Revue Internationale de

systémique

ISSN 0980-1472

ÉVOLUTION ET APPRENTISSAGE
EN ÉCONOMIE

L'hypothèse Lamarckienne remise en question

Sous la direction d'Ehud Zuscovitch

Vol. 7, N° 5, 1993

afcet

DUNOD

AFSCET

Revue Internationale de
systémique

Revue
Internationale
de Sytémique

volume 07, numéro 5, pages 573 - 592, 1993

A perspective on evolution and the Lamarckian hypothesis using artificial worlds and genetic algorithms

Harry Wechsler

REVUE INTERNATIONALE DE SYSTÉMIQUE
Vol. 7, N° 5, 1993, pp. 573 à 592

A PERSPECTIVE ON EVOLUTION AND THE LAMARCKIAN HYPOTHESIS USING ARTIFICIAL WORLDS AND GENETIC ALGORITHMS

Harry WECHSLER[1]

Abstract

This paper addresses evolution and the Lamarckian hypothesis using the framework of artificial worlds and evolutionary computation. The Lamarckian hypothesis is specifically concerned with whether acquired variations ("adaptations") can be inherited. Artificial worlds are large scale artifacts and represent visualization means for simulating collective behavior when analytical techniques break down. Analogies drawn from natural selection, optimization, and machine intelligence are brought to bear on the proper design of such artificial worlds. Artificial worlds predict future behavior, check which properties are arbitrary and which are not, and could possibly suggest means for preventing harmful behaviors from emerging. New conjectures can be tested for and the role that information plays in evolution and innovation can be properly assessed. Qualitative arguments suggest an updated version of the Lamarckian hypothesis involving knowledge transfer across developmental life cycles.

Résumé

L'article traite de l'évolution et de l'hypothèse lamarckienne dans le cadre analytique des mondes artificiels et du calcul évolutionniste. L'hypothèse lamarckienne est une manière de chercher à déterminer si les variétés acquises (adaptations) peuvent être transmises par le mécanisme d'hérédité. Les mondes artificiels sont des artifices analytiques de grande taille représentant un moyen de visualisation pour simuler le comportement collectif lorsque l'emploi des techniques analytiques s'avère impossible. Des analogies au processus de sélection naturelle, d'optimisation et d'intelligence des machines sont employées dans la

1. Department of Computer Science, George Mason University, Fairfax, VA 22030, USA.

conception de tels mondes artificiels. Ces mondes servent à prévoir des comportements futurs, à vérifier quelles propriétés sont arbitraires et celles qui ne les sont pas, et peuvent suggérer des moyens pour prévenir l'apparition des actions endommageantes. Des nouvelles conjectures peuvent ainsi être testées et le rôle que l'information joue dans l'évolution et l'innovation peut mieux se saisir. Des arguments qualitatifs suggèrent une version adaptée de l'hypothèse lamarckienne impliquant de transfert de connaissances entre des phases des cycles de vie.

I. INTRODUCTION

We survey in this paper what artificial world simulations using evolutionary computation, in general, and genetic algorithms in particular, can offer to economics in terms of insights and/or predictive power regarding innovation and emerging structures. The point of departure for our discussion is the thesis advanced recently by Brian Arthur (1992) that "the standard mode of theorizing assumed in economics is deductive – it assumes that human agents derive their conclusions by logical processes from complete, consistent and well-defined premises in a given problem using mostly static environments. This works well in simple problems, but it breaks down beyond a "problem complexity boundary" where human computational abilities are exceeded or the assumptions of deductive rationality cannot be relied upon. Beyond this problem complexity boundary, and despite problems becoming ill-defined, humans still continue to reason well, but by using induction rather than deduction."

Rumelhart and McClelland (1986) advance a similar thesis when they claim that "the time and space requirements of any cognitive theory are important determinants of the theory's plausibility. Complexity satisfaction provides a major source of constraints on the solution of the problem. Much of past analytical and theoretical work has tacitly assumed that the language of continuous mathematics is equivalent to the language of computation. Mathematical modeling, however, is not equivalent to computational modeling. There are still issues of representation, discretization, sampling, numerical stability, and computational complexity (at least) to contend with." To address these very issues, Simon (1982) has introduced a normative science (of the artificial) concerned with optimal engineering design and has defined the principle of bounded rationality, *i.e.*, that people have to accept "good-enough", possible suboptimal but still satisfying solutions, because computationally there is no other choice. Bounded rationality

belongs to imperative logic, the paradigm underlying procedural (behavioral) rationality. Specifically, given a set of constraints and (fixed) parameters, find those values of the implementation variables that maximize the utility of the design. Brian Arthur would probably summarize the same arguments concerning system design and analysis by suggesting that deductive rigor needs to be replaced by "behavioral rigor based on precise and observed actualities of human behavior."

Both the problem complexity boundary and the principle of bounded rationality ask for complex simulations if the economists were to become interested in pursuing behavioral rigor in terms of predictive power regarding innovation and emerging structures. Artificial worlds (AW) (Lane, 1992), computer simulated stochastic models consisting of agents that interact with each other and with their natural environment in prespecified ways, are then the only solution available to find real answers for real world situations. We emphasize that there is nothing artificial about AW except the use of an appropriate visualization medium for simulating the real world of economics. Artificial worlds allow for evolving preferences and nonfixed utility functions and they can be useful for simulating collective behavior. When analytical tools are lacking, the case can be made that heretical conjectures along with ortodox dogma should be allowed to compete and discharge questions regarding equilibrium theory through the proper conceptualization of space & time, information, and utility.

