

AN ANALYSIS OF THE BEHAVIOR OF A CLASS
OF GENETIC ADAPTIVE SYSTEMS

by
Kenneth Alan De Jong

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Computer and Communication Sciences)
in the University of Michigan
1975

Doctoral Committee:

Professor John H. Holland, Chairman
Associate Professor Larry K. Flanigan
Associate Professor Richard A. Volz
Associate Professor Bernard P. Zeigler

ACKNOWLEDGEMENTS

I would like to thank those persons who made this research not only possible, but also worthwhile and enjoyable. A special note of appreciation to John H. Holland whose enthusiasm for and insight into genetic algorithms provided a constant source of encouragement; to my committee members Larry K. Flanigan, Richard A. Volz, and Bernard P. Zeigler for their time and interest; and especially to my wife, Ruth, who chose to remain in that status even though called upon to be typist, editor, mother, housewife, teacher, and companion during this period.

This research was partially supported by the National Aeronautics and Space Administration under grant NSG 1176. The computer simulations were done on the excellent facilities provided by the University of Michigan Computing Center.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	11
LIST OF TABLES.	v
LIST OF ILLUSTRATIONS	vi
Chapter 1: FORMAL ADAPTIVE SYSTEMS	1
1.1: Introduction.	1
1.2: Some Problems for Adaptation.	2
1.2.1: Data Structure Design	2
1.2.2: Algorithm Design.	3
1.2.3: Game-playing Programs	4
1.2.4: Two-armed Bandits	4
1.3: A Formal Framework.	5
1.4: The Problem of Function Optimization.	7
1.5: A Reduction in Scope.	10
1.6: Summary	13
Chapter 2: GENETIC ADAPTIVE MODELS	15
2.1: Introduction.	15
2.2: Genetic Population Models	16
2.3: Reproductive Plans.	17
2.4: The Basic Reproductive Plan: R_1	20
2.5: K -armed Bandits	24
2.6: Hyperplane Analysis of R_1	40
2.7: An Example of R_1	44
2.8: Summary	47
Chapter 3: STOCHASTIC EFFECTS IN FINITE GENETIC MODELS.	48
3.1: Introduction.	48
3.2: The Problem of Premature Convergence.	48
3.3: Genetic Drift	53
3.4: The Effects of Population Size on R_1	58
3.5: The Effects of Mutation Rate on R_1	67
3.6: The Effects of Crossover Rate on R_1	72
3.7: The Effects of Generation Gap on R_1	77
3.8: Improving the Performance of R_1 on F_1	83
3.9: Summary	91
Chapter 4: PERFORMANCE EVALUATION OF GENETIC ADAPTIVE PLANS.	
4.1: Introduction.	96
4.2: The Performance of R_1 on E	96

TABLE OF CONTENTS

4.3:	Elitist Model R2	101
4.4:	Expected Value Model R3.	107
4.5:	Elitist Expected Value Model R4.	114
4.6:	Improving the Performance of R4 on F5.	128
4.7:	Crowding Factor Model R5	136
4.8:	Generalized Crossover Model R6	148
4.9:	Summary.	160
Chapter 5:	PERFORMANCE ANALYSIS OF FUNCTION OPTIMIZERS	163
5.1:	Introduction	163
5.2:	Local Optimization Techniques.	164
5.3:	Performance Evaluation Conventions	166
5.4:	Performance Evaluation of PRAXIS and DFP	169
5.5:	Summary.	185
Chapter 6:	SUMMARY AND CONCLUSIONS.	190
Appendix A:	THE ENVIRONMENT E	196
A.1:	Introduction	196
A.2:	Test Function F1	196
A.3:	Test Function F2	197
A.4:	Test Function F3	200
A.5:	Test Function F4	203
A.6:	Test Function F5	203
Appendix B:	RANDOM SEARCH ON E	211
B.1:	Introduction	211
B.2:	Validating RANDOM on E	212
B.3:	RANDOM on F1	214
B.4:	RANDOM on F2	215
B.5:	RANDOM on F3	219
B.6:	RANDOM on F4	224
B.7:	RANDOM on F5	226
Appendix C:	PLAN R1 ON E	233
C.1:	Introduction	233
C.2:	Plan R1 on F1.	236
C.3:	Plan R1 on F2.	239
C.4:	Plan R1 on F3.	243
C.5:	Plan R1 on F4.	246
C.6:	Plan R1 on F5.	249
C.7:	Robustness of Plan R1.	249
BIBLIOGRAPHY		253

LIST OF TABLES

<u>Table</u>		<u>Page</u>
3.1:	The data associated with a single run of R1 on F1	50
4.1a:	Off-line performance of R1 on E.	99
4.1b:	On-line performance of R1 on E.	100
4.2a:	Off-line performance of R4 on E.	120
4.2b:	On-line performance of R4 on E.	121
4.3a:	Off-line performance of R5 on E.	146
4.3b:	On-line performance of R5 on E.	147
4.4a:	Off-line performance of R6 on E.	158
4.4b:	On-line performance of R6 on E.	159
5.1a:	Off-line performance indices for PRAXIS and DFP on E	186
5.1b:	On-line performance indices for PRAXIS and DFP on E	187

