Book Reviews

Proceedings of the 1998 IEEE International Conference on Evolutionary Computation. (Piscataway, NJ: IEEE, 1998, ISBN 0-7803-4869-9, 845 pp., $156) *Reviewed by James Kennedy.*

The 1998 IEEE International Conference on Evolutionary Computation meetings in Anchorage, AK, represented an important step in the evolution of the field of evolutionary computation. Some hard questions were answered, some gaps were bridged, and some new hard questions were asked. In this review I comment on the state of the evolutionary computation (EC) paradigm as it is represented in this volume. These comments reflect my biases as a social psychologist and as an participant/observer in this important scientific enterprise. I take it as my assignment here to identify themes in these *Proceedings* that exemplify the directions in which evolutionary computation as a whole seems to be going, and I expect to bring a slightly different perspective to the enterprise.

The word "evolve" comes from a Latin root meaning "to roll out." Generally it has meant "to develop," though of course since Darwin the word has taken on a special meaning, having to do with the evolution of species. *Webster's Ninth New Collegiate Dictionary*'s first definition for evolution is, "a process of change in a certain direction." This seems a little ironic, as EC researchers would consider a population to be changing in an uncertain direction—if we knew the direction, we would just go there! I will define evolution nearer to its dictionary meaning, as a process of change in a direction that results in improvement; my intention is to readmit kinds of adaptation that have heretofore been considered to be outside the usual category of evolution.

A majority of the papers published in this volume discuss genetic algorithm (GA) theory and applications, though there are a good number of evolution strategies (ES) and evolutionary programming (EP) papers. Several papers represent genetic programming (GP) research, and an effort was made to emphasize the many innovative hybrids of the various EC methods, plus hybridization of EC with other techniques. There are papers on economic and game theory, immune-system models, brain-building, alife, fuzzy systems, molecular computing, neural networks, and many other subjects that might not be addressed in this review. A good part of the volume is made up of papers reporting applications of EC methods to practical problems. This review may appear to overlook these important contributions, as well as some important papers that lie outside the mainstream, this is a regrettable consequence of the reviewer's strategy of identifying emerging theoretical trends in EC.

I. Selection and Memory

As Schwefel and Kursawe point out in their paper in this volume

> There are two aspects to [the] modeling process: On the one hand, there is the desire to make use of life's tricks for solving difficult technical or managerial problems; on the other hand, doing so successfully, one hopes to gain some insight into why nature has gone the way it obviously did (p. 1).

Manuscript received April 12, 1999.

The author is with the U.S. Bureau of Labor Statistics, Washington, DC 20212 USA.

Publisher Item Identifier S 1089-778X(99)05421-1.

We learn *about* nature as we learn *from* nature in the evolutionary computation paradigm and in artificial life research generally. And while from a pragmatic engineering perspective there may be no motive for purism, it must be acknowledged that nature often succeeds where programmatic computing performs poorly, suggesting it might serve us well to figure out how it works. Further, since EC theorists draw from the language of genetics and evolution for description, explanation, and inspiration, it seems worthwhile to comment on the present work in that light—and to suggest expanding the vocabulary of natural concepts for this work.

It has been noted previously, e.g., by writers such as Dawkins, Csikszentmihalyi, Gell-Mann, Dennett, and many others, that the evolution of knowledge in a society resembles Darwinian evolution. For instance, new ideas spring up in a population; if they are good ones they propagate across individuals, mutating and merging with other ideas, and if they are not they die off. In such a view, "dying off" really means "being forgotten," and the method of selection is decision. Selection pressures then could be said to apply to ideas as well as to living organisms (keeping in mind that a problem solution might be a kind of idea), and it would seem reasonable to model the evolution of knowledge directly.

There is some confusion, though, in distinguishing between the knower and the known. If patterns of values in a row vector are considered to represent ideas, then they should be passed from one individual to another, with the better patterns increasing in frequency in the population while the worse ones become more rare. Thus, an idea may vanish while the individual who held it survives. It simply does not mean the same thing to say that 10##1#1 was forgotten as to say that an individual characterized by the genotype 1010111 has perished and failed to reproduce. The selection of knowledge is different from the selection of population members. An idea may be forgotten, but we discourage the execution of poorly adapted individuals.