Artificial worlds are quite different from artificial intelligence (AI). As we are going to encounter AI again it is worth now to define what it usually means. AI involves achieving, explaining, and/or simulating problem solving and decision making capabilities as those continuously exhibited by human subjects in meeting the world surrounding them. Searl (1984) has forcefully argued, however, against the idea of AI being able to duplicate human intelligence on grounds of functionality and intensionality. Judgments involving mathematical truths are not necessarily algorithmic and furthermore, consciousness, intentionality, and insights are needed to comprehend the full implication of truth statements (Penrose, 1989). Artificial worlds can provide a true perspective on behavioral rigor using complex simulations based on real assumptions, real interactions, and real consequences of specific behavior.

Terra firma for "intelligent" behavior in general, and economics, in particular, involves by default information processing and knowledge transfer. According to Marr (1982), the specific descriptive levels where tasks of varying complexity have to be understood are those of strategy, process (representation and algorithm), and mechanism (implementation). Basic computational theory, the first level, specifies the task, its appropriateness, and the implementation strategy. Next, the representation and algorithm specify the computational approach in terms of input and output representations, and the corresponding transformations. Consequently, the task determines the mixture of representations and algorithms, and a good match between the three levels is highly desirable. Furthermore, according to Simon (1982), "all mathematical derivation can be viewed simply as change of representation, making evident what was previously true but obscure. This view can be extended to all of problem solving – solving a problem then means representing (transforming) it so as to make the solution transparent." In other words, knowledge has to be made explicit if and when needed, and it would be up to the artificial worlds to accomplish this very task.

The natural way for handling complex decision problems is through learning and adaptation. Learning denotes systemic changes that are adaptive in the sense that they improve system performance on future endeavor involving tasks similar to those engaged in the past. We learn many times using analogies, and the result is that we constantly update the representations we hold on the world surrounding us. Knowledge acquisition is much more than accruing facts and it involves also their purposeful (functional) organization. Incremental learning and unlearning, adequate reinforcement and how to learn from mistakes as much as from positive experience, and credit (and critique) assignment are amongst the issues concerning learning and adaptation. Supervised learning is nevertheless quite restricted in its scope, and in fact, from a biological viewpoint, even less important for success than the emergence of adequate representations. It is up to the artificial worlds, however, to bring up discovery and innovation using self-organization and evolutionary computation. It is within the context of adaptation and evolutionary computation that we consider the Lamarckian hypothesis and its plausibility for economical markets. The Lamarckian hypothesis is specifically concerned with whether variations ("adaptations") acquired during the lifetime of the phenotype ("agent") can be incorporated in the genotype ("organization") and inherited later on through regular mechanisms of heredity ("inheritance").

Artificial worlds hold much potential for addressing problems whose solution would have great impact on modern economical theory. As an example, we could test conventional economic theory based on the assumption of diminishing returns vs. positive feedback whereby small chance ("chaotic") events (and necessity) early on in the history of an industry or technology can tilt the competitive balance (Brian Arthur, 1990). One should note that the

thesis supporting positive feedback for the economy is selective in terms of its applicability and that it appears to hold only for the high-tech, knowledge intensive sector of the economy. Early superiority leading to selectional advantage, an important ingredient for positive feedback, is also apparent during genetic programming where it leads to premature convergence in terms of fitness. It is the same knowledge factor that could make the Lamarckian hypothesis respectable again. If the gene pool (genotype) consists, among other things, of information, then inheritance of acquired characteristics becomes possible because humans can transmit innovation represented in terms of specific knowledge and experience through teaching and learning. Another important issue for artificial worlds to decide upon is what is the driving force behind innovation and novel economical organizational structures. Can the tendency toward relatively stable states of equilibrium, known as homeostatis, be held responsible for observed social and economical behavior, or maybe we should subscribe to the heterostasis thesis advanced by Kloft (1982), that "intelligence in complex systems is a concomitant of a striving for a maximal condition, whereby agents "seek" excitation and "avoid" inhibition." The change of paradigm, from homeostasis to heterostasis, should be tested for, and the role of internal drives such as motivation in general, and active search for innovation (novelty) by agents or entities, in particular, be properly assessed.

II. SELECTION, OPTIMIZATION AND ADAPTATION

We explore natural selection, computational sciences, and machine intelligence in order to draw analogies for how artificial worlds should be simulated and assessed using evolution, optimization, and adaptation criteria.

II.1. Natural Evolution and Selection

The process of natural evolution leads to change as a result of adaptive strategies being continuously tested for fitness by the environment as it should be the case for closed-loop feedback control. Evolution simply assumes selection ("survival of the fittest") but recent developments such as the neutralist theory, to be described later on, paint a much more complex picture. As it has been suggested by sociobiology, the agent and the environment mold each other in their continuous interactions. Evolution is not necessarily a smooth and continuous process. As an example, punctuated equilibrium describes evolution as a discontinuous process, where "revolutionary" change might be the norm rather than the exception. Gould (1989) recounts the events related to deciphering the Burgess shale, a limestone quarry rich in life's remnants after the Cambrian explosion, and concludes that there is no predefined ladder of evolution and that if one were to "replay the tape a million times and it is doubtful that anything like Homo sapiens would ever evolve again." The punctuated equilibrium eliminates species by "lottery", even when evolution, driven by contingency and opportunism, thrives towards increasing system complexity. Natural selection and the origin of species is the outcome of both chance and necessity.

The discoveries made by Darwin and Wallace on the mechanisms of natural selection, and by Mendel on genetics and inheritance, are fundamental to evolutionary theory and population genetics and they have led to neo-Darwinism or the synthetic theory. The next milestone was achieved by Watson and Crick when they broke the genetic code and set up the basis for the present explosion in molecular biology. Arguments on species, their taxonomy, and on heredity, explain how complex organisms emerge over space and time. Molecular genetics holds that inheritance of acquired characteristics is not possible and that there is no basis for the Lamarckian hypothesis.

