Evolutionary Algorithms, Fitness Landscapes and Search

by

Terry Jones

M.Math., Computer Science, University of Waterloo, 1988.

B.Sc. (Hons), Computer Science, Sydney University, 1986.

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy in Computer Science

The University of New Mexico
Albuquerque, New Mexico

May, 1995

For me.

Acknowledgments

I have somehow managed to spend six years, in three universities, working towards a Ph.D. After a false start at the University of Waterloo, I enrolled at Indiana University and subsequently transferred to the University of New Mexico after leaving to do research at the Santa Fe Institute. In the course of these wanderings, I have been fortunate to interact with many people who have influenced me greatly. One of the pleasures of finally finishing is this opportunity to thank them.

At Waterloo, Charlie Colbourn was the first professor who regarded me as a friend rather than as a chore. Charlie inspired me to work on algorithms in a rigorous way, taught me what a proof is, and treated me as though what I thought was important. He apparently effortlessly supervised ten graduate students and produced many papers while always having time to talk, eat lunch, share a joke, and even help students move house.

Soon after arriving at Waterloo, I was adopted by Gregory Rawlins, who unexpectedly invited me to live at his apartment, after noticing that I was living in the Math building, using the suspended ceiling as a wardrobe. Gregory introduced me to posets and optimality in comparison-based algorithms. More than anyone, he has influenced the research directions I have taken over the years. He subsequently became my supervisor at Indiana and remained on my dissertation committee when I transferred to UNM. Gregory taught me to ask interesting questions, showed me what quality was, and encouraged my occasionally odd approaches to problem solving.

I was also greatly influenced by Ian Munro, who supervised my masters work and my first attempt at a Ph.D. Ian's ability to rapidly assess the worth of ideas and algorithms is amazing. I would spend a week or two on an approach to a problem and Ian could understand it, deconstruct it and tell me (correctly) that it wouldn't work in about a minute. On one memorable occasion, I went into his office with an algorithm I had devised and worked on furiously for at least two weeks. Ian listened for his customary minute, leaned back in his chair, and said "Do you want me to tell you smallest n for which it will not be optimal?"

Andrew Hensel, with whom I shared so much of my two and a half years at Waterloo, was the most original and creative person I have ever known well. Together, we dismantled the world and rebuilt it on our own crazy terms. We lived life at a million miles an hour and there was nothing like it. Five years ago, Andrew killed himself. There have been few days since then that I have not thought of him and the time we spent together.

At Indiana University, I was fortunate to be supervised by Douglas Hofstadter. Doug has both wide-ranging interests and the ability to think (and write!) clearly about his interests, in a highly original and probing way. Doug was very supportive of my extracurricular activities. When I began spending six hours a day juggling and unicycling; rather than telling me to do some work, Doug became the faculty advisor for the IU juggling club. Doug is an intelligent, kind, thoughtful and generous person. Perhaps most importantly, it was through Doug that I met Ana.

When I was offered a one year position as a graduate fellow at the Santa Fe Institute, I had barely heard of the place. Assured by many that this was an opportunity that shouldn't be missed, I left Indiana. After two days at SFI I had decided that I never wanted to leave. I have been extraordinarily lucky to have spent the last three years at SFI. There are so many wonderful things that could be said about SFI—the people, the lack of hierarchy, the open doors, the never-ending conversations, the atmosphere, the lack of red tape, the many workshops, countless interesting visitors and the diversity of interests. It is hard to imagine a more stimulating and encouraging academic environment. It will be difficult and sad to leave. Mike Simmons, the academic vice-president, is one of the many extremely competent people running SFI. The ongoing health of the institute is, I believe, largely due to Mike's vision of what SFI should be, to his refusal to allow that vision to be compromised, to his warm and friendly personality, and to his ability to deal calmly and rationally with any situation. He is fortunate to be surrounded by a team of people who combine to make SFI the unique place that it is.

At SFI I began to work with John Holland and Stephanie Forrest developing John's "Echo" system. My Echo experience has taught me many things about evolutionary biology, modeling and computer science. John always encouraged me to "follow my nose." When it led me away from Echo, he continued to support my investigations and take an active interest in them. His intuitions about Echo and genetic algorithms are extremely sharp. I have lost count of the number of times John suggested correct explanations for my observations and made accurate predictions about these systems.