This suggests an adaptive algorithm that looks a lot like an evolutionary algorithm (EA). (In this review I will use this term to include all algorithms that simulate evolution, whether at the genetic level or above.) The major difference is in the type of adaptation. Forgetting is not accomplished by deleting population members, but, instead, individuals change as they move toward better regions of the problem space. *Learning* comprises changes in perpetuated structures to accommodate information in the environment, while in selection a population is made up of the descendants of previous populations, i.e., it is a different set of individuals. Individual organisms have memory, but Darwinian evolution does not, and that is the difference.

Which brings us to elitism. In the present *Proceedings* and elsewhere, we note that almost every so-called evolutionary algorithm introduces wise immortals into the population. The idea that the very best members of the population or their clones always survive from generation to generation is alien to anything in natural evolution. It is apparent from the papers in this volume that the paradigm is moving toward inclusion of models which maintain and modify durable individuals, as compared to iteratively replacing population members. Surviving over time, elite individuals can retain information (this is memory), modify it (something like mutation), and communicate it to other population members (something like crossover). Several recent paradigms address the topic directly, and an increasing number of researchers seem to be pushing toward the idea or searching for it: the idea is social learning.

Perhaps the best-established, most comprehensive, and most explicit social-learning approach is the cultural algorithm method promulgated by Reynolds. A cultural algorithm comprises a population of individuals who interact with one another and with knowledge (informational entities called "mappa") stored in a "belief space," which is dynamically and stochastically updated as new information about the problem is found. A belief space can be broadly considered a kind of elaboration of elitism, as superior schemata are retained over time for interaction with ordinary population members. Reynolds and his colleagues conceptualize the belief space as an analog of cultural traditions and artifacts that encode successful solutions to complex problems faced by a population. Cultural algorithms have been shown to be a flexible adjunct to various EA's and a powerful method for theoretical insights and practical applications. Reynolds and Al-Shehri's paper describes the use of cultural algorithms with evolutionary programming to guide decision tree induction in large databases.

A number of other papers in these *Proceedings* implement similar types of hybridization of population members with some preserved subset of excellent individual structures or schemata. In Hidrobo and Aguilar's method, details of successful individuals' structures comprise a "collective memory." Better values in the collective memory are saved in "trace individuals" to be transferred probabilistically to "reference individuals" and recombined with members of a GA population. Puppala *et al.* also maintain a repository of successful problem solutions. Their coevolutionary method optimizes the coordinated behaviors of a pair of actors, a "whitewasher" and a "painter," who must execute a task together. Successful individuals are stored in a "shared memory" and the algorithm pairs individuals from one population (e.g., whitewashers) with individuals from the other population's shared memory. Nakashima *et al.*'s application for constructing linguistic rules in a fuzzy system incorporates a kind of GA in which a set of elite solutions is stored; in each generation, a few solutions are randomly selected from that set to be introduced into the GA population.

Voigt and Lange argue that good local search methods have memory. They suggest and implement a "random memorizing" technique for use in evolution strategies, or any algorithm which has a global step-size control. A memory, i.e., a stored previous best solution, is retrieved at random to compute a search direction for the current population members. Harik *et al.* make the interesting observation that the population in a GA plays the role of a memory, compensating for decision errors made by random crossover and mutation by maintaining successful schemata across generations.

It is not hard to see why researchers would want to retain the best structures from each generation for recombination with later generations. For instance, in these *Proceedings* Kemenade's analysis of the performance of GA's on NK-XOR landscapes shows that elitist GA's simply perform better than the comparable nonelitist version. Memory, while it is not part of the evolution of species in nature, fits well with the kinds of evolution performed in artificial systems.

There may be some misconception in the usage of the concept of memory. Elite individuals, if they persevere over iterations, may be thought to *possess* memory, but it is incorrect to call them memories. This is especially true in those paradigms where the "memories" are structurally identical to the population members, and interact with them through crossover or other operators. It is again the difference between good ideas and individuals with good ideas. Where schemata are stored, that is subsets of individual vectors, these should be thought of as salient arguments of persuasive individuals, rather than disembodied ideas. In social interaction we do not adopt random beliefs of others, but salient ones.

A number of researchers have adopted the framework of coevolution to implement a kind of durable population like a belief space, where successful individuals or schemata are perpetuated across generations, to recombine with population members. For example, Handa *et al.* solve constraint satisfaction problems using a coevolutionary algorithm which they describe in terms of two populations; a host population searches for good regions in the problem space, and a parasite population searches for good schemata on the hosts. "Transcription" is used to insert parasitic schemata into hosts. They compare their approach to cultural algorithms, with the parasite population of proven schemata serving as a kind of belief space.

