STUDIA UNIV. BABEŞ–BOLYAI, INFORMATICA, Volume **XLV**, Number 1, 2000

GENETIC CHROMODYNAMICS

D. DUMITRESCU

ABSTRACT. A new evolutionary search and optimization metaheuristics, called *Genetic Chromodynamics* (**GC**), is proposed. The **GC**-based methods use a variable sized solution population and a local interaction principle. Local interactions induce a restricted mating scheme and permit detection of multiple optimal solutions.

The main idea of the **GC** strategy is to force the formation and maintenance of stable sub-populations. Proposed local interaction scheme ensures sub-population stabilization in the early search stages.

Sub-populations co-evolve and eventually converge towards several optimal solutions. The number of individuals in the current population decreases with the generation. Very close individuals are merged. At convergence the number of sub-populations equals the number of optimal solutions. Each final sub-population hopefully contains a single individual representing an optimum point (a solution of the problem).

The GC approach allows as solution representation any data structure compatible with the problem and any set of meaningful variation operators.

GC-based techniques can be used to solve multimodal, static and dynamic, optimization problems.

Keywords: Evolutionary algorithms, Genetic chromodynamics, Multimodal optimization

1. Introduction

Evolutionary computing (EC) deals with adaptive search and optimization techniques that simulate biological evolution and adaptation processes. EC mainly includes Genetic algorithms (GAs), Evolution strategies, Evolutionary programming and Genetic programming. Genetic algorithms represent the most typical instance of EC (see [4]).

Unfortunately standard GAs can not solve all kinds of optimization and search problems, like GA – hard or deceptive problems [8]. While one of the main difficulties arises from the premature local convergence, other difficulties concern the multimodal optimization problems. Standard GAs, as well as the other usual evolutionary procedures, generally fail to detect multiple optimum points.

2000 *Mathematics Subject Classification.* 68T05.

1998 *CR Categories and Descriptors.* I.2.8 [**Computing Methodologies**]: Artificial Intelligence – *Problem Solving, Control Methods, and Search.*

Several methods have been proposed to solve premature convergence and multimodal optimization problems.

Virus-evolutionary genetic algorithm (VEGA) [10] has been considered to prevent the premature local convergence due to the lack of diversity in the solution population. The VEGA approach is based on the *virus theory of evolution* according to which viruses transport segments of DNA across the species. The VEGA approach implies two populations: a *host population* and a *virus population*. The virus population realizes a horizontal propagation of genetic information in the host population. This propagation is realized by *virus infection*, i.e. by caring solution fragments (substrings) between the individuals in the host population. Therefore the VEGA technique simulates evolution with horizontal propagation and vertical (i.e. usual) inheritance of genetic information.

Multimodal genetic algorithms generally use another biological idea, namely the *niche* concept [1, 2, 6, 7, 8, 9]. Each optimum region in the search space will be considered as a niche. Niching genetic algorithms are able to form and maintain multiple, diverse, optimal solutions. Usually the niche concept is implemented through the use of fitness sharing. The niching process is accomplished by degrading the fitness of an individual according to the presence of nearby individuals.

The *sharing functions* [6] are used to calculate the extent of sharing to be performed between two individuals. For each individual the value of the sharing function is calculated with respect to the individuals in the population. The niche count of an individual is the sum of the corresponding sharing values. The fitness of an individual is divided by its niche count. The obtained updated value is the *shared fitness* of that individual.

The radius s of the estimated niches is considered. The individuals separated by distance greater than s do not degrade each other's fitness.

Sharing tends to spread the population over different optima proportionally to the values of these optima. Unfortunately the niching methods do not guarantee an appropriate selection of all useful solutions, for any situation [8]. The detection of the number of optimal solutions could be a problem, as well.

In this paper we consider a different, non-niching, strategy to prevent premature local convergence and to detect multiple optimal solutions. The proposed approach is called *Genetic Chromodynamics* (**GC**). Let us note that **GC** does not represent a particular evolutionary technique but merely a metaheuristics for solving (multimodal) optimization/search problems.

2. Genetic Chromodynamics principles

Genetic Chromodynamics metaheuristics uses a variable sized population of solutions (chromosomes or individuals) and a local mating scheme. Several solution representations can be considered. For instance solutions may be represented as real- component vectors. Solution representation as binary strings can be also used. Proposed **GC** strategy allows any data structure suitable for a problem together with any set of meaningful variation/search operators. Moreover the proposed approach is independent of the solution representation.