The material basis for heredity is DNA, a ladder-like molecule whose message is encoded using the "letters" (chemical bases) A, G, T, and C. The message is read in triplets, the words are of two kinds, "stop" or "aminoacid X", and twenty different amino acids are encoded by triplets. The genetic message specifies sequences of amino acids terminated by stop signs, and, when translated and properly expressed, the result is proteins, which are chains of amino acids. Genes are sections of DNA which specify a discrete amino acid chain. Chromosomes, are composed of genes, which may take on some number of values called alleles. One can talk of a particular gene, for example an animal's eye color gene, its locus, position, and its allele value, blue eyes (Goldberg, 1989). As a result of its interactions with the environment, the genotype, the genetic package some prototype organism ("entity") is endowed with, is expressed as some specific phenotype ("agent") following the development ("embriologic") stage. The genotype is the blueprint for "solution" and the phenotype is its expressed instantiation.

According to Gregory (1987) selective mating and variation are held responsible for natural selection (through the survival of the fittest) and they ultimately lead to evolutionary change across the "genepool". "Inheritance is not blending: genes are passed unchanged from generation to generation once chromosomes pair off and random crossing-over takes place." According to the original theory espoused by Wallace, hereditary variations advantageous in a changing environment will be selected, and the species will change to remain in harmony (like constant numbers) with the environment. Variation, measuring the degree of diversity and corresponding to internal entropy, comes from the shuffling of genetic material taking place during crossing-over, from hidden genetic variation, and from mutations ("mistakes") in replication of DNA. It is assumed that mutations are rare and that most of them cannot be passed on to the next generation. Selection can be thought of as either eliminating variation (purifying, directional) or maintaining it (balancing, stabilizing); and also as either promoting change (directional) or maintaining the status quo (purifying, stabilizing).

Selection of the balancing type ensures that the norm maintained is that of a polymorphic population. An alternative explanation is the "non-Darwinian" proposal that the observed protein variants are selectively neutral; they confer no real advantage or disadvantage and are maintained purely by chance, or by random genetic drift. On this neutralist theory, harmless or even slightly deleterious mutations may spread and become fixed (or eliminated) purely by chance. As with natural selection, experiments show that genetic drift occurs especially in small populations where biased samples are more likely. "In essence, the neo-Darwinian theory expects that a given variation is correlated with some environmental variable, and the neutralist theory expects that it is not" (Gregory, 1987). Good things need to be taken in moderation and selection makes no exception. Excessive selection can be harmful too so genetics features such as diploidy (pairs of chromosomes) and dominance shield alternate solution from extinction. Dominant genes are always expressed (heterozygous or homozygous modes of reproduction) while a recessive gene is only expressed when it shows up in the company of another recessive (homozygous reproduction). Long-term memory expressed as recessive allele combinations protects past experience and "permits alternate solutions to be held in abeyance – shielded against overselection" (Goldberg, 1989).

Natural selection modifies and preserves old species but this alone will not produce new species ("structures"). Models for the origin of new forms of life include (i) gradual adaptive divergence (Darwin) and/or speciation possibly primed by gradual selective change within geographic isolation (neo-Darwinism) and involving small populations, (ii) accidents and macromutations (multiplication of chromosomes – polyploidy) involving individuals rather than populations, and (iii) punctuated equilibrium or quantum speciation, an intermediate alternative, where no geographic isolation is necessary, accident rather than natural selection is the cause, inbreeding amongst descendants of a single individual is necessary, and a new species may arise in a few generations with the establishment of a homozygous population for mutation (Gregory, 1987). As the paleontologists generally fail to find evidence of gradual transformation in the fossil record, punctuated equilibrium is gaining increased respectability. "All" what it took to move from chimpanzee to *Homo Sapiens* is a minute amount of chromosomal change ("mutation"). One should find it normal then, that as the rate of technological change and knowledge dissemination heats up (corresponding to higher entropy), the rate at which new companies and forms of production appear (and old ones disappear) should accelerate. Random mutations throw up material on which selection acts; most offshoots perish and but a few succeed as new species. Since the majority of inferred change in DNA is (silent and) due to mutations, small populations should be frequent in the history of species, since drift is more effective in small populations, and most effective in inbreeding (Gregory, 1987).

Since silent substitutions in the DNA are more frequent than the non-silent ones (with an effect on the phenotype) it appears that the majority of evolutionary change is "immune" to natural selection. The effect of the large neutral or "non-Darwinian" component discovered in molecular evolution is to downgrade further the role of design and increase the role of chance (Gregory, 1987). Embryology and epigenetics need still to explain development or the unfolding of form time after time using genetic blueprints. Control genes, self-organization, and possible neo-Lamarckian modes of changes have been suggested so far but none is yet conclusive and widely accepted. It is commonly accepted that ontogeny, an individual organism's embryonic development, follows phylogeny, the evolution of the species. This law is subject to exceptions, and there are cases suggesting an evolution in the opposite direction. The evolution of the skull and the face in higher primates and man seems to be one of those exceptions. It seems as if what is a transitional stage in the ontogeny of other primates became a terminal stage in man (Changeux, 1983). The mystery of transformation lies then at the interface between genotype and its expression, the phenotype.

II.2. Computational Sciences and Optimization

Optimization and evolution of successful design fit to survive competition, and its intrinsic computational complexity, require sophisticated means for information processing. Useful analogies come from population genetics, sociobiology, operation research and probabilistic computation, simulated annealing and Boltzmann machines, and non-linear dynamical systems and chaos.

