

# Distributed Algorithms

An Intuitive Approach

Wan Fokkink

The MIT Press  
Cambridge, Massachusetts  
London, England

---

# Contents

<b>Preface</b> .....	ix
<b>1 Introduction</b> .....	1
<hr/>	
<b>I Message Passing</b>	
<hr/>	
<b>2 Preliminaries</b> .....	7
<b>3 Snapshots</b> .....	13
3.1 Chandy-Lamport algorithm .....	14
3.2 Lai-Yang algorithm .....	15
<b>4 Waves</b> .....	19
4.1 Traversal algorithms .....	19
4.2 Tree algorithm .....	23
4.3 Echo algorithm .....	24
<b>5 Deadlock Detection</b> .....	27
5.1 Wait-for graphs .....	27
5.2 Bracha-Toueg algorithm .....	29
<b>6 Termination Detection</b> .....	37
6.1 Dijkstra-Scholten algorithm .....	38
6.2 Weight-throwing algorithm .....	39
6.3 Rana's algorithm .....	40
6.4 Safra's algorithm .....	42
<b>7 Garbage Collection</b> .....	47
7.1 Reference counting .....	47
7.2 Garbage collection implies termination detection .....	50
7.3 Tracing .....	51

<b>8</b>	<b>Routing</b> .....	53
	8.1 Chandy-Misra algorithm .....	53
	8.2 Merlin-Segall algorithm .....	55
	8.3 Toueg's algorithm .....	58
	8.4 Frederickson's algorithm .....	61
	8.5 Packet switching .....	65
	8.6 Routing on the Internet .....	67
<b>9</b>	<b>Election</b> .....	73
	9.1 Election in rings .....	73
	9.2 Tree election algorithm .....	77
	9.3 Echo algorithm with extinction .....	79
	9.4 Minimum spanning trees .....	80
<b>10</b>	<b>Anonymous Networks</b> .....	87
	10.1 Impossibility of election in anonymous rings .....	87
	10.2 Probabilistic algorithms .....	88
	10.3 Itai-Rodeh election algorithm for rings .....	89
	10.4 Echo algorithm with extinction for anonymous networks .....	91
	10.5 Computing the size of an anonymous ring is impossible .....	93
	10.6 Itai-Rodeh ring size algorithm .....	94
	10.7 Election in IEEE 1394 .....	96
<b>11</b>	<b>Synchronous Networks</b> .....	101
	11.1 A simple synchronizer .....	101
	11.2 Awerbuch's synchronizer .....	102
	11.3 Bounded delay networks with local clocks .....	105
	11.4 Election in anonymous rings with bounded expected delay .....	106
<b>12</b>	<b>Crash Failures</b> .....	111
	12.1 Impossibility of 1-crash consensus .....	112
	12.2 Bracha-Toueg crash consensus algorithm .....	113
	12.3 Failure detectors .....	115
	12.4 Consensus with a weakly accurate failure detector .....	116
	12.5 Chandra-Toueg algorithm .....	116
<b>13</b>	<b>Byzantine Failures</b> .....	121
	13.1 Bracha-Toueg Byzantine consensus algorithm .....	121
	13.2 Mahaney-Schneider synchronizer .....	125
	13.3 Lamport-Shostak-Pease broadcast algorithm .....	127
	13.4 Lamport-Shostak-Pease authentication algorithm .....	130
<b>14</b>	<b>Mutual Exclusion</b> .....	135
	14.1 Ricart-Agrawala algorithm .....	135
	14.2 Raymond's algorithm .....	137
	14.3 Agrawal-El Abbadi algorithm .....	140

---

**II Shared Memory**

---

<b>15 Preliminaries</b> .....	145
<b>16 Mutual Exclusion II</b> .....	147
16.1 Peterson's algorithm .....	147
16.2 Bakery algorithm .....	150
16.3 $N$ registers are required .....	152
16.4 Fischer's algorithm .....	152
16.5 Test-and-test-and-set lock .....	153
16.6 Queue locks .....	155
<b>17 Barriers</b> .....	161
17.1 Sense-reversing barrier .....	161
17.2 Combining tree barrier .....	162
17.3 Tournament barrier .....	165
17.4 Dissemination barrier .....	168
<b>18 Self-Stabilization</b> .....	171
18.1 Dijkstra's token ring for mutual exclusion .....	171
18.2 Arora-Gouda spanning tree algorithm .....	175
18.3 Afek-Kutten-Yung spanning tree algorithm .....	177
<b>19 Online Scheduling</b> .....	181
19.1 Jobs .....	181
19.2 Schedulers .....	182
19.3 Resource access control .....	188
<b>Pseudocode Descriptions</b> .....	193
<b>References</b> .....	221
<b>Index</b> .....	225