 6 部份，共有 25 章，這 6 部份彼此是相互獨立，讀者可依自己的意願研讀，多數的章節之後附有習題和部份解答，加上本書重點是敘述關連性系統，配合非關連性系統作比較，內容方面是強調理論與實際應用並重。有助於讀者對資料庫系統觀念的建立和了解，這就是選擇本書最主要的目的。

不過，本書只是第一冊，即將出版的第二冊除了討論其他主題外，對於本書第 3，4，6 部份的主題亦有更深入的探討。這一連貫性的介紹與討論很適合系統程式師、系統分析師、應用程式師以及電腦相關科系的教師和學生研讀。

本書是利用公餘時刻編譯而成，在文辭和內容方面，雖力求貼切、完善，但疏漏不適之處，在所難免，尚祈專家學者不吝指正幸。

編譯者蔡易章、陳榮源 謹識

編輯部序

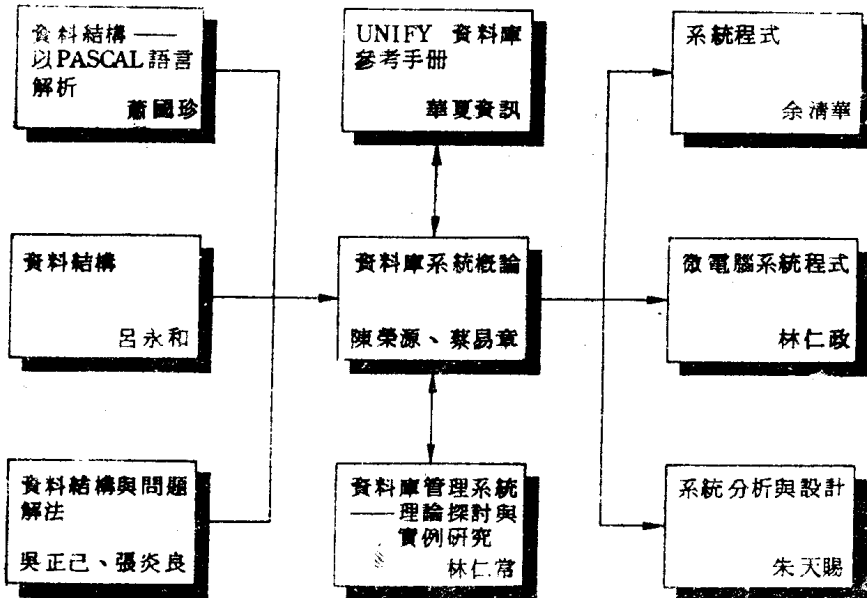


「系統編輯」是我們的編輯方針，我們所提供給您的，絕不只是一本書，而是關於這門學問的所有知識，它們由淺入深，循序漸進。

本書譯自 C. J. Date 資料庫系統概論的第四版，是資料庫系統理論的經典之作，新版增加的項目計有：查詢最佳化、復元、一致性、安全、整合、以格式為基礎的系統、自然語言介面、報告撰寫器、繪圖介面、應用產生器、語言模型製作、分散式資料庫、及資料庫機器等題材，是大學電子計算機系、資訊系資料庫系統課程的最佳教材。

同時，為了使您能有系統且循序漸進研習資料庫系統方面叢書，我們以流程圖方式，列出各有關圖書的閱讀順序，以減少您研習此門學問的摸索時間，並能對這門學問有完整的知識。若您在這方面有任何問題，歡迎來函連繫，我們將竭誠為您服務。

流 程 圖



目 錄

第 1 部份 基本概念

第 1 章

資料庫管理系統概觀

3

1.1 範例介紹 3

1.2 什麼是資料庫系統？ 5

1.3 運算用資料 9

1.4 為何要有資料庫？ 12

1.5 資料獨立性 15

1.6 關連性系統及其他 20

習 題 22

參考文獻和書目 23

部份習題解答 26

第 2 章

資料庫系統的架構

29

2.1 目 的 29

2.2 架構的三個層次 29

2.3 外部層 33

2.4 概念層 35

2.5 內部層 36

2.6 對 映 37

2.7 資料庫管理師 37

2.8 資料庫管理系統 40

2.9	資料通訊	40
	習題	41
	參考文獻和書目	42
第3章	內部層	45
3.1	介紹	45
3.2	資料庫存取：概觀	47
3.3	頁集合及檔	51
3.4	索引	59
3.5	雜亂	67
3.6	指標鍊	73
3.7	壓縮技術	76
3.8	結論短評	80
	習題	80
	參考文獻和書目	83
	部份習題解答	89

第2部份 一個關連性系統範例

第4章	DB2概觀	97
4.1	背景	97
4.2	關連性資料庫	98
4.3	SQL 語言	103
4.4	系統主要成分	110
	參考文獻和書目	116
第5章	資料定義	119
5.1	介紹	119
5.2	基表	119
5.3	索引	123

5.4	結論短評	125
	習題	126
	部份習題解答	127
第6章	資料處理	129
6.1	介紹	129
6.2	簡單的查詢	130
6.3	聯合查詢	134
6.4	內建函數	139
6.5	高等特色	143
6.6	更新操作	151
6.7	結論短評	156
	習題	159
	部份習題解答	163
第7章	系統目錄	171
7.1	介紹	171
7.2	目錄查詢	173
7.3	更新目錄	174
	習題	176
	部份習題解答	176
第8章	觀點	179
8.1	介紹	179
8.2	觀點定義	182
8.3	觀點之DML操作	184
8.4	邏輯資料獨立	189
8.5	觀點之優點	191
	習題	192
	部份習題解答	192