From population genetics comes the Fisher – Eigen fundamental theorem (of natural selection) stating that the average fitness increases (proportional to the variance of the fitness) (Hofbauer and Sigmund, 1988). It can be shown that the theorem expresses a mixed evolutionary strategy that combines mutation in terms of thermodynamic search (see simulated annealing below) and selection (Boseniuk *et al.*, 1987). (This theorem should be hedged, however, with the comment made by Isaiah Berlin regarding the "fox that knows many things, but the hedgehog knows one big thing. Foxes have a diverse vision of life. They are skeptics, agnostics, tolerant, centrifugal. Hedgehogs have a central, systematic vision of life. They are believers, doers, committed, centripetal, often fanatical. Certain moments in history allow foxes to prosper. They tend, however, to be rare."

Sociobiology (Wilson, 1982) involves itself with the evolutionary interpretation of behavior and draws much from ethology, cellular automata such as the "game of life" using dynamical behaviors resulting from different initial configurations (Cowan, 1982), and/or evolutionary social games (Axelrod, 1984). It interprets behavior in terms of strategies which could have selective advantage when they increase the chance of survival of those sharing their genes – altruism is one intriguing example of such behavior while dumping and gaining market share would be another one. Among the issues that come up we mention limited conflict ("competition") that benefits individual individuals ("agents") as well as their species ("organizations") and are characteristic of evolutionary stable strategies (ESS) (Maynard Smith, 1982), growth rates and ecological models (Hofbauer and Sigmund, 1988) involving the Lotka – Volterra equations tuned for different types of ecosystems such as competition, symbiosis, and/or host ("prey") – parasites ("predator"). Co-evolution of parasites has been shown to prevent genetic algorithms from being trapped in local optima and to make testing for fitness more efficient (Hillis, 1991). The benefits accruing from co-evolution are the result of two independent gene pools striving to develop evolutionary stable strategies, similar to the permanent flux alternating between epidemics and immunity. The parasites are scored according to how well they find flaws in the host and cannibalize it, while a host ("pray") is successful according to its ability to starve the parasites. The question whether behavior is genetically determined is highly controversial and it is reminiscent of old reductionism arguments such as nature vs nurture and/or heredity vs. environment. Whether the tyranny of the genes should be held responsible for society's evils or not has become a legitimate and highly loaded issue. The Lamarckian hypothesis, if it were to be true, would diminish the role of the genes and let some fresh air blow in.

Solutions for optimization problems are found using perturbations methods and belong to operation research. Standard techniques include hill climbing and greedy algorithms (vs. blind algorithms) along successful perturbations, and branch and bound where search is aborted along those paths whose estimated cost is higher than some *a priori* threshold. Hill climbing considers all perturbations from some given current state and would choose the best one as the next state configuration. Hill climbing, a local method whose horizons are limited could possibly underlie the Lamarckian hypothesis. Problems (and possible solutions) related to hill climbing and its inherent locality include local maxima traps (backtrack), plateau where local comparisons are not enough (big jumps/perturbations) and ridges (move in several directions at once) (Rich and Knight, 1991).

Complexity analysis usually assumes worst case behavior. Optimization solutions whose worst behavior is "good" are obviously the most desirable ones. Due to finite resources in many instances one might also be satisfied with solutions that on the average are well behaved. In fact, one might even accept solutions that have a small likelihood for wrong answers. Computations involving instructions that make random choices are called probabilistic. Most probabilistic algorithms fall within two classes – Las Vegas and Monte Carlo. Las Vegas algorithms never return an incorrect answer, but sometimes they do not find an answer at all, while Monte Carlo algorithms always give an answer, but the answer is not necessarily right. The probability for success increases as the time available to perform random choices increases.

The motivation for random (and non-deterministic) moves using probabilistic choice comes from thermodynamics. Simulated annealing optimization, whose goal is to achieve increased complexity (higher entropy states), proceeds along energy gradients characteristic of fitness landscapes. For reasons involving search performance and to avoid being trapped in local maxima, the process can be occasionally perturbed using non-optimal moves that temporarily lead to lower energy states. Non-optimal moves are accepted according to a control parameter labeled as "temperature", and the higher the "temperature" the more likely it is that perturbations of decreased fitness will be accepted. (Annealing done at zero temperature corresponds to hill climbing and greedy algorithms.) Another paradigm, that employing Boltzmann machines, facilitates deriving behavioral distributions fit for some predetermined environment. If forced ("clamped") behavior at

equilibrium fits the "free" run of the network structure then good match between network organization and its environment results (Ackley *et al.*, 1985). As it has been already suggested, mixed strategies taking advantage of individual development and built around life cycles can be quite beneficial. During "childhood" exploration of alternative solutions is carried out using simulated annealing performed at high temperatures, while during "maturity", the temperature is gradually reduced, stochastic search becomes feeble (so unfavorable perturbations become less likely), and natural selection takes over. The hybrid strategy, corresponding to competition and selection subject to conservation of total population, yields better results than pure simulated annealing for problems such as the Travel Salesman Problem (TSP) (Boseniuk and Ebeling, 1988).

What about predicting or forecasting phenomena whose behavior is modeled by non-linear dynamical systems, and starting from some initial conditions? If everything was predictable from the very beginning, this would be a large blow to the very notion of free will. Sensitive dependence on initial conditions makes it impossible to make long term predictions because non-linear dynamical systems separate initially closed in space/time trajectories exponentially fast. The effect, known as the "butterfly effect", provides access to novel behavior. Nonlinearity is necessary but not sufficient to make dynamical systems unpredictable. The unpredictability is due to random elements, and the corresponding fluctuations and the resulting behavior is known as chaos. Chaotic behavior prevents us from associating cause and effect. Still, chaos is deterministic, and it can result from rules where chance does not play any role. Chaos can organize randomness and impose structure where none would be available otherwise.

