

# 目 次

はじめに

1	データベースの世界	1
1.1	データベースの利用分野	2
1.2	データベース概観	9
1.3	本書の内容の概観	12
2	基本的事項	17
2.1	集合とその演算	18
2.2	関係とグラフ	21
(a)	関係	21
(b)	グラフ	23
(c)	木構造	24
2.3	有限状態のモデル	24
(a)	有限オートマトン	25
(b)	ワークフロー	25
(c)	一重マルコフ連鎖	28
2.4	データの符号化	28
(a)	暗号	29
(b)	誤り検出符号と誤り訂正符号	30
(c)	データ圧縮	32
2.5	データの検索とソート	34

3 関係データベース言語	39
3.1 データベース操作の基礎	40
(a) クロスワードと結合	40
(b) クロスワードと選択	41
(c) クロスワードと射影	41
(d) 関係の定義	42
(e) 射影	42
(f) 選択	44
(g) 制約	44
(h) 結合	45
(i) 集合演算	45
(j) 関係除算	46
(k) 集約演算	46
(l) 質問の種類	47
(m) データ操作機能	47
3.2 標準データベース言語 SQL の基礎	48
(a) 検索	48
(b) ブール演算	48
(c) 区間演算	49
(d) 順序づけ	49
(e) グループ化	49
3.3 各種結合演算	50
(a) 直積	50
(b) 結合	51
(c) 外結合	51
(d) 自己結合	52
(e) 入れ子問合わせ	53

3.4	オンライン解析処理 OLAP	54
(a)	集約関数	56
(b)	OLAPの実現方法	56
(c)	データキューブとその操作	56
(d)	GROUPING SETS	58
(e)	ROLL UP	58
(f)	CUBE 演算	59
(g)	区間合計の効率化	60
(h)	スタースキーマ	61
(i)	トップ $n$ 質問	62
(j)	オンライン集約	63
4	情報検索とウェブ情報処理	65
4.1	情報検索の基礎	66
(a)	文書の特徴ベクトル	68
(b)	質問の種類	69
(c)	距離	70
(d)	類似度	70
(e)	情報検索の質	71
4.2	集合検索	73
(a)	転置ファイル	73
(b)	ハッシュ	73
(c)	ハッシュ関数による処理の効率化	75
4.3	距離検索とクラスタリング	76
(a)	クラスタリングの目的	77
(b)	基本的なクラスタリング手法	77
(c)	利用者のクラスタリング	78
4.4	部分マッチ検索	78

(a)	有限オートマトンによる部分マッチ	78
(b)	部分系列の索引	80
(c)	比較	82
4.5	ウェブ上の情報検索	82
(a)	ページの重要性	83
(b)	ハブとオーソリティ	84
5	データベース設計とデータモデル	87
5.1	関係データベースで表現できる一貫性制約	88
5.2	実体関連モデルによる表現	97
(a)	データベース設計への利用	99
(b)	設計のための例として用いる実体関連図	100
5.3	関係モデルの設計	101
5.4	種々のデータモデル	114
(a)	非正規表現	115
(b)	多値従属性と非正規表現	115
(c)	非正規表現から正規表現に	116
(d)	ネットワークモデル	117
(e)	階層データモデル	118
(f)	ネットワークモデルにおける意味制約	118
(g)	半構造データモデル	119
(h)	オブジェクト指向モデル	121
6	情報構造	125
6.1	記憶階層とデータ構造	126
(a)	直接アクセスと順アクセス	127
(b)	揮発性・不揮発性と信頼性	128
(c)	記憶装置の特性	128

(d)	記憶階層	128
(e)	キャッシュの考え方	129
(f)	記憶装置とファイル構成	130
6.2	ディスクの基本的性質と高速化	131
(a)	ブロック	132
(b)	バッファ	132
(c)	外部ソート	133
(d)	キャッシュに記憶すべきデータの決定	134
(e)	索引の種類	136
6.3	木構造索引	137
(a)	節点の種類	138
(b)	おもな木構造索引	139
(c)	B <sup>+</sup> 木の定義	140
(d)	データ挿入の例	142
(e)	多次元情報を扱うためのファイル構造	143
6.4	ハッシュにもとづくファイル構成	143
(a)	拡張可能ハッシュ	144
(b)	線形ハッシュ	146
6.5	複数ディスクの利用による高速化と高信頼化	147
(a)	RAIDの手法	148
(b)	RAIDのレベル	150
7	トランザクション処理	153
7.1	トランザクション	154
(a)	トランザクションに要求される基本的性質	155
(b)	トランザクションの並行処理	156
7.2	直列可能性の実現	158

(a) 直列可能性	158
(b) 直列可能性の保証	159
(c) 直列可能性の実現法	160
(d) 二相施錠方式	160
(e) 二相施錠方式における直列可能性	161
(f) デッドロック	162
(g) デッドロックの発見法	163
(h) 基本的な施錠	163
(i) 狭義二相施錠	164
(j) 時刻印による並行処理制御	165
(k) 比較	166
7.3 各種高性能データベースと並行処理	167
(a) アクティブデータベース	167
(b) 実時間データベース	168
(c) 主記憶データベース	168
7.4 協調処理のためのトランザクション処理	169
(a) 多バージョン法	169
(b) 読み出し専用トランザクションの定義	170
(c) 確定前のデータの読み出し	171
(d) ネストトランザクション	171
(e) 補償トランザクション	171
(f) 分散システムにおける処理	171
(g) ワークフローにおけるトランザクション処理	172
さらに勉強するために	175
索引	179