Evaluating and Improving Steady State Evolutionary Algorithms on Constraint Satisfaction Problems

Koen van der Hauw

9 August 1996

Abstract

Currently there is a growing interest in the evolutionary algorithm paradigm, as it promises a robust and general search technique. Still, in spite of much research, for many people the question remains how good evolutionary algorithms really are. Therefore, in this research, a successful class of evolutionary algorithms, Steady State evolutionary algorithms, is thoroughly examined to find optimal settings on two NP-complete problems: Graph 3-Coloring and 3-Satisfiability. Several versions of the evolutionary algorithm are tested and evaluated and the best version for each NP-complete problem is compared to a good existing algorithm for each problem. Then extensions for the evolutionary algorithm are presented that make the evolutionary algorithms perform better than the more traditional algorithms on the hardest problem instances.

Preface

This research was done as a Master's Thesis for graduating in Computer Science at Leiden University. It is about evolutionary algorithms, search algorithms based on Darwin's evolution theory, and was motivated by my own interest in biology and evolution and my curiosity about its effectiveness in a simulation on a computer. As seeing is believing and because of the opportunity to do this research at the Department of Computer Science in Leiden, I decided to test a class of evolutionary algorithms on some very hard problems to partially answer at least for myself the question about the usefulness of evolutionary algorithms.

I would like to thank my supervisor, Guszti Eiben, for his support and for the freedom he gave me in doing this research. Further thanks to Han La Poutré, my second supervisor, for his corrections and to Zsofi Ruttkay for some suggestions. Lastly I would like thank all the people whose computer I used for the many time consuming experiments that needed to be done and who possibly got annoyed with it.

Contents

Chapter 1

Introduction

This Master's Thesis is about evolutionary algorithms, a class of search algorithms based on biological evolution. Because evolutionary algorithms are supposed to be robust and general search algorithms, there has been a growing interest in evolutionary algorithms during the last decades. Especially the last few years, they are becoming known to a wider public in the computer science community. Still, for many people, the question remains how good evolutionary algorithms are compared to existing algorithms. To get an idea about the answer, in this research two difficult problems are used as a testcase to compare existing algorithms for these problems with an evolutionary algorithm.

We restricts ourselves to the class of Constraint Satisfaction Problems (CSPs), that, simply stated, consist of a number of variables and a number of constraints on the values of these variables. Therefore the conclusions in this research are really only valid for CSPs. The two CSPs that are examined in this research are: Graph 3-Coloring and 3-Satisfiability. Both are known to be NP-complete problems and so, unless P=NP, no complete algorithms with polynomial time complexity are known to exist. In other words no efficient algorithms are known for these hard problems and therefore they are interesting as a testcase. The Graph 3-Coloring problem will be the main problem and the 3-Satisfiability problem will be used to verify the main conclusions.

The objective of this thesis is first to find a good setting for the parameters in the evolutionary algorithm, so that a reasonably optimized version is obtained. This version will be compared with a known, good algorithm to get an idea of the power and use of evolutionary algorithms. Then, when the evolutionary algorithm turns out to be inferior to the existing method, ideas will be suggested and tested to possibly improve the performance of the evolutionary algorithm. Throughout this research, as little domain knowledge as possible will be used, as a main advantage of evolutionary algorithms is their general applicability.

The next chapter is an introduction to the principles of evolutionary algorithms. Chapter 3 describes the class of evolutionary algorithms that is used in this report in detail. In chapter 4, the measures for comparing the algorithms will be defined and discussed. In chapter 5, the Graph 3-Coloring problem and existing algorithms will be described. Chapter 6 is the most important chapter in this report. In it the evolutionary algorithm is optimized, compared to the existing algorithms and some improvements are suggested and tested. In Chapter 7, the main results of the previous chapter will be verified on 3-Satisfiability. Finally, chapter 8 gives a summary of this research, a discussion of the main conclusions and some ideas for further research. In appendix A, a summary of the abbreviations and some notation is given.

Chapter 2

What are Evolutionary Algorithms?

Evolutionary algorithms (EAs) are search algorithms based on biological evolution. Darwin was the first to clearly state the idea of *natural selection* as a principle behind biological evolution [15]. Natural selection or "survival of the fittest"[1] is simply the process by which the best or most fit members of some species have more probability of surviving or by which the weakest members have more probability of dying out so that the weakest individuals will be selected out. As fit members live longer, they will have more offspring[2] so that on average the whole species will become fitter.