Self-organization, to be encountered later on in the context of neural system assemblies, is complementary to selection and has the potential to enhance the scope of evolution by increasing the intrinsic complexity across individual agents. Kaufmann (1991) and others have suggested that adaptation takes place at the boundary between order and chaos and that evolution propels complex adaptive systems towards that very boundary. Phase transitions are well known in physics for leading to qualitative behavioral changes; transitions from order to randomness could play a similar role for evolutionary change in terms of information contents. Self-organization manifests itself also as self-organized criticality (Bak *et al.*, 1988), which is characteristic of critical states whereby minor events hold the potential for starting chain reactions ("avalanches") and strike thus disaster ("black Mondays") for complex systems such as economic markets. Some connection between chaos and self-organized criticality comes through the (l/f) flicker noise characteristic of a system whose dynamics are influenced by past events and of fractal (self-replicating) behavior.

II.3. Machine Intelligence and Learning

An important distinction needs to be made between those situations where the optimal response is known and where learning then amounts to the system merely being able to reproduce it, vs. those situations where the optimal response has yet to be identified. Supervised learning with a teacher corresponds to the first case, while self- organization, *i.e.*, unsupervised learning, corresponds to the latter case. Supervised learning, characteristic of error-correction strategies, seeks system equilibrium through negative feedback, and corresponds to homeostasis as defined earlier. Self-organization, characteristic of processes involving (novelty) discovery using extreme-searching strategies, is much more important for evolution, and corresponds to heterostasis.

Selection presumes categorization and categorization allows for adaptation within some given environment. Categorization methodology covers many approaches, ranging from essentialism to polymorphism, and it entails usually several (4 to 6) layers of hierarchical taxonomy. "The selective point of view maintains that anatomical variability is unavoidable from a developmental point of view and, moreover, is essential to the functioning of the system – for variability provides the substrate upon which selection acts. The nervous system probably continues to generate functional variability throughout the lifetime of the organism. Neuronal groups exemplify another main principle of neuronal group selection, that of degeneracy. Degeneracy endows the nervous system with a great deal of combinatorial richness and allows selection to guide the evolution of the network through the adaptive landscape" (Reeke *et al.*, 1990). Degeneracy goes beyond recessive genes, phenotype changes are occurring, and a revised Lamarckian hypothesis for economics concerned with inheriting successful ("neural") assemblies of ideas and innovation becomes possible. As it has been the case with the polymorphic approach for categorization, the case can also be made for a manifold of solutions, and the overall design is that of distributed "computation" (information processing) and adaptation.

The very existence of the nervous system, a higher qualitative stage of development on the evolutionary scale, suggests that nature, aware of its shortcomings, like lacking the ability to forecast future contingencies, has also allowed for the possibility of reactive planning and control (using adaptation, selection and synaptic plasticity) for when the need shall arise. Learning and adaptation can still be primed across eons through prewired drives and thus reflect innate, evolutionarily determined behavioral biases. For animals, *de novo* synthesis of new responses is critical for survival (Reeke *et al.*, 1990). *Tabula rasa* learning, "model-free" estimation, is slow to converge and requires prohibitively large training sets in order to reduce the variance contribution. The only way to control the variance in complex inference problems is to use model-based estimation, which is biased-prone (Geman *et al.*, 1992). The tradeoffs between bias and variance are well known and can be thought also in terms of variations of interpolate vs. extrapolate schemes. The means to overcome this dilemma and the limitations it imposes on performance should involve some limited bias suggested as boundary conditions by the genotype. Also characteristic of the nervous system is the high degree of reentry, whereby dynamic and integrative processes tie together seemingly independent processes. It is within this very context that the Lamarckian hypothesis becomes plausible as an example of evolutionary change. Rational beings ("organizations") in charge of their "own" fate have to innovate continuously in order to survive and the way to do that, usually within a short span of time, is by using hill climbing methods and analogical reasoning. Major breakthroughs or "paradigm" shifts, such as the industrial revolution or the advent of the information age, are quite rare and they correspond conceptually to punctuated equilibrium as described earlier.

Artificial intelligence (AI) (Winston, 1992), connectionism ("neural networks") learning (Hertz *et al.*, 1991), and their hybrids, are characteristic for machine intelligence and learning. The background for adaptation comes from symbolic (Weiss and Kulikowski, 1991) and fuzzy logic (Klir and Folger, 1988), statistics and estimation, control, and information theory. Connectionist learning (Rumelhart and McClelland, 1986; Anderson *et al.*, 1988 and 1990) is particularly suited to handle analytical tasks involving numerical information but too complex to render themselves to closed form solution. Collective behavior takes the form of distributed and hierarchical computation and, by using realistic assumptions, it can simulate basic tasks such as functional approximation, (stimulus-response) mappings and generalization tasks, and competitive and reinforcement learning. Performance evaluation in terms of accuracy and efficiency, compactness of representations using entropy concepts, capability to generalize beyond given training sets, and incremental learning and forgetting are basic for assessing amongst the different techniques proposed so far.

III. EVOLUTIONARY COMPUTATION AND GENETIC ALGORITHMS

Evolutionary computation mimics what nature has done all along and it does that using similar principles. There are three basic classes of evolutionary algorithms and they involve genetic algorithms (Holland, 1975) using Mendelian genetics, Darwinian algorithms using mutation/selection dynamics (Rechenberg, 1973; Edelman, 1987), and replicator (differential) equations based on the Fisher – Eigen theorem. Evolutionary computation is characteristic of "weak" AI methods, and it is usually used when a strong domain theory is lacking. It is obvious that further crafting of such methods using specific domain knowledge can only improve their performance. We briefly survey below evolutionary computation and consider some of the major issues involved in.