After working on Echo with John and Steph for a year, I transferred schools, to the University of New Mexico. Steph became my "advisor" and I began work on fitness landscapes. Steph was my seventh advisor and, fortunately, she carried a large whip. Steph is probably the only one of the seven that could have made me finish. She has a refreshingly healthy, carefree, somewhat cynical view of the world and is great to laugh and joke with. Yet when she says "Jump!" I can only answer "How high?" This combination of fun and a kick in the pants once a week was the ideal foil to my recreational tendencies and limited attention span. It was Steph who insisted that I look for a connection between my landscapes and AI search techniques. I have come to respect and trust her judgement and to think of her as a good friend, collaborator and confidant.

Of the universities I have attended, UNM's computer science department is by far the smallest and least well funded. However, it has been at UNM though that I have finally found myself surrounded by other graduate students with whom I was inclined to interact academically. The Friday meetings of the "adaptive" group have been the highlight of my UNM time and I thank all the members of the group for keeping life interesting. In particular, I'd like to thank my evil twins, Derek Smith and Ron Hightower. Together we have spent many enjoyable hours rollerblading at high speed on campus and talking about our research. Derek in particular has always been very quick to understand my ideas about fitness landscapes. He has helped me to see where I was wrong and why I was occasionally right. He read this entire work and greatly helped with its content, style, and appearance. This dissertation would have been possible without Derek, but far less likely. Derek and Ron both have broad knowledge of computer science and possess high-speed, accurate, research quality detectors. They are full of interesting ideas and suggestions, and are models of academic generosity.

It is not possible to summarize Ana Mosterín Höpping and her influence on me in one paragraph, but I will try. Ana has completely changed my life over the last three years. Her love, intelligence, honesty, goodness, healthiness, humor, taste, liveliness and beauty have given me something to live for. She is the most balanced and well-adjusted person I have ever known. The fact of her existence is a continual miracle to me. She has supported me in hundreds of ways throughout the development and writing of this dissertation.

My parents are also very special people. My mother once told me that the last thing she and I had in common was an umbilical cord. Despite this, we are very close and I love them dearly. They have given their unconditional support, knowing that doing so contributed greatly to my absence these last nine years. They were strong enough to let me go easily, to believe in me, and to let slip away all those years during which we could have been geographically closer and undoubtedly driving each other crazy.

The other members of my dissertation committee, Ed Angel, Paul Helman, George Luger and Carla Wofsy were very helpful. They were all interested in what I was doing and improved the quality of what I eventually wrote. Ed's very healthy, even robust, cynicism and his willingness to give me a hard time at every opportunity, is the kind of attitude that I appreciate and enjoy the most. Paul made very detailed comments on all aspects of the dissertation, and tried to keep me honest with respect to connections to dynamic programming and branch and bound. George refused to let me get away with murder in the research hours I took with him, and made sure that I did a good job of understanding AI search algorithms. Though Carla was a late addition to the committee, she quickly and happily read the dissertation.

Joe Culberson, of the University of Alberta, is one of the few people who appreciated and fully understood what I was trying to achieve in this dissertation. Joe and I share similar views on many aspects of fitness landscapes. I had known Joe for years before I was pleasantly surprised to discover that we had independently conceived of a crossover landscape in essentially the same way. We have exchanged many ideas about landscapes for the last two years, keeping careful track of each other's sanity. My thinking and writing are clearer as a result of Joe's rigor and insistence that I make myself clear to him.

At SFI I have profited from many discussions with Melanie Mitchell, Chris Langton, Richard Palmer, Una-May OReilly, Rajarshi Das, Michael Cohen, Aviv Bergman, Peter Stadler, Walter Fontana, Stu Kauffman, and Wim Hordijk (who also read and commented on the first two chapters of this dissertation). I have made friends and received encouragement from many people in the research community, including David Ackley, Ken De Jong, David Fogel, Jeff Horn, Rich Korf, Nils Nilsson, Nick Radcliffe, Rick Riolo, Günter Wagner and Darrell Whitley. Jesús Mosterín read my thesis proposal, cheerfully scribbled "nonsense!" across it, and then sent me a useful page of formal definitions. Many members of the staff at SFI, especially Ginger Richardson, have become good friends. The CS department at UNM has been wonderful; I have been particularly helped by Ed Angel, Joann Buehler, Jim Herbeck and Jim Hollan.

