DISTRIBUTED ALGORITHMS AND PROTOCOLS

Michel Raynal

Professor of Computer Science
IRISA (INRIA-CNRS, University of Rennes), France

Translated by Jack Howlett, *St Cross College, Oxford*

Technische Hochschule Darmstadt
FACHBEREICH INFORMATIK
BIBLIOTHEK
Inventar-Nr.: 7560
Sachgebiete:
Standort:

JOHN WILEY & SONS

Chichester · New York · Brisbane · Toronto · Singapore

CONTENTS

Preface ix

1 Introduction to distributed algorithms 1
1 Parallelism and distribution 1
2 Networks and distributed systems 2
2.1 General and local networks 2
2.2 Characteristics of distributed systems 3
2.3 What should be distributed, and why? 4
3 Distributed algorithms 5
3.1 Basic elements: processes, communication paths 5
3.2 Features of distributed algorithms 8
3.3 Classifying distributed algorithms 10
4 Some concepts and techniques 10
4.1 The approach to design 10
4.2 Concepts and techniques 11
4.3 Communication + ordering = control 14
References 15

2 Election and mutual exclusion algorithms 19
1 Introduction 19
2 The mutual exclusion problem 19
3 The Ricart and Agrawala/Suzuki Kasami algorithm 21
3.1 Overview of other algorithms 21
3.2 Assumptions 21
3.3 Principle of the algorithm 22
3.4 The algorithm 22
3.5 Proof of the algorithm 24
3.6 Messages and time-stamping 25
4 An algorithm for regenerating the token 26
4.1 A token circulating on a logical ring 26
4.2 Loss of token: Misra's algorithm 26
5 Elective algorithms 29
5.1 Introduction 29
5.2 The Chang and Roberts algorithm 30
5.3 The Hirschberg and Sinclair algorithm 33
5.4 The algorithm of Dolev, Klawe and Roden 37
5.5 Other algorithms 40
References 41

3 Algorithms for detection and resolution of deadlock **43**
1 Introduction 43
1.1 The problem of deadlock 43
1.2 Characterization of deadlock situations 44
2 Distribution of a centralized algorithm: Lomet's algorithms 49
2.1 Global and local states 49
2.2 Lomet's first algorithm: replication of the global state 50
2.3 Lomet's second algorithm: use of partial state information 51
3 The Rosenkrantz, Stearns and Lewis algorithm 53
3.1 Principle of the algorithm: use of time-stamping 53
3.2 'Wait-die' method 54
3.3 'Wound-wait' method 55
3.4 Comments on these algorithms 55
4 Algorithms for detecting deadlocks 56
4.1 Proposed methods 56
4.2 Algorithms with centralized control 57
4.3 Algorithms with hierarchical control 57
4.4 Algorithms with distributed control 58
5 Deadlocks due to communications: algorithm of Chandy, Misra and Haas 58
5.1 Resumé: features of the problem 58
5.2 Assumptions and principles underlying the algorithm 59
5.3 The CMH algorithm 59
5.4 Other algorithms 64
References 64

4 Algorithms for detecting termination **67**
1 Introduction: the problem of termination 67
1.1 Distributed termination 67
1.2 Termination and deadlock 68
1.3 Principles underlying the solutions 69
2 Use of diffusing computation: algorithm of Dijkstra and Scholten 69
2.1 Assumptions 69
2.2 Basis of the algorithm 71
2.3 The algorithm 71
2.4 Comments on the algorithm, proof of validity 72
3 Termination on a ring: algorithm of Dijkstra, Feijen and van Gasteren 74
3.1 Assumptions 74
3.2 Principles of the algorithm 74
3.3 Actions at a site 75
3.4 The algorithm 76
3.5 Disadvantages: Topor's algorithm 77
4 Misra's algorithm 78
4.1 Basis and assumptions 78

4.2 The token in Misra's algorithm 79
4.3 The algorithm 79
4.4 Some comments 81
5 *Use of time-stamping: Rana's algorithm* 81
5.1 Context and assumptions 81
5.2 Principle of the algorithm 82
5.3 The algorithm 84
5.4 Proof of correctness 85
6 *A note on some other algorithms* 87
References 87

5 Protocols for data transfer **89**
1 Introduction 89
2 Protocols for the implementation of CSP 90
2.1 The CSP language: a short introduction 90
2.2 Silberschatz's protocol 93
2.3 Bernstein's protocol 96
2.4 Other protocols 103
3 Methods for reliable broadcasting of messages 103
3.1 The problem 103
3.2 Context of the problem, assumptions 104
3.3 Principles of protocol 104
3.4 Schneider, Gries and Schlichting's protocol 106
3.5 Comments 109
References 110

6 Management of distributed data **113**
1 Introduction 113
1.1 Nature of the data 113
1.2 Distribution of data 114
1.3 Problems to be discussed 115
2 Consistency of duplicated data 115
2.1 Context of the problem 115
2.2 Detection of mutual inconsistency: algorithm of Parker *et al.* 116
2.3 Maintaining mutual consistency 122
2.4 Initializing a new site 127
3 Distribution of control algorithms 128
3.1 Introduction 128
3.2 Construction of a total ordering 128
3.3 Distributed atomicity 130
References 133

7 Problems of gaining consensus in the presence of uncertainties (or how to avoid Byzantine quarrels) **137**
1 The problem of consensus 137
1.1 The problem and its formulation 137

1.2 Features of the solutions 139
2 *The Lamport, Shostak and Pease algorithm* 141
2.1 Assumptions 141
2.2 A criterion for impossibility 142
2.3 Underlying principles of the solution 143
2.4 The algorithm 144
2.5 Proof of correctness 147
2.6 Complexity of the algorithm 148
2.7 Other assumptions concerning the network 149
2.8 Other algorithms 150
3 *Solutions using signed messages* 151
3.1 Assumptions, importance of signatures 151
3.2 The Lamport, Shostak and Pease algorithm 152
3.3 The Dolev and Strong algorithm 154
3.4 The Dolev and Reischtug algorithms 156
4 *Broadcasting in a bus-connected system* 157
4.1 The problem: assumptions 157
4.2 The Babaoglu and Drummond algorithm 158
5 *Conclusion* 160
References 162