The class of genetic algorithms includes adaptation, optimization and selection techniques that maintain a constant-sized population of candidate solutions, known as individuals. The initial seed population can be chosen randomly or on the basis of heuristics, if those are available for a given application. The number of offsprings for some candidate solution is proportional to its fitness relative to the rest of population. The power of a genetic algorithm lies in its ability to explore and exploit, in a highly efficient manner, information about a large number of individual solutions. Candidate solutions are encoded using either binary notation, for simplicity, or real numbers. The search underlying GAs is such that breadth and depth are balanced according to observed fitness. By allocating more reproductive occurrences ("chances") to above average individuals, the overall effect is to increase the population's average fitness. New individuals are created using mostly two genetic operators known as *crossover* and *mutation*. Crossover operates by selecting a random location in the genetic string of the parents (crossover point) and concatenating the initial segment of one parent with the final segment of the other parent to create a new child. A second child is simultaneously generated using the remaining segments of the two parents. Mutation provides for occasional disturbances in offspring by inverting "bits" if binary encoding are employed or by slightly altering one or more genetic elements using random fluctuations (from some normal distribution) during reproduction if real numbers are used. Mutation ensures diversity for genetic strings over the long haul and it thus prevents stagnation in the evolution of optimal solutions. Note that performance of GAs is improved if Monte Carlo methods are used to evaluate fitness, and that the class of GAs belongs to simulated annealing when the population size is one.

Some of the major problems affecting genetic algorithms and their possible solution include poor fitness scaling possibly compensated by ranking, and slow convergence and premature convergence possibly overcome by parallel runs of the algorithm. Specifically, Muhlenbein and Kindermann (1989) have used parallel genetic algorithms for combinatorial optimization whereby selection is done locally in a neighborhood and each individual (phenotype) is active (rather than passive) using hill climbing ("Lamarckian adaptation") so recombination is eventually performed in the space of genotypes representing a local fitness maxima. The conclusions drawn suggest that premature convergence can be avoided, that "evolution with small isolated groups of individuals and migration is faster than in large (nonhomogeneous) populations with random mating", that "local inbreeding within a species has extremely important evolutionary consequences, too close inbreeding leads merely to extinction, and some crossbreedings is favorable but not too much", and, relevant to the Lamarckian hypothesis, that "genotype and phenotype learning seem to be equally powerful, but genetically specified *a priori* conditions give individuals a head start in solving complex problems." The observer affects the measurement by the very process of measurement – if parallel competitive and reinforcement schemes of evolution are considered there appears to be support for the Lamarckian hypothesis. There is also support for the fact that a population having spatial structure, restricted mating rules and increased selective pressure, yields more diversity and "percolates" faster than a random but less robust ("panmictic") population. The explanation is that "the creative forces of evolution take place at migration and few generations afterwards." One can expand on migration using diffusion processes in order to further improve fitness.

Another aspect bearing on genetic algorithms is related to their internal structure and credit assignment. Individual solutions are ranked once fitness is measured but the question still begs as to whom the merit (or blame) should be assigned. The degree of dependency between genes, known as epistasis, is the major factor affecting credit assignment, and has led to many alometric studies. High epistasis, characteristic of polygeny, makes credit assignment difficult, and allows for "bad" mutations to survive as "hitchhikers" due to their location being adjacent to that of "successful" genes. The credit assignment problem for classifier systems has been approached using the bucket brigade algorithm (Holland, 1975). The algorithm operates in analogy to economical markets where goods ("genetic material") are bid and exchanged by traders between manufacturers (environment) and consumers (Goldberg, 1989). It has been pointed out that "when there are epistatic fitness interactions, sexual reproduction can actually slow down evolutionary progress by breaking up co-adapted groups of genes" (Maynard-Smith, 1982).

Epistasis is closely related to representations. For representations exhibiting low epistasis "greedy" (hill climbing) algorithms should be more appropriate, while for high epistasis "random" algorithms are best. The Lamarckian model has been suggested by Davidor (1991) for possibly coping with epistasis, as defined above, by estimating the degree of co-adaptation between adjacent elements. The Lamarckian hypothesis, using the AI jargon, achieves sub-goal reward through the usual hill climbing method where "local" errors are estimated and regions that require further co-adaptation are identified. If epistasis is high, the question still begs on how to infer actual performance using only local estimates. One could possibly conjecture that the Lamarckian hypothesis operates then at the level of small island populations and precedes migration and major selection.

Darwinian algorithms, known also as evolutionary systems (ES), are described as (a/ b,c) and/or (a/ b+c), where a is the number of parents during some generation ("cycle"), b is referred to as the "mixing" number (if two parents mix their genes b=2), and c is the number of children. Selection is deterministic (vs. random for GAs) and is done so only the a best of the individuals are allowed to produce offspring. A comma "," indicates that parents are not included in the selection ("pure strategy") and it is appropriate for dynamical and noisy situations, while a "+" indicates that parents are included ("elitist strategy"). The range of mutation ("stepsize") is not fixed, but inherited, and it will be tuned by the evolution process itself (Muhlenbein and Kindermann, 1989). For Darwinian algorithms, mutation is more important than crossover because once the population starts to converge what is needed is minor tuning rather than major changes likely to destroy the solutions derived so far. Mutation is also more effective in small population and as a consequence parallel algorithms are especially suited for the Darwinian algorithms.

