HIGHLY PARALLEL EVOLUTIONARY ALGORITHMS FOR GLOBAL OPTIMIZATION, SYMBOLIC INFERENCE AND NON-LINEAR REGRESSION

S. Arnone, M. Dell'Orto and A. Tettamanzi*

Dipartimento di Scienze dell'Informazione, Universitá degli Studi di Milano, Italy

M. Tomassini

Centro Svizzero di Calcolo Scientifico, Via Cantonale, 6928 Manno, Switzerland

ABSTRACT

This work describes massively parallel genetic algorithms as a means for solving difficult global optimization, symbolic expression inference and curve fitting problems. Results are shown for hard test cases and the implications for computational physics problems are discussed. The approach, eminently parallel, is believed to be competitive with more familiar methodologies used by computational scientists

1. Introduction

The name *Evolutionary Algorithm* encompasses a family of stochastic optimization techniques based on the key concept of *evolution*. The first proposals in this direction date back to the mid sixties, when John Holland, of the University of Michigan, introduced Genetic Algorithms and Ingo Rechenberg, of the Technical University of Berlin, independently began to work on Evolution Strategies. Their pioneering works eventually gave rise to a broad class of optimization methods particularly well suited for hard problems where little is known about the underlying search space. The last development of this research thread is the so called Genetic Programming, introduced by John Koza [1]. A recent work of reference and synthesis in the field of Evolutionary Algorithms is [2].

An Evolutionary Algorithm mantains a population of candidate solutions for the problem at hand, and makes it evolve by iteratively applying a (usually quite small) set of stochastic operators, known as *mutation*, *recombination*, *reproduction* and *selection*.

Mutation randomly perturbs a candidate solution, recombination decomposes two distinct solutions and then randomly mixes their parts to form a novel solution, reproduction replicates the most successful solutions found in a population at a rate proportional to their relative goodness, while selection purges poor solutions from a population.

The initial population may be either a random sample of the solution space or may be seeded with solutions found by simple local search procedures, if these are available.

The resulting process tends to find globally optimal solutions to the problem much in the same way as in Nature populations of organisms tend to adapt to their surrounding environment.

*Part of this research has been supported by Dipartimento di Scienze dell'Informazione, under the supervision of Prof. G. Degli Antoni.

2. Parallel Cellular Evolutionary Algorithms for Global Optimization and Symbolic Inference of Formulae

There is great scope for parallelism in genetic algorithms as they are naturally suited for both coarse-grain and fine-grain parallel implementation. Some experiments showed that good results can be achieved by providing populations with a spatial structure and basing the evolutionary process on local interactions between group of individuals (the so-called *island* model or between single individuals. A convenient algorithm architecture for the latter model makes use of a so-called *apomictic* topology, implemented by distributing individuals on a planar grid, where genetic interactions are restricted to the immediate neighbors of each individual. Since each individual can evolve asinchronously and indipendently of the others, a massively parallel implementation follows straightly; for example a virtual (or physical) processor could be assigned to each of them, thus letting the evolution go on in the fastest possible way. This is particulary important because usually the evaluation of a candidate solution, that each individual rappresents, is a time-consuming task. In some cases locality of selection and mating can lead to premature convergence of the population towards one or a few local extrema. To compensate for this effect, we may introduce a diffusion-like process in which individuals periodically take a random walk to another grid point where the highest scoring individual survives and may reproduce. Mutation is applied in the usual way to help jumping out of local minima by randomizing the system.

In what follows two examples of the usefulness of this approach will be described. These are global function optimization and the symbolic inference of formulae. All the models have been implemented and tested on a Connection Machine CM-200. These grid models map nearly ideally on the CM-200 due to the fact that there is a one-to-one correspondence between processor grid topologies and population topologies and to the existence of very efficient nearest neighbour communications.

The global function minimization problem can be stated as follows. Given a function $f(x)$:

$$f : R^n \mapsto R$$

and a set $D \in R^n$, find x^* such that:

$$f(x^*) = \min\{f(x) \mid \forall x \in D\}$$

where $x = (x_1, x_2, \ldots, x_n)^T$.

Genetic algorithms have been quite successful on this kind of problems, especially when standard mathematical methods are hard to apply e.g., for noisy, discontinuos or nondifferentiable functions.

Each of the n coordinates is represented by a binary string. The concatenation of the n coordinate substrings form a chromosome that is the representation of a particular point $x \in D$. The length of the bit-string depends on the required precision. This is a standard encoding for real quantities in GAs [2] however, a straight floating-point encoding can also be used and is very advantageous in terms of precision and efficiency [2]. However, it requires that the semantics of genetic operator be adapted to the new representation.