Kubota and Fukuda's virus-evolutionary genetic algorithm is based on a similar coevolutionary concept. The premise is that viruses are sometimes responsible for the transmission of DNA segments from one species to another. Thus a kind of GA is implemented, in which the host population looks for an optimal problem solution while a parasitic population of viruses searches for effective schemata on the host. A virus can pick up a schema through an operator the authors call "incorporation " and can overwrite genes on the host through "reverse transcription." Thus viruses can propagate good schemata; Kubota and Fukuda's results suggest that the method can contribute to the performance of a GA.

Though it is not the primary method of adaptation in social systems, selection of population or group members does play a part. Particle swarm optimization is a method that is based on social interaction, in which every individual retains a memory of its previous best performance. Individuals are attracted toward their previous best positions and toward the previous successes of their neighbors. In a reverse twist of what has been said here, Angeline suggests adding selection to that algorithm and demonstrates an improvement in performance when the poorest half of the population is replaced by the best half at each generation. This is analogous to a group process where poorly adjusted members drop out and are replaced with individuals who more closely resemble the well-adjusted group members.

The papers cited in this section provide evidence of an expansion of the EC paradigm to include social-psychological group processes taking place in populations of enduring individuals. Terms such as "culture," "collective memory," "shared memory," etc., refer not at all to the evolution of species; writers who use these terms are overtly citing social and cognitive processing as the basis for their adaptive algorithms. In other instances, writers who retain the language of evolution and genetics seem to have simply transferred the meanings of terms from one paradigm to another. For instance, to say that populations "learn" is to introduce a whole new meaning for the word. In ordinary use, learning is a change of state of an individual organism: epigenetic evolution, perhaps.

II. Recombination and Social Interaction

In *Bully for Brontosaurus*, Gould wrote that

> The basic topologies of biological and cultural change are completely different. Biological evolution is a system of constant divergence without subsequent joining of branches. Lineages, once distinct, are separate forever. In human history, transmission across lineages is, perhaps, the major source of cultural change. [1, p. 68]

Of course evolution diverges between species, but sexual reproduction promotes convergence within them, enabling fine-tuning of genetic problem solutions. In some ways then the introduction of crossover to an evolutionary algorithm invites comparison to cultural cross-lineage transmission of information.

People talk to one another, adopt one another's explanations for things, get facts from one another, copy one another's mannerisms: homo sapiens (literally, "intelligent man") is fundamentally a social creature. As individuals integrate facts and explanations gathered from others, the global result is a kind of cultural convergence on patterns that allow members of the population to make sense of the world. In EC this kind of interaction is approximated through crossover.

From the evidence of these *Proceedings*, it appears that recombination in artificial genetic systems has simply been liberated from nature. Children are routinely immaculately conceived in single-parent families, or as a consequence of orgiastic practices involving multitudes of parents. Recombination operators in this volume include single-point, double-point, uniform, and diagonal crossover, a "grouping" genetic operator, "subschedule" crossover, "part type exchange" crossover, "multistep crossover fusion," "harmonic" crossover, "unimodal normal distribution" crossover, "center of mass" crossover, "multi-parent feature-wise" crossover, a "seed" crossover operator—the examples of nature are disregarded, they seem irrelevant. Various EC paradigms implement crossover of genes, strategy parameters, phenotypic traits—again, it seems apparent that researchers are pulling away from the root metaphor of biological evolution, toward a framework that allows richer and more versatile interactions of individuals in a population: a society.

As we have been talking about the evolution of knowledge and ideas through a social system, we should note that innovations sometimes occur simultaneously to dispersed members of a scientific community. Historical examples are well known. It appears that three independent sets of EC researchers have published papers in this volume introducing nearly the same new algorithm. Corno *et al.*'s "selfish gene" algorithm, Harik *et al.*'s "compact genetic algorithm," and, to a lesser extent, Abbattista and Dalbis' "scout algorithm," each presented here for the first time, represent the binary population in a vector of probabilities, adjusted as a result of the performance of ephemeral individuals in competitions. Probabilities of winners' alleles are increased slightly, while—at least in the selfish gene and scout algorithms—the loser's probabilities are decreased. The scout algorithm contains some elements not found in the others, and there are small differences, e.g., in the formulas for updating probabilities, between the other two, but the three are very similar. The most obvious precedent to these methods is population-based incremental learning, or PBIL, which is cited in only one of the three papers, and which differs from these algorithms in important ways. This appears to be an efficient and effective way to simulate evolution, which will almost certainly become more prominent in the EC paradigm over time. It also appears to be a case of social evolution within a scientific community converging on an interesting local optimum.