Specific questions have been raised about what the genetic architectures and strategies should look like. As it seems to be the case for the nervous system, Muhlenbein and Kindermann (1989) consider two conflicting viewpoints, one suggesting that higher functions are emergent properties of flat but highly connected networks, and the other proposing that higher functions result from specific substructures and connections genetically determined. "For fixed environments there is a clear disadvantage for large flat organizations. However, if the environment were to change radically, then these architectures would have an advantage because they are capable of mutating quickly into large hierarchical structures."

IV. THE LAMARCKIAN HYPOTHESIS REVISITED

The Lamarckian hypothesis becomes an alternative to reckon with only if it is possible to pass on to future generations information acquired through phenotypical adaptation. Heterostatic behavior at the phenotype level, characteristic of agents seeking explicit information within the confines imposed by the genetic pool, can be easily implemented using hill climbing fit for a particular problem domain. While it is obvious that inheritance imposes specific bounds on behavior, it is also clear from ethology that such bounds may be wide open to allow for adaptation (Gould, 1982) because time and memory are limited. As it has originally been suggested by Lorenz the genotype determines patterns of growth subject to the modifying effects of experience (Muhlenbein and Kinderman, 1989). It is also obvious that the effects of such experience have to be encoded somehow and as a consequence it should not take a great leap of faith to subscribe to the notion that informative coding can be passed on. Based on the "fossil" record (coming from paleontology or from cases of failed industries) the case could also be made that the Lamarckian hypothesis provides means for testing different phenotypes before closing on genotype solutions, and it helps thus with preparing the code for higher, yet to emerge, structures.

Some discussion should clarify in conceptual terms the usefulness of the Lamarckian hypothesis for genetic algorithms (GAs) and how to implement it. As for most GAs crossover reproduction and mutations are random, the potential for decrease in fitness is high due to epistasis. While natural evolution appears to be irreversible the same case does not hold necessarily for economical performance. Hill climbing using local error estimates in fitness as feedback control should be in principle able to locate genetic substrings prone for further improvement without disrupting those substrings already adapted. In terms of development and life cycles, the Lamarckian hypothesis could be exercised as an option at the stage when the genotype is expressed and it assumes its own identity (epigenetics) and/or during the early stages of development ("youth"). It is the expression of the genotype and the elements bearing on such expression that could be possibly affected by Lamarckian factors. Much innovation can be achieved by new and usually small companies; mutations are frequent while the incentive for reproduction (spinoffs) is almost non-existent. For mature companies the reverse holds true, mutations are seldom while reproduction (selection) occurs frequently. Another possibility to account for the Lamarckian hypothesis would be that acquired characteristics are passed on to the next evolutionary level rather than to the same level the agent belongs to. Structural changes taking place could be possibly passed onto the more advanced but surviving phylum; as an example, transitions from the soon to be extinct dinosaurs were passed on to the birds. Finally, another possibility to contend with is that the Lamarckian hypothesis operates at the level of ecosystems and that the medium of exchange, that of knowledge, is now changing at ever increasing speeds, and the inherent limitations characteristic of hill climbing can be overcome.

V. CONCLUSIONS

This paper provides a perspective on the Lamarckian hypothesis using artificial worlds and evolutionary computation. Artificial worlds are large scale artifacts useful to simulate and visualize collective behavior. The goal of simulation is to both predict future behavior and to possibly prevent it from emerging if deemed harmful. New conjectures can be tested for and the role that information plays in evolution and innovation can be properly assessed. Instead of exploring merely how markets work, artificial worlds can probe why the markets are the way they are and how would they possibly evolve in the future. Topics of particular interest for simulation studies, beside the Lamarckian hypothesis, include positive feedback, how individual rationality is molded by the rules of the games into aggregate rationality, top-down vs bottom-up strategies, recombination (high diversity but an expensive proposition) vs. asexual reproduction (fast proliferation of offspring), government intervention vs. *laissez faire*. Qualitative arguments seem to suggest that factors such as knowledge transfer, hierarchical structures and development life cycles involving hybrid ("genotype" and "phenotype") learning could possibly support some modern version of the Lamarckian hypothesis.

References

D. H. ACKLEY, G. E. HINTON, T. J. SEJNOWSKI, A learning algorithm for Boltzmann machines, *Cognitive Sciences*, Vol. 9, 1985, p. 147-169.

R. AXELROD, *The Evolution of Cooperation*, Basic Books, 1984,

J. A. ANDERSON, E. ROSENFELD, Eds., *Neurocomputing*, MIT Press, 1988.

J. A. ANDERSON, A. PELLIONISZ, E. ROSENFELD, Eds., *Neurocomputing*, 2, MIT Press, 1988.

W. B. ARTHUR, On learning and adaptation in the economy, TR 92-07-038, Santa Fe Institute, 1992.

P. BAK, C. TANG, K. WIESENFELD, Self-organized criticality, *Physical Review A*, Vol. 38, 10, 1988, p. 364-374.

T. BOSENIUK, W. EBELING, A. ENGEL, Boltzmann and Darwin strategies in complex optimization, *Physics Letters A*, Vol. 125, 6-7, 1987, p. 307-310.

T. BOSENIUK, W. EBELING, Optimization of NP-complete problems by Boltzmann-Darwin strategies including lyfe cycles, *Europhys. Lett.*, Vol. 6, 2, 1988, p. 107-112.

J. P. CHANGEUX, *L'Homme Neuronal*, Fayard, 1985.

J. COWAN, Game of life, in E. BERLECKAMP, J. CONWAY, R. GUY, Eds., *Winning Ways*, Academic Press, 1982.

Y. DAVIDOR, Lamarckian sub-goal reward in genetic algorithms, in AIELLO L. C., Ed., *Proc. of the 9th European Conference on Artificial Intelligence*, Stockholm, Sweden, 1991, p. 189-194.