To test the cellular parallel GA models we used, among others, functions f_6, f_7 and f_8 from [3]. These are hard multimodal test functions for which the location of the global minimum is known and they are defined as follows:

$$f_6(x) = \sum_{i=1}^{6}[x_i^2 - \cos(18x_i)],$$

$$f_7(x) = \sum_{i=1}^{10} -x_i \sin(\sqrt{\mid x_i \mid}),$$

$$f_8(x) = \sum_{i=1}^{10} x_i^2/4000 - \prod_{i=1}^{10} \cos(x_i/\sqrt{i}) + 1,$$

With $x_i \in [-5.12, 5.12]$ for f_6, $x_i \in [-500, 500]$ for f_7 and $x_i \in [-600, 600]$ for f_8. Function f_8 is particularly hard. The global minimum $f = 0$ is at $x_i = 0, i = 1, \ldots, n$. In the ten-dimensional case there are four suboptimal minima with $f(x) \approx 0.0074$ at $x \approx (\pm\pi, \pm\pi\sqrt{2}, 0, \ldots, 0)$.

The results, averaged over 50 runs for each problem with a population size of 8K show that the percentage of success is very high. If periodic random migration of individuals is allowed the success rate is 100% for all the test cases. A fuller account is given in [4].

There are many possible applications of these techniques to computational physics problems. Free energy minimization in proteins and crystals are well-known cases. We are experimenting with Lennard-Jones clusters of atoms with encouraging results. The method does not employ any knowledge about the solution to the problem, contrary to what is usually the case in many large scale molecular configuration problems. However, knowledge can be added when available by seeding the initial population with known sub-optimal structures or fragments, instead of starting with a randomly chosen population.

Symbolic regression, linear and non-linear interpolation and more generally curve fitting problems are of great interest both from a mathematical and a physical point of view. Furthermore, due to the vastness of the search spaces involved, they seem to be particulary suited for an evolutionary approach. The above described parallel approach seems to be particulary suited for smoothing problems that, involving large amounts of data, need large computational capabilities.

An interesting example of similar problems is the symbolic inference of mathematical functions, that can be formulated as follows: Given an objective function, whose values are known in a set of points, we have to find its symbolic expression. A particular case of the latter problem, currently under study by the authors, is the induction of the closed form of some special case of recurrence equation of one variable. The equations considered are of the form:

$$f(i) = f(i-1) + h(i) \quad i = 0, 1, \ldots, \tag{1}$$

where h is known and f is the unknown function to infer. Equation **??** can be rewritten as

$$f(n) = \sum_{i=1}^{n} h(i);$$

then suppose that $h(i)$ is of the form i^k, $k = 1, 2, \ldots$: in this case we obtain the family of functions

$$f_k(n) = \sum_{i=1}^{n} i^k. \tag{2}$$

In a genetic approach to this kind of problem, each genotype corresponds to a symbolic expression represented in inverse polish notation with a string of characters containing operators and operands. The evolution takes place by means of a simulated "natural" selection such that each individual has a probability to reproduce proportional to its fitness. Thus genotypes that correspond to a poor formula are replaced by better ones, as generations go by. New genotypes are generated when existing ones undergo crossover and mutation. More details of this approach are given in [5]. Some of the results obtained follow:

$$f_1(n) = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

$$f_2(n) = \sum_{i=1}^{n} i^2 = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}$$

$$f_3(n) = \sum_{i=1}^{n} i^3 = \frac{n^4}{4} + \frac{n^3}{2} + \frac{n^2}{4}$$

$$f_4(n) = \sum_{i=1}^{n} i^4 = \frac{n^5}{5} + \frac{n^4}{2} + \frac{n^3}{3} - \frac{n}{30}$$

$$f_5(n) = \sum_{i=1}^{n} i^5 = \frac{n^6}{6} + \frac{n^5}{2} + \frac{5n^4}{12} - \frac{n^2}{12}$$

3. Conclusions

In the last few years the field of *evolutionary algorithms* has been steadily growing. Today they are currently used as a replacement or together with existing methods for a wide class of difficult optimization problems both in the academic and the industrial world. According to our own experience and that of other groups, parallel evolutionary algorithms in particular are very effective for both numerical and symbolical optimization problems as they can cut down significantly on computing time while offering a natural and easy approach to parallelism. Nevertheless, these methods are still little used in computational science work although they are certainly competitive with simulated annealing and other heuristics for the same tasks and can be more advantageously parallelized thus allowing to tackle more difficult problems.

Acknowledgments

We would like to thank Prof. G. Degli Antoni for encouragement and helpful advice and Drs. B. Chopard of the University of Geneva and P. Rossi of CRS4, Cagliari, for kindly making available to us the CM-200 environments at their centers.

References

1. J.R. Koza, *Genetic Programming: on the programming of computers by means of natural selection*, The MIT Press, Cambridge, Massachussets, 1993.
2. Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag, Berlin, 1992.
3. A. Törn and A. Žilinskas, *Global Optimization*, Springer-Verlag, 1989.
4. M. Tomassini, *The Parallel Genetic Cellular Automata: Application to Global Function Optimization*, Proceedings of the International Conference on Artificial Neural Networks and Genetic Algorithms, Springer-Verlag, Wien, 385, 1993.
5. S. Arnone, M. Dell'Orto and A. Tettamanzi, *A Massively Parallel Evolutionary Algorithm for the Symbolic Inference of Inductive Formulae*, Submitted to ESA '94.