LIST OF ILLUSTRATIONS

<u>Figures</u>	<u>Page</u>
2.1: Expected losses over 50 trials on bandits B1(9,1) and B2(8,1)	27
2.2: Optimal losses incurred over T trials on two bandits B1(9,1) and B2(8,1)	29
2.3: Optimal distribution of T trials between two bandits B1(9,1) and B2(8,1)	30
2.4: DTS expected losses over 50 trials on bandits B1(9,1) and B2(8,1)	32
2.5: A comparison of the expected losses for DTS and the optimal on two bandits B1(9,1) and B2(8,1)	34
2.6: Simulated losses over 100 trials using TVS on two bandits B1(9,1) and B2(8,1)	36
2.7: A comparison of expected losses over T trials on two bandits B1(9,1) and B2(8,1)	38
2.8: A comparison of the allocation of T trials to two bandits B1(9,1) and B2(8,1)	39
3.1: The rate of allele loss due to genetic drift as a function of population size	56
3.2: The rate of allele loss due to genetic drift as a function of population size	57
3.3: The rate of allele loss due to genetic drift as a function of the mutation rate	59
3.4: The rate of allele loss due to genetic drift as a function of the mutation rate	60
3.5: The effects of population size on allele loss for R1 on test function F1	63
3.6: The effects of population size on off-line performance of R1 on test function F1	65
3.7: The effects of population size on on0line performance of R1 on test function F1	66
3.8: The effects of mutation rate on allele loss for R1 on test function F1	69
3.9: The effects of mutation rate on off-line performance of R1 on test function F1	70
3.10: The effects of mutation rate on on-line performance of R1 on test function F1	71
3.11: The effects of crossover rate on allele loss for R1 on test function F1	74
3.12: The effects of crossover rate on off-line performance of R1 on test function F1	75
3.13: The effects of crossover rate on on-line performance of R1 on test function F1	76
3.14: The effects of generation gap on allele loss of R1 on test function F1	80

LIST OF ILLUSTRATIONS

<u>Figures</u>	<u>Page</u>
3.15: The effects of generation gap on off-line performance of R1 on test function F1. . . .	81
3.16: The effects of generation gap on on-line performance of R1 on test function F1. . . .	82
3.17: Off-line performance of R1 on F1 as a function of crossover rate and generation gap. . .	86
3.18: On-line performance of R1 on F1 as a function of crossover rate and generation gap. .	87
3.19: Off-line performance of R1 on F1 as a function of mutation rate.	89
3.20: On-line performance of R1 on F1 as a function of mutation rate.	90
3.21: Off-line performance of R1 on F1 as a function of population size.	92
3.22: On-line performance of R1 on F1 as a function of population size.	93
4.1: Allele loss for R2 on F1	103
4.2: Off-line performance curves for R2 on F1	104
4.3: On-line performance curves for R2 on F1. . . .	105
4.4: Allele loss for R3 on F1	110
4.5: Off-line performance curve for R3 on F1. . . .	111
4.6: On-line performance curve for R3 on F1	112
4.7: Allele loss for R4 on F1	115
4.8: Off-line performance curve for R4 on F1. . . .	116
4.9: On-line performance curve for R4 on F1	117
4.10: A comparison of off-line performance curves generated on F1.	123
4.11: A comparison of off-line performance curves generated on F2.	124
4.12: A comparison of off-line performance curves generated on F3.	125
4.13: A comparison of off-line performance curves generated on F4.	126
4.14: A comparison of off-line performance curves generated on F5.	127
4.15: Off-line performance curves for genetic plans on F5.	129
4.16: Off-line performance for R4 on F5 as a function of mutation rate.	134
4.17: On-line performance for R4 on F5 as a function of mutation rate.	135
4.18: Allele loss for R5 on F1	139
4.19: Off-line performance for R5 on F5 as a function of the crowding factor.	140
4.20: On-line performance for R5 on F5 as a function of the crowding factor.	141

