

目 次

第1章	はじめに	1
1.1	背景	1
1.2	簡潔データ構造の歴史	2
1.3	本書の構成	2
第2章	基本事項	5
2.1	計算モデル	5
2.2	標準的な記号と関数	7
2.2.1	オーダ表記	7
2.2.2	イェンセンの不等式	8
2.2.3	スターリングの公式	9
2.2.4	文字列と集合	10
2.2.5	順序木と探索順	10
2.3	情報理論的下限	12
2.4	簡潔データ構造	13
2.5	エントロピー	14
2.6	整数の符号化	16
2.6.1	1進数符号	17
2.6.2	ゴロム符号	17
2.6.3	ガンマ符号	17
2.6.4	デルタ符号	18
2.7	整数列の符号化	18
2.7.1	ハフマン符号	18

viii — 目 次

2.7.2	インターバル符号	18
2.7.3	MTF 符号	19
2.7.4	Elias-Fano 符号	20
第 3 章	基本的な簡潔データ構造	23
3.1	ビットベクトルの簡潔データ構造	23
3.1.1	<i>rank</i> の計算	24
3.1.2	<i>select</i> の計算	26
3.2	パターンに対する <i>rank/select</i>	28
3.3	疎なベクトルの簡潔データ構造	30
3.4	非常に疎なベクトルの簡潔データ構造	33
3.4.1	ランダムアクセス可能な Elias-Fano 符号	34
3.4.2	ランダムアクセス可能なガンマ符号	36
3.4.3	ランダムアクセス可能なデルタ符号	36
3.4.4	疎なベクトルでの定数時間 <i>rank</i> 索引	37
3.4.5	非常に疎なベクトルでの <i>rank</i> 索引	42
3.5	下限	42
3.6	実装上の工夫	43
3.6.1	<i>rank</i> の高速化	44
3.6.2	<i>select</i> の高速化	45
3.7	文献ノート	45
第 4 章	ウェーブレット木	49
4.1	文字列での <i>rank/select</i>	49
4.2	アルファベットサイズが大きいとき	54
4.3	その他の演算	57
4.3.1	区間内の異なる文字の列挙	57
4.3.2	ある文字より小さい文字の数	59
4.4	ハフマン型ウェーブレット木	59
4.5	多分岐ウェーブレット木	61

4.6	直接アドレス可能符号	61
4.7	直交領域探索	63
4.8	文献ノート	68
第 5 章	区間最小値問い合わせ	69
5.1	問題の定義	69
5.2	RMQ を LCA に帰着	70
5.3	LCA を RMQ に帰着	72
5.4	± 1 RMQ 問題	74
5.5	RMQ 問題の定数時間アルゴリズム	76
5.6	RMQ 問題の $4n$ ビットデータ構造	77
5.7	RMQ 問題の $2n$ ビットデータ構造	79
5.8	サイズの下限	81
5.9	文献ノート	82
第 6 章	順序木	85
6.1	順序木の基本操作	85
6.2	LOUDS 表現	87
6.2.1	LOUDS 表現の定義	87
6.2.2	LOUDS を用いた木の演算	89
6.2.3	LOUDS を用いたラベル付き木	91
6.3	括弧列 (BP) 表現	92
6.3.1	BP 表現の定義	92
6.3.2	<i>findclose</i> のデータ構造	96
6.3.3	<i>enclose</i> の計算	100
6.3.4	最近共通祖先の計算	102
6.4	DFUDS 表現	103
6.4.1	DFUDS 表現の定義	103
6.4.2	DFUDS での基本操作	105
6.4.3	最近共通祖先の計算	106

x — 目 次

6.4.4	DFUDS 表現の圧縮法	108
6.4.5	全2分木の効率的な表現	109
6.5	BP 表現のより簡単なデータ構造	111
6.5.1	超過配列	111
6.5.2	$O(\lg n)$ 時間データ構造	113
6.5.3	区間最大最小木	118
6.5.4	大きな木に対するデータ構造	119
6.6	動的な簡潔順序木	127
6.7	文献ノート	130
第7章	文字列検索のデータ構造	133
7.1	文字列検索の基本問題	133
7.2	接尾辞配列	136
7.3	接尾辞木	137
7.4	圧縮接尾辞配列	142
7.4.1	接尾辞配列の圧縮	142
7.4.2	自己索引化	145
7.4.3	後方探索	146
7.4.4	Ψ の圧縮	149
7.5	圧縮接尾辞木	151
7.5.1	木構造	151
7.5.2	ノードの文字列深さの表現	152
7.5.3	枝ラベルの表現	155
7.5.4	木の巡回操作	155
7.5.5	圧縮接尾辞木の計算量	158
7.6	文書集合に対するデータ構造	158
7.6.1	文書列挙問題のための索引	159
7.6.2	単語頻度の計算法	161
7.6.3	文書頻度の計算法	162
7.7	文献ノート	164

第 8 章 BW 変換	167
8.1 ブロックソート圧縮法	167
8.2 逆 BW 変換と <i>LF</i> 関数	170
8.3 FM-index	173
8.4 圧縮接尾辞配列と FM-index の関係	175
8.4.1 Ψ から <i>BW</i> を計算	176
8.4.2 <i>BW</i> から Ψ を計算	176
8.4.3 <i>BW</i> と Ψ の相互変換	177
8.5 双方向 BW 変換	177
8.6 ラベル付き木の圧縮	179
8.6.1 XBW 変換	179
8.6.2 XBW を用いた木の巡回操作	182
8.6.3 部分パス問い合わせ	184
8.7 de Bruijn グラフの圧縮	185
8.7.1 グラフの定義	186
8.7.2 簡潔 de Bruijn グラフ	188
8.7.3 簡潔 de Bruijn グラフのデータ構造	190
8.7.4 グラフ上の操作の実現	194
8.8 文献ノート	197
参考文献	199
索 引	211