III. Species and Cultures

Algorithms that incorporate some form of crossover—and that includes most of the contents of this volume—commonly suffer from premature convergence of the population. The effect of this is that progress stops as diversity is eliminated in the population.

An opportunistic approach to the premature-convergence problem was taken by a number of researchers in this volume. For instance, Evans' Hypergamous Parallel Genetic Algorithm (HPGA) "embraces premature convergence as exponential-like convergence to some (probabilistic) local optimum" (p. 621). In HPGA, subpopulations or "islands" are evolved using crossover. These subpopulations, also called "demes," converge on local optima and are used to seed higher-level populations called "melting pots."

Deme evolution, while not new, seems to be gaining adherents. In a robotic application, van de Logt and Walter want to find all the solutions for a given multimodal optimization problem. After beginning their GA with isolated subpopulations, they apply "exchange" to compare the best members of all subpopulations. If they are too similar then they are replaced with randomly generated new individuals. Second, "migration" is used to move individuals from one subpopulation into another. The researchers want the populations to converge in different directions, and this is what they report happens. Pedroso and Wako-shi use "niche search" with an interesting kind of elitism to route a vehicle through a complicated graph. Subpopulations are evolved, with elitism within them. When multiple subpopulations evolve the same elite individual, then only one of them is allowed to be elitist, so the others continue searching. As with the previous authors, their aim is to search intensively in good regions—elitism ensures that—while maintaining diversity—nonelitist evolution within subpopulations ensures that.

Eiben *et al.* use migration as well, plus an operation they call "redivision." Evolving subpopulations contribute some randomly selected members to a migration pool to be distributed back to the subpopulations randomly. This increases the sizes of some populations while reducing the sizes of others. Redivision is used to give the smallest subpopulations another chance. Members of subpopulations are allocated to a redivision pool, with smaller populations contributing fewer individuals, and then individuals are taken from the redivision pool and distributed among the subpopulations in such a way that the smaller ones increase in size and the larger ones become smaller.

Shoaf and Foster also evolve subpopulations in their stock portfolio application and migrate the most-fit members periodically. They argue that the deme model is better than sequential single-population trials because it allows alternating periods of hill-climbing and competition among local optima.

Some discussions of deme models have described these in terms of species formation on islands where isolated populations converge genetically. One problem with this is that crossover can only occur between individuals who are members of the same species; once populations have speciated they could not migrate and recombine with populations on other islands. We could stretch the point by describing it in terms of the evolving of different breeds or races of a species, with slightly different characteristics, but this metaphor seems rather weak, as the similarities among natural deme populations of the same species would overwhelm their differences. Also, presumably artificial subpopulations might bear no genetic resemblance to one another, so it would be hard to classify them as one species. Another problem with the analogy is that island populations in nature evolve in different environments. In the usual EC implementation, subpopulations evolve to solve the same problem. It is easier to think of these methods in terms of formation of cultures and interactions among them. Migration then is emigration and immigration of individuals between societies transmitting patterns of beliefs and behaviors from one community to another, as remarked by Gould, above. (We note also that the word "deme" comes from a Greek root meaning "people," as in "demographic," "democratic," etc.)

IV. Selection and Cognition

Two distinct communities study the mechanisms of intelligence, these being psychologists and computer scientists. Early in the history of electronic computing, it was realized that some aspects of intelligent human information-processing could be simulated, and perhaps even exceeded, on a digital machine. Thus the relationship between mind as a human phenomenon and artificial or computational intelligence has always been strongly implied, though it is rare for any investigator to have scholarly knowledge of both fields. It is possible—in fact, it is common—to come up with a theory of psychology based on introspection and "common sense," to implement the model in a computer program, and to demonstrate "intelligence" through successful application of the model to some problem domain. Happily, explicit examples of this faux pas are rare in these *Proceedings*.