L. DAVIS, Ed., *Genetic Algorithms and Simulated Annealing*, Morgan Kaufmann, 1987.

L. DAVIS, Ed., *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, 1991.

G. EDELMAN, *Neural Darwinism*, Basic Books, 1987.

S. FORREST, Ed., *Emergent Computation*, MIT Press, 1991.

S. GEMAN, E. BIENENSTOCK, R. DOURSAT, Neural networks and the bias/variance dilemma, *Neural Computation*, 4, 1992, p. 1-58.

D. E. GOLDBERG, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.

J. L. GOULD, *Ethology*, W. W. Norton, 1982.

S. J. GOULD, *Wondeful Life: The Burgess Shale and the Nature of History*, W. W. Norton, 1989.

R. L. GREGORY, Ed., *The Oxford Companion to the Mind*, Oxford University Press, 1987.

J. HERTZ, A. KROGH, R. G. PALMER, *Introduction to the Theory of Neural Computation*, Santa Fe Institute Studies in the Sciences of Complexity, Addison-Wesley, 1991.

W. D. HILLIS, Co-evolving parasites improve simulated evolution, in S. FORREST, Ed., *Emergent Computation*, MIT Press, 1991, p. 228-234.

J. HOFBAUER, K. SIGMUND, *The Theory of Evolution and Dynamical Systems*, Cambridge University Press, 1988.

J. H. HOLLAND, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, 1975.

J. H. HOLLAND, J. H. MILLER, Artificial adaptive agents in economic theory, TR 91-05-025, Santa Fe Institute, 1991.

S. KAUFMANN, *Antichaos and adaptation*, Scientific American, 1991.

G. J. KLIR, T. A. FOLGER, *Fuzzy Sets, Uncertainty, and Information*, Prentice Hall, 1988.

H. A. KLOFT, *The Hedonistic Neuron*, Hemisphere Publishing, 1982.

D. A. LANE, Artificial worlds and economics, TR 92-09-048, Santa Fe Institute, 1992.

D. MARR, *Vision*, W. H. Freeman, 1982.

J. MAYNARD SMITH J., *Evolution and the Theory of Games*, Cambridge University Press, 1982.

H. MUHLENBEIN, J. KINDERMANN, The dynamics of evolution and learning. Toward genetic neural networks, in R. PFEIFER, Z. SCHRETER, F. FOGELMAN-SOULIE, L. STEELS, Eds., *Connectionism in Perspective*, Elsevier Science, 1989, p. 173-197.

R. PENROSE, *The Emperor's New Mind*, Oxford University Press, 1989.

I. RECHENBERG, *Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution*, Frommann-Holzboog Verlag, 1973.

G. N. RECKE, L. H. FINKEL, G. M. EDELMAN, Selective recognition automata, *in* S. F. ZORNETZER, J. L. DAVIS, C. LAU, Eds., *An Introduction to Neural and Electronic Networks*, Academic Press, p. 203-226.

E. RICH, K. KNIGHT K, *Artificial Intelligence*, McGraw-Hill, 1991.

D. E. RUMELHART, J. L. MCCLELLAND, Eds., *Parallel and Distributed Processing*, Vols. 1 and 2, MIT Press, 1986.

J. SEARL, *Minds, Brains, and Science*, Harvard University Press, 1984.

H. SIMON, *The Science of the Artificial*, MIT Press, 1982.

S. M. WEISS, C. A. KULIKOWSKI, *Computer Systems That Learn*, Morgan Kaufmann, 1991.

E. O. WILSON, *Sociobiology*, Harvard University Press, 1982.

P. W. WINSTON, *Artificial Intelligence*, Addison-Wesley, 1992.

REVUE INTERNATIONALE DE SYSTÉMIQUE
Vol. 7, N° 5, 1993, pp. 593 à 612

APPRENTISSAGE, TEMPS HISTORIQUE ET ÉVOLUTION ÉCONOMIQUE

Bernard ANCORI[1]

Résumé

Une formalisation adéquate de l'état et la dynamique d'un système économique complexe, défini dans les catégories de l'information et de la communication permet de conférer à l'évolution économique les principales caractéristiques du temps historique : contingence, accumulation en une mémoire constructiviste pour laquelle l'oubli, loin de se réduire à une perte d'information, joue un rôle structurant.

Abstract

The state and the dynamics of a complex economic system, expressed in terms of information and communication, endow economic evolution with the main features of historical time. Contingencies, irreversibility and cumulative forces create an active memory in which forgetfulness is not reducible to a pure loss of information and fulfils a structural role.

Les entités que nous avons l'habitude de désigner globalement par l'expression d'« agents économiques individuels » possèdent en réalité une triple dimension : ce sont d'abord des *individus*, considérés sous l'angle des motivations indivisibles et singulières de leurs décisions ; ce sont ensuite des *sujets*, qui agissent souverainement à l'intérieur d'espaces que leur ménagent diverses contraintes et sont responsables des conséquences de ces mêmes décisions ; ce sont enfin des *personnes*, qui renferment l'ensemble organisé de pratiques et d'attitudes conférant intériorité et unité à ces individus-sujets. Individu, sujet, personne : la réunion de ces trois dimensions est le produit d'une très longue évolution historique qui détermine simultanément l'apparition du concept d'agent économique (L. Dumont,

1. Centre de Recherche en Sciences Économiques et Gestion (BETA), Université Louis-Pasteur, 38, boulevard d'Anvers, 67070 Strasbourg Cedex.

Rev. intern. systémique. 0980-1472 Vol. 7/93/05/$ 4.00/© Afcet Gauthier-Villars