LIST OF ILLUSTRATIONS

<u>Figures</u>	<u>Page</u>
4.21: Off-line performance for R5 on F5 as a function of the crowding factor.	143
4.22: Off-line performance for R5 on F5 as a function of the crowding factor.	144
4.23: Loss probability curves for second order hyperplanes on chromosomes of length 30 as a function of the number of crossover points	152
4.24: Allele loss for R6 on F1 as a function of the number of crossover points	154
4.25: Off-line performance curves for R6 on F1 as a function of the number of crossover points	156
4.26: On-line performance curves for R6 on F1 as a function of the number of crossover points	157
5.1: Off-line performance curves for PRAXIS and DFP in local mode on F1.	170
5.2: On-line performance curves for PRAXIS and DFP in local mode on F1.	171
5.3: Off-line performance curves for PRAXIS and DFP in local mode on F2.	173
5.4: On-line performance curves for PRAXIS and DFP in local mode on F2.	174
5.5: Off-line performance curves for PRAXIS and DFP in local mode on F3.	175
5.6: On-line performance curves for PRAXIS and DFP in local mode on F3.	176
5.7: Off-line performance curves for PRAXIS and DFP in local mode on F4.	178
5.8: On-line performance curves for PRAXIS and DFP in local mode on F4.	179
5.9: Off-line performance curves for PRAXIS and DFP in local mode on F5.	180
5.10: On-line performance curves for PRAXIS and DFP in local mode on F5.	181
5.11: Off-line performance curves for PRAXIS and DFP in global mode on F3	183
5.12: Off-line performance curves for PRAXIS and DFP in global mode on F5	184
A.1a: Top surface defined by the 2-dimensional version of F1.	198
A.1b: Bottom surface defined by the 2-dimensional version of F1.	199
A.2a: Top surface defined by test function F2.	201
A.2b: Bottom surface defined by test function F2	202

LIST OF ILLUSTRATIONS

<u>Figures</u>	<u>Page</u>
A.3: Surface defined by the 2-dimensional version of test function F3	204
A.4a: Top surface defined by the 2-dimensional version of test function F4	205
A.4b: Bottom surface defined by the 2-dimensional version of test function F4	206
A.5a: Top surface defined by test function F5	209
A.5b: Bottom surface defined by test function F5.	210
B.1a: Off-line performance curves for random search on test function F1.	216
B.1b: On-line performance curves for random search on test function F1.	217
B.2a: Off-line performance curve for random search on test function F2.	218
B.2b: On-line performance curves for random search on test function F2.	220
B.3a: Off-line performance curves for random search on test function F3.	222
B.3b: On-line performance curves for random search on test function F3.	223
B.4a: Off-line performance curves for random search on test function F4.	225
B.4b: On-line performance curves for random search on test function F4.	227
B.5a: Off-line performance curves for random search on test function F5.	230
B.5b: On-line performance curve for random search on test function F5.	231
C.1a: Off-line performance curve for plan R1 on test function F1.	237
C.1b: On-line performance curve for plan R1 on test function F1.	238
C.2a: Off-line performance curve for plan R1 on test function F2.	240
C.2b: On-line performance curve for plan R1 on test function F2.	241
C.3a: Off-line performance curve for plan R1 on test function F3.	244
C.3b: On-line performance curve for plan R1 on test function F3.	245
C.4a: Off-line performance curve for plan R1 on test function F4.	247
C.4b: On-line performance curve for plan R1 on test function F4.	248
C.5a: Off-line performance curve for plan R1 on test function F5.	250

LIST OF ILLUSTRATIONS

<u>Figures</u>	<u>Page</u>
C.5b: On-line performance curve for plan R1 on test function F5.	251

ABSTRACT

AN ANALYSIS OF THE BEHAVIOR OF A CLASS OF GENETIC ADAPTIVE SYSTEMS

by

Kenneth Alan De Jong

Chairman: John H. Holland

This thesis is concerned with the design and analysis of adaptive systems, particularly in the area of adaptive computer software. To that end a formalism for the study of adaptive systems is introduced and, within this framework, a means of evaluating the performance of adaptive systems is defined. The central feature of the evaluation process is robustness: the ability of an adaptive system to rapidly respond to its environment over a broad range of situations. To provide a concrete measure of robustness, a family E of environmental response surfaces was carefully chosen to include a wide variety of surfaces, including multimodal and discontinuous ones. The performance of an adaptive system is evaluated over E by computer simulation by monitoring two distinct performance curves: on-line and off-line performance. With on-line performance every response of the adaptive system is evaluated, reflecting situations in which an adaptive system is used to dynamically improve the performance of a system. With off-line performance only responses which improve performance are evaluated, reflecting situations

in which testing can be done independently of the system being controlled.

Within this evaluation framework, a class of genetic adaptive systems is introduced for analysis and evaluation. These artificial genetic systems, called reproductive plans, generate adaptive responses by simulating the information processing achieved in natural systems by means of the mechanisms of heredity and evolution. This is accomplished internally by maintaining a population of individuals whose "genetic" material specifies a particular point on the response surface. New individuals (responses) are produced by simulating population development via mating rules, production of offspring, mixing of genetic material, and so on.

Even the most elementary genetic adaptive plan is shown to produce performance on E which is superior to pure random search of the response surfaces. However, these elementary genetic plans were shown to be easily affected by stochastic side-effects resulting from internal random processes. By suitable adjustments in parameters and modifications to the basic genetic plan, a considerable improvement in the performance was achieved on E.

As a final point of comparison, the performance of two standard function optimization techniques was evaluated on E. Their performance is shown to be superior on the continuous quadratic-like functions for which they were

designed. However, the genetic plans are shown to be superior on the discontinuous and multimodal surfaces, suggesting that genetic plans hold a valid position between specialized local adaptive techniques and pure random search.