The main idea of the **GC** strategy is to force the formation and maintenance of sub-populations of solutions. Sub-populations co-evolve and eventually converge towards several (local and global) optimal solutions.

The number of individuals in the population decreases with the generation. Very similar individuals (solutions) are merged. At convergence the number of sub-populations equals the number of optimal (local and global) solutions. In the standard case each final sub-population contains a single individual representing an optimum point (a solution of the problem).

A different color is assigned to each solution in the initial population. In the standard **GC** approach every solution in each generation is selected for recombination or mutation. The recombination mate of a given solution is selected within a determined *mating region*. Before recombination all solutions in a given mating region will receive the color of the best individual within that region.

A (2,1) recombination mechanism is used. The first parent is dominant and the second one is recessive. The unique offspring is labeled as the descendent of its dominant parent. The offspring will inherit its parent color. It is expected that at convergence only different colored solutions will remain in the population.

Two sub-populations will generally have different color sets. The number of colors per region tends to decrease with the time. Hopefully a dominant color will be established in each sub-population.

We may consider that the method encounters two interacting dynamics: a micro and a macro dynamics. The system micro-dynamics is associated with solution modifications. The macro-dynamics is associated with sub-populations formation, modification and stabilization. Macro-dynamics induces a dominant color within each sub-population.

We may consider each of the micro and macro-dynamics as expressing a particular aspect of the global system dynamics.

As **GC** strategy uses a variable-sized solution population, the underlying population dynamics is more complicated than in usual evolutionary algorithms. Therefore the corresponding search process may also be supposed to be more powerful. This feature makes **GC**-based searching methods appealing for solving difficult tasks, like time-dependent, multimodal and multiobjective optimization problems. The macro-dynamics of a variable sized population seems to be adequate to deal with a changing environment. Hence the proposed approach is potential useful for tackling distributed AI applications, like cooperative multi-agents.

3. Subpopulation emergence

GC-based optimization techniques start with a large arbitrary population of solutions. Dimension of the solution population decreases at each generation. There is a highly probability that each new generation will contain some individuals better than the individuals in the previous generation.

Using a local mating scheme the formation and maintenance of solution subpopulation is favored or even forced.

Sub-populations evolve towards compact and well separated solution clusters. Sub- populations within each generation $P(t)$ induce a hard partition (or at least a cover) of the set $P(t)$.

In defining sub-populations we may consider a biological point of view. Recombination of individuals in the same sub-population is highly expected. The probability of mating individuals belonging to the same sub-population is greater than the probability of mating individuals from different sub-populations. Recombining individuals in different sub-populations is not definitively forbidden but usually it is very improbable. Therefore we may say that sub-populations are composed of highly compatible (with respect to recombination) individuals.

GC approach is essentially based on local interactions in a variable-sized population. The role of local mating scheme (local solution interactions) and that of variable sized population may be summarized as follows:

(i) to ensure early sub-population formation and stabilization;
(ii) to avoid massive migration between sub-populations approximating different optimum points (migrations could affect the quality of some already obtained 'pure' or high quality solutions);
(iii) to prevent destruction of some useful (high quality) sub-populations;
(iv) to ensure a high probability of obtaining all useful problem solutions.

Local interaction principle needs a slight modification of the variation search operators.

4. Mating region

Let us consider a distance concept (a metric or a pseudo-metric) δ defined on the solution space Y. Consider an initial population in which each solution has a different color. Let f be the fitness function. As usual $f(c)$ evaluates the quality of the solution c.

As only short range interactions between solutions are allowed, the mate of each solution c has to belong to a neighborhood of c. It is usually convenient to consider this neighborhood as the closed ball $V(c, r)$ of center c and radius r.

We may interpret the parameter r as the *interaction radius* (or *interaction range*) of the individual c. Short range (or local) interactions will ensure an appropriate co-evolution of the sub-populations.

All the individuals within the region $V(c,r)$ receive the color of the best individual in that mating region. The search process starts with a population whose individuals have all different colors.

An adaptation mechanism can be used to control the interaction range r, so as to support sub-population stabilization. Within this adaptation mechanism the interaction radius of each individual could be different. In this way the flexibility of the search process may increase significantly. Each sub-population may have a more independent evolution (more freedom degrees of its behavior). To control the domain interactions we can use a general (problem independent) method or a particular heuristic. Problem dependent approaches seem to be appealing for dealing with some particular situations.

Let us note that the meaning of the mating region $V(c,r)$ is not that of a niche. The resources of this region are not shared between its members, as in the niche approach. It is more suitable to interpret the *mating region* $V(c,r)$ as the *interaction domain* of the individual c.

