
Contents

1	Introduction	1
1.1	Purpose	2
1.2	Organization	4
2	Representations for Genetic and Evolutionary Algorithms ..	9
2.1	Genetic Representations	10
2.1.1	Genotypes and Phenotypes	10
2.1.2	Decomposition of the Fitness Function	11
2.1.3	Types of Representations	13
2.2	Genetic and Evolutionary Algorithms	15
2.2.1	Principles	15
2.2.2	Functionality	16
2.2.3	Schema Theorem and Building Block Hypothesis	18
2.3	Problem Difficulty	22
2.3.1	Reasons for Problem Difficulty	22
2.3.2	Measurements of Problem Difficulty	25
2.4	Existing Recommendations for the Design of Efficient Representations	28
2.4.1	Goldberg's Meaningful Building Blocks and Minimal Alphabets	28
2.4.2	Radcliffe's Formae and Equivalence Classes	29
2.4.3	Palmer's Tree Encoding Issues	31
2.4.4	Ronald's Representational Redundancy	31
3	Three Elements of a Theory of Representations	33
3.1	Redundancy	35
3.1.1	Redundant Representations and Neutral Networks	35
3.1.2	Synonymously and Non-Synonymously Redundant Representations	38
3.1.3	Complexity Model for Redundant Representations	45

3.1.4	Population Sizing for Synonymously Redundant Representations	47
3.1.5	Run Duration and Overall Problem Complexity for Synonymously Redundant Representations	49
3.1.6	Analyzing the Redundant Trivial Voting Mapping	50
3.1.7	Conclusions and Further Research	57
3.2	Scaling	59
3.2.1	Definitions and Background	59
3.2.2	Population Sizing Model for Exponentially Scaled Representations Neglecting the Effect of Genetic Drift	61
3.2.3	Population Sizing Model for Exponentially Scaled Representations Considering the Effect of Genetic Drift	65
3.2.4	Empirical Results for BinInt problems	68
3.2.5	Conclusions	72
3.3	Locality	73
3.3.1	Influence of Representations on Problem Difficulty	74
3.3.2	Metrics, Locality, and Mutation Operators	76
3.3.3	Phenotype-Fitness Mappings and Problem Difficulty	78
3.3.4	Influence of Locality on Problem Difficulty	81
3.3.5	Distance Distortion and Crossover Operators	84
3.3.6	Modifying BB-Complexity for the One-Max Problem	86
3.3.7	Empirical Results	89
3.3.8	Conclusions	93
3.4	Summary and Conclusions	95
4	Time-Quality Framework for a Theory-Based Analysis and Design of Representations	97
4.1	Solution Quality and Time to Convergence	98
4.2	Elements of the Framework	99
4.2.1	Redundancy	99
4.2.2	Scaling	100
4.2.3	Locality	101
4.3	The Framework	104
4.3.1	Uniformly Scaled Representations	104
4.3.2	Exponentially Scaled Representations	106
4.4	Implications for the Design of Representations	108
4.4.1	Uniformly Redundant Representations Are Robust	109
4.4.2	Exponentially Scaled Representations Are Fast, but Inaccurate	111
4.4.3	Low-locality Representations Are Difficult to Predict, and No Good Choice	113
4.5	Summary and Conclusions	115

5	Analysis of Binary Representations of Integers	117
5.1	Integer Optimization Problems	118
5.2	Binary String Representations	120
5.3	A Theoretical Comparison	123
5.3.1	Redundancy and the Unary Encoding	123
5.3.2	Scaling, Modification of Problem Difficulty, and the Binary Encoding	126
5.3.3	Modification of Problem Difficulty and the Gray Encoding	127
5.4	Experimental Results	129
5.4.1	Integer One-Max Problem and Deceptive Integer One-Max Problem	129
5.4.2	Modifications of the Integer One-Max Problem	134
5.5	Summary and Conclusions	139
6	Analysis and Design of Representations for Trees	141
6.1	The Tree Design Problem	142
6.1.1	Definitions	142
6.1.2	Metrics and Distances	144
6.1.3	Tree Structures	145
6.1.4	Schema Analysis for Graphs	146
6.1.5	Scalable Test Problems for Graphs	147
6.1.6	Tree Encoding Issues	150
6.2	Prüfer Numbers	151
6.2.1	Historical Review	152
6.2.2	Construction	154
6.2.3	Properties	156
6.2.4	The Low Locality of the Prüfer Number Encoding	157
6.2.5	Summary and Conclusions	169
6.3	The Characteristic Vector Encoding	171
6.3.1	Encoding Trees with Characteristic Vectors	171
6.3.2	Repairing Invalid Solutions	172
6.3.3	Bias and Non-Synonymous Redundancy	173
6.3.4	Summary	177
6.4	The Link and Node Biased Encoding	178
6.4.1	Motivation and Functionality	179
6.4.2	Bias and Non-Uniformly Redundant Representations	183
6.4.3	The Node-Biased Encoding	184
6.4.4	A Concept for the Analysis of Redundant Representations	187
6.4.5	Population Sizing for the Link-Biased Encoding	191
6.4.6	The Link-and-Node-Biased Encoding	195
6.4.7	Experimental Results	197
6.4.8	Conclusions	200
6.5	Network Random Keys (NetKeys)	201

6.5.1	Motivation	202
6.5.2	Functionality	202
6.5.3	Properties	207
6.5.4	Uniform Redundancy	208
6.5.5	Population Sizing and Run Duration for the One-Max Tree Problem	210
6.5.6	Conclusions.....	212
6.6	Conclusions.....	213
7	Analysis and Design of Search Operators for Trees	217
7.1	NetDir: A Direct Representation for Trees	218
7.1.1	Historical Review.....	218
7.1.2	Properties of Direct Representations	219
7.1.3	Operators for NetDir	220
7.1.4	Summary.....	223
7.2	The Edge-Set Encoding	224
7.2.1	Functionality	225
7.2.2	Bias	227
7.2.3	Performance for the OCST problem	230
7.2.4	Summary and Conclusions.....	237
8	Performance of Genetic and Evolutionary Algorithms on Tree Problems	241
8.1	GEA Performance on Scalable Test Tree Problems.....	242
8.1.1	Analysis of Representations.....	242
8.1.2	One-Max Tree Problem	246
8.1.3	Deceptive Trap Problem for Trees	251
8.2	GEA Performance on Optimal Communication Spanning Tree Problems	256
8.2.1	The Optimal Communication Spanning Tree Problem..	257
8.2.2	Optimization Methods for the Optimal Communication Spanning Tree Problem.....	258
8.2.3	Description of Test Problems	260
8.2.4	Analysis of Representations.....	262
8.2.5	Theoretical Predictions on the Performance of Representations	264
8.2.6	Experimental Results	266
8.3	Summary.....	272
9	Summary and Conclusions	275
9.1	Summary	275
9.2	Conclusions.....	277

A	Optimal Communication Spanning Tree Test Instances	281
A.1	Palmer's Test Instances	281
A.2	Raidl's Test Instances	285
A.3	Berry's Test Instances	289
A.4	Real World Problems	291
	List of Symbols	317
	List of Acronyms	321
	Index	323