Sometimes people who do not want to accept evolution theory, say that natural selection is a purely negative force that just weeds out failures, but is not capable of building up the complexity that we see around us everywhere. They argue that it only subtracts from what is already there and that a more creative "designer" is needed that adds something as well. In Darwin's theory, *mutation* is the force that can add and all it needs is enough *time* to stumble upon the right mutations. Mutations, induced by for example radiation, can change an organism's *genes*, that form the *instructions* for building and regulating an organism and are found in each cell of the organism. By mutating the genes of new offspring, the way the offspring looks, behaves, et cetera can be changed. It should be realized that a small change in genetic material (genotype) can have great effect on the resulting organism (phenotype). So again, all that is needed is enough time for mutation to add to the complexity of organisms. Because our brains are built (by evolution itself) to deal with events on radically different timescales (hours, years or perhaps decades) from those that evolution had available (millions of decades), our intuitive judgments of what is probable turn out to be wrong by many orders of magnitude. Still, however much time we have, it seems impossible to maintain that, for example, an eye could come to existence in one step just by mutation. The answer to this is that evolution is a *gradual* process that cumulatively builds up complex designs from small changes that come from mutations. Of

[1] When a fit organism is defined as an organism which is good at surviving, we seem to get a trivial remark: survival of the organisms that are best at surviving. But maybe as it is so trivial, there's no way of denying it.

[2] In nature often the stronger individuals also have more probability of mating and because of this will create even more offspring.

course each of these small changes should have enough advantage to surviving or it would be selected out again. One could ask what the advantage is of half an eye as the eye is so complex that it does not work properly anymore if changed too much. The advantage of half an eye is simply that it's better than seeing nothing and even one lightcell could be an advantage. In [19], Dawkins very clearly and gradually explains how evolution can build up complex organisms with mutation, natural selection and enough time.

An example, drawn from [57], should clarify the basic ideas from evolution theory. Let's suppose that somewhere and sometime we have a population of cute, furry, little rabbits. Now some of them will be faster and smarter than other rabbits. These faster, smarter rabbits are less likely to be eaten by foxes (we assume a population of foxes as well) and therefore more of them survive to do what rabbits do best: make more rabbits. Of course, some of the slower, dumber rabbits will survive just because they are lucky. This surviving population starts breeding which results in a good mixture of rabbit characteristics: some slow rabbits breed with fast rabbits, some fast with fast, some smart with fast rabbits, and so on. And sometimes, nature throws in a "wild hare" by mutating some of the rabbit genetic material so that perhaps an exceptional smart, dumb or fast rabbit is created. The resulting baby rabbits will (on average) be faster and smarter than the ones in the original population, because more faster, smarter parents survived the foxes. (Fortunately for the foxes, the foxes are undergoing a similar process, otherwise the rabbits might become too fast and smart for the foxes to catch any of them).

Evolutionary algorithms (EAs) are inspired by natural evolution, though simplificate on many issues. Contrary to natural evolution, EAs are *goal-driven*: they are trying to evolve to a prespecified goal. The goal can be the finding of solutions for any problem that needs a solution and the EA maintains a *population* of potential solutions for this problem. In each member or *individual* the necessary information is stored to represent these partial solutions. The information is stored in *genes* and the individuals solely consist of genes, each gene representing a characteristic of the solution. A user defined *fitness function* is used for natural selection and for each individual gives a calculated *fitness* value that specifies how good the individual is. Now breeding is simulated by selecting *parents* from the existing population, called the current *generation*, that are allowed to mate and produce offspring (that are again potential solutions). Mating is done by selecting two parents, that produce one or more children by *crossing over* their genes so that information from both parents is combined in the *children*. The children are *mutated* by random changes in some of their genes to bring in new *diversity* in the population. Often a fixed population size is used and from the original population and offspring population, individuals have to be selected to form the next generation. The individuals are selected by favoring the fitter individuals and so giving them higher probability for surviving. In some EAs, individuals from one generation will not be allowed to exist in the next generation (similarly to many low organisms which reproduce only once). In other EAs, individuals from one generation are also allowed to exist in the next generation (similarly to many higher organisms which can mate several times). Because the fitter individuals are favored, the average population will become fitter and fitter as the generations go by so that the potential solutions will become better and better, approaching the final solution closer and closer.