For some particular problems we may admit migrations (meaning that recombination is permitted) between different interaction domains. Allowing permeable interaction domains may lead to better solutions by increasing population diversity.

5. Termination condition

Various termination conditions for the **GC** search process may be identified. Some stopping conditions may be formulated according to the particular problem considered. Other stop conditions are problem independent. Here we are interested in the second class.

A good, general, problem independent heuristics is to stop the search process if the solution population remains unchanged for a fixed number of generations. This condition represents a natural termination criterion ensuring that the search process continues only how long is necessary.

6. Selection and recombination

Genetic chromodynamics involves two types of selection schemes. Global selection supplies the parent population. Local selection is a mechanism for choosing a mate of a solution in the respective mating region.

6.1. Global selection.

Within standard version of Genetic chromodynamics approach each solution c in the population $P(t)$ will be considered for recombination. More sophisticated global selection mechanisms may be used. Their efficiency in this context is questionable.

6.2. **Local selection.** According to the proposed local interaction scheme the mate of the solution c will be chosen from the neighborhood (mating region) $V(c,r)$ of c. Local mate selection is done according to the values of the fitness function f.

For selecting the mate of a given solution we may use proportional selection. Let m be a solution in the interaction domain (mating region) $V(c,r)$ of the solution c. The probability that m is selected as the mate of c is denoted by $p(m)$ and is defined as

$$p(m) = \frac{f(m)}{\sum_{a \in V(c,r)} f(a)}.$$

Any other type of selection compatible with the particular considered problem is permitted. Tournament selection is a very powerful mechanism and may be successfully use for mate selection.

6.3. **Recombination.** Let a be the selected partner of c. The ordered pair (c,a) generates by recombination a unique offspring. The first parent is dominant, whereas the second one is recessive.

Let d be the offspring generated by c and a. The offspring d will inherit the color of its (dominant) parent c and will be labeled as the descendent of c only.

The form of the recombination operator will be chosen according to the solution representation and the nature of the problem.

For a real valued solution representation a convex combination of the genes in c and a can be used to obtain the components of d. In the case of convex recombination the i^{th} position of the offspring d has the expression:

$$d_i = qc_i + (1-q)a_i,$$

where q is a real number in the unit interval [0,1].

7. Mutation operator

If the closed ball $V(c,r)$ – the interaction domain of c – is empty then the solution c will be selected for mutation. In this way recombination and mutation are mutually exclusive operators. Mutation may be considered as acting mainly on stray points.

An additive normal perturbation seems to be appropriate for general optimization purposes. By mutation (stray) solutions are usually drawn closer to local optimum points of the objective function. As a side effect solutions are forced towards one of the existing sub-populations.

Various solution components may suffer perturbation with different standard deviation values. In every situation the mutated solution will inherit the color of its parent, as well.

8. Mutation acceptance scheme

Within usual evolutionary algorithms generated mutations are generally unconditionally accepted. Within Genetic chromodynamics based techniques a more sophisticated acceptation mechanism will be considered.

8.1. **General acceptance mechanism.** Standard **GC** approach considers that in each generation every solution is involved in recombination or mutation. Each solution will produce, and possibly be replaced by, an offspring. Whichever is better between a dominant parent and its offspring will be included in the new generation.

According to the proposed mechanism a mutated solution (offspring), which is better than its parent, is unconditionally accepted. This acceptance scheme induces a rapid convergence of the search process.

It seems that no restriction on mutation parameter is needed if the best from parent and offspring survives. This strategy can be useful in the first stages of the search process. In the last stages it may cause a drawback of the search process.

Let us consider a solution representing an optimum point. Its descendant obtained by mutation could belong to a region corresponding to a different optimum point, having a higher fitness. The offspring could surpass its parent fitness. Therefore the offspring will survive and a useful optimum point represented by its parent is lost.

To prevent the extinction of some optimum points – especially in the last search stages - we may admit that a mutated offspring have to belong to the interaction range of its parent. We may fulfill this requirement by choosing an appropriate value of the standard deviation parameter (which ensures a high probability the offspring belongs to the interaction range). This strategy is another facet of the local interactions principle.

According to the particular implementation or to the problem at hand other acceptation mechanisms may be considered.

We may also associate an acceptance probability p to each offspring worse than its parent. A simulated annealing scheme (see [11]) may be used to control the mutated solution acceptance according to the probability value p.