Where intracranial single-agent systems are the tradition in artificial and computational intelligence research, the model of intelligence is individual cogitation. This is true for standard fuzzy systems and neural networks, as well as the heuristic algorithms usually associated with good old-fashioned artificial intelligence. If we are to consider evolutionary computation in such a way, where inside-the-head populations of hypotheses are compared, mutated, and interwoven, an internal cognitive process is suggested that is highly parallel and irrational. This is not a negative criticism. Ghiselin, in the introduction to his classic collection of essays by creative thinkers including mathematicians, musicians, artists, and scientists, noted that creative thought seems never to occur by a rational method, but rather arises out of a kind of "commerce with disorder." Almost 40 years ago D. Campbell, a psychologist, proposed a model of creative thought through "blind variation and selective retention" which was deliberately based on evolutionary concepts. Reversing the perspective, anthropologist G. Bateson (son of the nineteenth-century geneticist W. Bateson) suggested in his 1979 volume *Mind and Nature* that evolution is a type of mind. Some versions of evolutionary epistemology and psychology have existed at least since Darwin's time. Adaptive trial and error can be a very powerful cognitive method.

While the iterative mutation/crossover/selection processes typical of EA's are obviously not identical to the processes of creative problem solving, they may be similar to the nonconscious algorithms that drive some automatic human information processing, i.e., the spontaneous problem solving that lies outside of awareness. Thus, EC researchers might be able to learn a trick or two from cognitive psychology. If we were to conceptualize an EC population as a kind of cacophonic interior multilogue, then it should be possible to adopt or at least investigate operations that have been discovered in cognitive research.

V. Scientific Methods in Evolutionary Computation

One thing is shamefully lacking in this collection of papers, and that is scientific methodology. Rarely are falsifiable hypotheses specified, almost never is an experimental design described, and rigorous data analysis is practically nonexistent. In a casual conversation it may be sufficient to say to a friend, I tried such-and such, and it seemed like this one was better than that one, but this kind of reporting does not meet the standard for scientific research. If independent variables are manipulated, then the reader should be told what they are. Besides reporting means (or worse, graphs without means), authors should give standard deviations, at least, or confidence intervals, so that readers can judge whether differences between conditions appear to be reliably greater than chance. Even better, inferential statistical tests, especially analysis of variance but including t-tests and regression analysis, should be conducted and reported.

Clearly much progress has been made using informal research methods, but even after several decades (albeit following a slow start) the knowledge of the field is largely limited to a handful of theorems and some rules of thumb. This is a field which combines deductive mathematics and inductive empirical investigation in a unique way; the reliance on empirical evidence to compare search strategies absolutely necessitates the adoption of rigorous research methods.

VI. Conclusions

The EC paradigm is typified by systems comprising the following:

- populations of independent-variable vectors;
- interactions among population members;
- methods for deriving new vectors to test new points in the function space;
- reward for good performance;
- an explanation by analogy (i.e., biological evolution) for choosing and understanding the methods used.

Populations may be implied or virtual, as in the case of the compact, selfish-gene, and scout algorithms mentioned above. The interaction among population members may be as simple as tournament selection or as complicated as some of the crossover methods mentioned above. Methods for making new vectors might view the next generation individuals as perpetuations, modifications, or descendants of previous ones, or as entirely new members disumbilicated from history. Good performance may be rewarded by the transmission of genotypic or phenotypic features to the next generation, by elitist immortality of the individual, or by increasing the influence of the individual through adjustment of probabilities of crossover, numbers of offspring, etc. The aim is to increase the effect of superior vectors and decrease the effect of lesser ones, recognizing that the contribution of the lesser individuals might turn out to be critical, especially in allowing escape from local optima.

The explanation by analogy in EC is uniquely important, but at the same time it is unusually flexible. The various EC paradigms are based on very different zoom angles of biology, and their explanations are hardly comparable, but researchers hybridize methods anyway—and with good results.

It is clear from the current set of research reports that those in the field have decidedly extended operations beyond those that can be explained in terms of biological evolution. As a social psychologist, I note that many of the extensions of EC methods seem to borrow from the template of human society. Of course there may be many varieties of adaptive algorithms—probably the number is not finite. I am watching for theorists to move the discussion up a level, to come up with a model of collective adaptation that transcends these subcategories, and fits them together as coequal methods for changing in a direction that results in improvement.

References

[1] S. J. Gould, *Bully for Brontosaurus: Reflections in Natural History*. New York: Norton, 1991.