Finally, I have made many friends along the way. They have helped me, one way or another, in my struggle to complete a Ph.D. Many thanks to Sandy Amass, Marco Ariano, Marcella Austin, Amy Barley, Greg Basford, Dexter Bradshaw, Ted Bucklin, Peter Buhr, Susy Deck, Emily Dickinson, Elizabeth Dunn, Beth Filliman, Bob French, Julie Frieder, Elizabeth Gonzalez, Fritz Grobe, Steve Hayman, Ursula Höpping, Helga Keller, The King, Jim Marshall, Gary McGraw, Heather Meek, Eric Neufeld, Luke OConnor, Lisa Ragatz, Steven Ragatz, Kate Ryan, Philip San Miguel, John Sellens, Francesca Shrady, Lisa Thomas, Andre Trudel and the one and only Françoise Van Gastel.

Evolutionary Algorithms, Fitness Landscapes and Search

by

Terry Jones

B.Sc. (Hons), Computer Science, Sydney University, 1986
M.Math., Computer Science, University of Waterloo, 1988
Ph.D., Computer Science, University of New Mexico, 1995

ABSTRACT

A new model of fitness landscapes suitable for the consideration of evolutionary and other search algorithms is developed and its consequences are investigated. Answers to the questions "What is a landscape?" "Are landscapes useful?" and "What makes a landscape difficult to search?" are provided. The model makes it possible to construct landscapes for algorithms that employ multiple operators, including operators that act on or produce multiple individuals. It also incorporates operator transition probabilities. The consequences of adopting the model include a "one operator, one landscape" view of algorithms that search with multiple operators.

An investigation into crossover landscapes and hillclimbing algorithms on them illustrates the dual role played by crossover in genetic algorithms. This leads to the "headless chicken" test for the usefulness of crossover to a given genetic algorithm and to serious questions about the usefulness of maintaining a population. A "reverse hillclimbing" algorithm is presented that allows the determination of details of the basin of attraction of points on a landscape. These details can be used to directly compare members of a class of hillclimbing algorithms and to accurately predict how long a particular hillclimber will take to discover a given point.

A connection between evolutionary algorithms and the heuristic search algorithms of Artificial Intelligence and Operations Research is established. One aspect of this correspondence is investigated in detail: the relationship between fitness functions and heuristic functions. By considering how closely fitness functions approximate the ideal for heuristic functions, a measure of search difficulty is obtained. This measure, fitness distance correlation, is a remarkably reliable indicator of problem difficulty for a genetic algorithm on many problems taken from the genetic algorithms literature, even though the measure incorporates no knowledge of the operation of a genetic algorithm. This leads to one answer to the question "What makes a problem hard (or easy) for a genetic algorithm?" The answer is perfectly in keeping with what has been well known in Artificial Intelligence for over thirty years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

ABBREVIATIONS

The following abbreviations are used in this dissertation:

§x(y)	Section x on page y.
β	The bit-flipping operator. §2.6.1(31)
μ	A generic mutation operator. §2.6.2(32)
ϕ	A generic operator. §2.3(24)
ψ	A generic operator. §2.3(24)
$\phi_{k \to l}$	Class of fixed cardinality operators with cardinality k and l. §2.5.9(29)
χ	A generic crossover operator.
$\chi_n^{k \to l}$	n-point crossover producing l offspring from k parents. §2.6.3(34)
$\chi_u^{k \to l}$	Uniform crossover producing l offspring from k parents. §2.6.3(34)
AI	Artificial Intelligence.
$B_\phi(v)$	The basin of attraction of a point v under an operator ϕ. §2.5.8(29)
BH	Bit-flipping Hillclimbing. §3.6(55)
CH	Crossover Hillclimbing. §3.7(57)
EC	Evolutionary Computation. §1.1(4)
EP	Evolutionary Programming. §1.1(4)
ES	Evolution Strategy. §1.1(4)
$\mathcal{F}$	Fitness space. §2.3(24)
GA	Genetic Algorithm. §1.1(4)
GP	Genetic Programming. §1.1(4)
$G_\mathcal{L}$	The graph corresponding to a landscape $\mathcal{L}$. §2.4(26)
$\mathcal{L}$	A landscape. §2.4(25)
LSA	Local search algorithm. §4.2(86)
$\mathcal{M}(S)$	The set of all multisets whose elements are drawn from a set S. §2.2.1(18)
$\mathcal{M}_q(S)$	Those elements of $\mathcal{M}(S)$ with cardinality q. §2.2.1(18)
$N_\phi(v)$	The neighborhood of a vertex v under the operator ϕ. §2.3(24)
$\mathcal{O}$	Object space. §2.3(20)
OR	Operations Research.
$>_S$	A partial order on a set S. §2.2.2(18)
P	Probability.

$\mathcal{R}$	Representation space. §2.3(20)
TM	Turing Machine.
TSP	Traveling Salesperson problem.
$u(x)$	The number of ones in a binary string x.