8.2. **Simulated annealing acceptance.** In some situations, it is important to have an additional mechanism for preventing premature local convergence. This task may be accomplished by allowing an offspring that is worse than its parent to be accepted in the new generation. In this regard, an acceptance mechanism analogous to simulated annealing technique (see [11]) may be used.

The *cost* associated with the acceptance (maintenance) of a solution c in the new generation is defined as:

$$C(c) = K - f(c),$$

where the real constant K is chosen such that $C(c) \geq 0$, for each solution c.

Remark. The higher the fitness of a solution, the lower the cost to keep that solution in the next generation.

Let d be an offspring (obtained by recombination or by mutation) which is worse than its parent c, we have:

$$f(d) < f(c).$$

The associated cost variation is:

$$\Delta C = C(d) - C(c).$$

It is easy to see that this cost is positive. The probability p of accepting the offspring d in the new generation is

$$p = e^{-\frac{\Delta C}{kT}},$$

where $k > 0$, and T is a positive parameter signifying system temperature.

The values of the parameter k and T controlling the acceptance probability are chosen depending on the specific problem.

By subsequently lowering the temperature, the acceptance probability decreases over time. In the final search process stages very small acceptation probabilities of worse solutions are needed.

By the proposed acceptance mechanism the solutions will generally get closer to the points corresponding to small cost values (high fitness values). Let us observe that the considered acceptance mechanism does not ensure the system reaches thermodynamic equilibrium at each generation (for each value of the parameter T), like in Metropolis algorithm (see [11]) normally used in simulated annealing. We may suppose the equilibrium will be achieved only at the end of the search process.

The equilibrium corresponds to slow temperature variations. We may consider temperature decreasing according to the schedule:

$$T_g = \frac{T_1}{1 + \ln g},$$

where T_1 is the initial temperature and $g > 1$ is the generation index.

To implement the proposed mechanism a random number R having uniform distribution in [0,1] is generated. If $R < p$ then the offspring (worst than its parent) is accepted in the new generation. Otherwise its parent is accepted.

9. Adapting mutation parameter

An important problem with respect to the proposed evolutionary technique is to choose an appropriate perturbation range for the mutation parameter. A related problem concerns the development of suitable adapting technique for this parameter.

We may consider several adaptation mechanisms for the perturbation standard deviation (representing the perturbation amplitude).

To ensure the *fine tuning* of the search process in its final stages we may allow perturbation amplitude decreasing with time.

Another strategy to control the standard deviation parameter may be realized by a self- adapting process. In this case the standard deviation is included in the solution structure (genotype) and it is adapted by the effect of the variation operators.

10. Interaction- range adaptation

Usually the interaction-range is the same for all the solutions. To control sub-population stabilization we may use a mechanism to adapt the interaction radius depending on the specific problem under consideration. Generally it seems useful the interaction radius be a time decreasing parameter.

A radius control mechanism could also ensure a supplementary tuning of the search process right from the first stages.

A possibility for evolving interaction radius is to consider a symbiosis of the current population $P(t)$ and a secondary population whose individuals represent interaction ranges.

We also may consider each solution has its own interaction radius. This parameter may be included in the genotype and evolved during the search process.

11. Population decreasing and stabilization

Short-range interactions permit early solution clustering in sub-populations. Local interactions also favor sub-population stabilization. As a side effect, after a few generations, some solutions might overlap, or become very close, as two or more sub- populations might evolve towards the same optimum point. To detect the correct number of optima is necessary to have only one solution per optimum. To this end, the population size is subsequently reduced by merging similar (close in terms of distance δ) solutions.

If distance between two solutions is less than an appropriate threshold, then the two solutions are merged. This verification will be done at each insertion of a new solution in the population.

The search process stops if after a (previously fixed) number of generations no significant change occurs in the population. Here a significant change is the acceptation of a new generated offspring.

We obtain the number of optimum points as the number of solutions in the final population. Each solution in the final population gives the position of a global or local optimum point.

Therefore we may consider the **CG** approach as being merely a class of optimisation and search techniques based on the local interaction principle. Any useful heuristic may be incorporated.

12. Local and infra-local optima

By maintaining a diversity of sub-populations the Genetic chromodynamics search methods are expected to avoid the problems due to local premature convergence. The proposed approach seems also to be robust with respect to very close optimum points. Close optima may not represent distinct useful solutions, since they are merely local perturbations (due to noise, for instance) of a certain optimum point. We may call them *infra-local optima*.

For most practical problems infra-local optima are solutions of no interest. Local optima of fractal functions may represent an interesting example of such useless solutions. Infra-local optima represent parasite solutions. Their detection is a time-consuming task. Furthermore parasite solutions can also generate confusion in interpreting the results.

13. Applications

Genetic chromodynamics is intended as a general optimization/search technique. GC-based methods are particularly suitable for solving multimodal and multiobjective optimization problems.

Genetic chromodynamics can also be used to solve mathematical problems that traditionally are not treated by evolutionary approaches. Examples of such problems are: equation solving (algebraic, differential or integral equations), fixed point detection and equation systems solving.

The GC approach may be used to solve real - world optimization problems. Genetic chromodynamics flavor methods can be also applied in various scientific, engineering or business fields involving static or dynamic (process) optimization.

Clustering, data compression and other data mining problems are very suitable for a GC treatment. Genetic chromodynamics clustering based methods can be particularly useful to detect the optimal number of clusters in a data set and the corresponding set of useful prototypes. The method is effective even for a very few number of data points (one data point per class, for instance).

14. Conclusions

An evolutionary metaheuristics is proposed. This metaheuristics is called *Genetic Chromodynamics strategy*. **GC** implementations generate a new class of search/optimization techniques. The **GC** approach uses a variable-sized population and local interactions among solutions. Within the methods in the **GC** family solutions are supposed to have different colors. Population dynamics is accompanied by a color dynamics. Short-range interactions permit early sub-populations emergence. The considered local interactions also guarantee the sub-populations maintenance and stabilization.

The solution sub-populations evolve towards the local and global optimum points. The final population contains as many solutions as (global and local) optimum points are detected.

Genetic Chromodynamics strategy is intended to prevent local premature convergence and to solve multimodal optimization and search problems. One of the important features of the **GC**-based techniques is their robustness with respect to local perturbations of the optimum points.

Genetic chromodynamics is a flexible method allowing the incorporation of different general or problem-depending heuristics. We have already exemplified this ability by using a version of simulated annealing to control the acceptance mechanism of a new solution. A similar mechanism could be used to control the mutation process. For some particular problems considering elements of tabu search (see [5]) could ameliorate the performance of the GC method.

Therefore we can consider the Genetic chromodynamics approach as being merely a class of optimization and search techniques based on the principle of local interactions and using a variable- sized population. Each particular chromodynamics technique may also incorporate any useful heuristic.

References

[1] Booker, L., Improving the performance of genetic algorithms in classifier system, J.J. Grefenstete (Editor), *Proceedings of the First International Conference on genetic Algorithms*, Lawrence Erlbaum Associates, 1985, pp 80-92.

[2] Davidor, Y., A naturally occurring niche and species phenomenon: the model and the first results, in R.K. Belew , L.B. Booker.(Editors), *Proceedings of the Fourth International Conference on Genetic Algorithms*, Morgan Kaufmann, 1991, pp 257-273.

[3] Dumitrescu, D., Bodrogi, L., A new evolutionary method and its applications in clustering, Babes-Bolyai University Seminar on Computer Science, 2, 1998, pp. 127-134.

[4] Dumitrescu, D., Lazzerini, B., Jain, L.C, Dumitrescu, A., *Evolutionary Computation*, CRC Press, Boca Raton, 2000.

[5] Glover ,F., Laguna,.M., *Tabu Search*, Kluwer Academic, Publishers, Boston, 1997.

[6] Goldberg, D., E., Richardson J., Genetic algorithms with sharing for multimodal function optimization, Proc. 2^{nd} Conference on Genetic Algorithms, 1987, 41- 49.

[7] Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, Reading, MA, 1989.

[8] Goldberg, D. E., Deb, K., Horn, J., Massive multimodality , deception , and genetic algorithms , in R. Manner , B. Manderick (Editors) , Parallel Problem Solving from Nature, Elsevier,1992, 37-46.

[9] Mahfoud, S.W., *Niching Methods for Genetic Algorithms* Ph.D. Thesis, University of Illinois, 1995.

[10] Shimojima, K., Kubota, N., Fukuda, T., Virus-evolutionary genetic algorithms for fuzzy controller optimization, in F. Herrera, J.L.Verdegay, Eds., Genetic Algorithms.

[11] Van Laarhoven, P.J.M., Aarts, E.H.L., *Simulated Annealing: Theory and Applications*, D. Reidel Publishing.1987.

Faculty of Mathematics and Computer Science, "Babeş-Bolyai" University, RO-3400 Cluj-Napoca, Romania
E-mail address: ddumitr@cs.ubbcluj.ro