

Viability Evolution: Elimination and Extinction in Evolutionary Computation

Claudio Mattiussi and Dario Floreano

April 2003 - Technical Report

Laboratory of Intelligent Systems, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Abstract—Current mainstream Evolutionary Algorithms (EA) are based on the concept of selection, encapsulated in the definition of a fitness function. Besides selection, however, the natural evolutionary process includes a phenomenon of elimination, which is linked to the ideas of viability and contingency. In this paper, we show how elimination can be modeled and integrated with selection to give rise to a new evolutionary scheme that we call Selection-Elimination EA (SE-EA). Comparing conventional EA to the newly defined SE-EA we show that SE-EA can exploit naturally some prior information about the problem at hand that is seldom, if at all, exploited by conventional EA, while avoiding the assumption of knowledge that EA based on the fitness function require but that is usually not really available. We discuss the fact that the introduction of elimination in Evolutionary Computation gives rise to the possibility of obtaining a multi-level evolutionary process that includes as a central component the phenomenon of extinction. We suggest the interpretation of an evolutionary process that includes elimination, in terms of an error-driven process and derive from it a new appreciation of the role of the interaction with the environment in the determination of the outcome of the process and of the possibility of achieving an open-ended evolution. The working of the SE-EA are illustrated with two examples that model its application to multi-objective engineering problems.

I. INTRODUCTION

THE methods of Evolutionary Computation (EC) [9], [10], [44] are the result of an attempt to distill in algorithmic form the workings of the process of biological evolution. The ultimate goal is the development of a set of engineering and problem-solving tools displaying the adaptive and creative potential usually attributed to the evolutionary process operating in nature. Many evolutionary algorithms (EA) inspired to this approach have been developed and are currently in use within the EC community, for example, Genetic Algorithms, Evolutionary Strategies, and Genetic Programming [8]. The ideal starting point for the development of EC method is obviously our current understanding of the evolutionary

process and its formalization in terms of an evolutionary theory. Although many aspects of the workings of the evolutionary process are still controversial [50], there is a general consensus on attributing a central role in it to the interplay of variation, inheritance, and differential reproductive success [20], [69].

Differential reproductive success in nature is the result of multiple factors that operate at many levels of organization to influence the probability of survival and reproduction of the individuals and of their progeny. Since biological organisms are, in general, self-reproducing autonomous entities, the differential reproductive success of individuals is an implicit consequence of their existing and behaving in their environment. Due to technological limitations, in present-day EC experiments the individuals are never truly self-reproducing autonomous agents. Therefore, their differential reproductive success cannot be obtained simply as a consequence of their existence and behavior. This means that the experimenter is forced to intervene in the process to explicitly determine - at least partially - the reproductive success of the individuals. The presence of this explicit intervention entails the risk of formulating algorithms that do not mirror the actual workings of the natural evolutionary process and are unable to achieve its efficacy.

Of the many possible discrepancies between the structure of the evolutionary process and that of its algorithmic implementations we consider here the case of the explicit elimination of the individuals subjected to the action of the algorithm. We show that this aspect is usually omitted in current mainstream EA formulations and try to expose the reasons of this state of things. Then, we investigate the consequences that this absence produces on the potential of the resulting algorithms. Particular emphasis is put on two improvements that the implementation of explicit elimination can bring to EC: a substantial enlargement of the range of evolutionary experiments that can be performed and of the causes of differential reproductive success that can be modeled, and the possibility to obtain a multi-level evolutionary

process. To remedy the neglect of elimination in EC we introduce a new EA structure that explicitly implements elimination and integrates it with the selection and reproduction process that is usually modeled in current EA. Finally, we discuss the properties of this extended algorithm and present an example of its application to multi-objective problems.

Besides elimination there are in the natural evolutionary process other aspects that are likely to be instrumental to its achieving the results that we observe in the biological world and that are poorly if at all represented in EC models. For example, we can mention the role played by the structure and dynamics of the genome [68] in the production of variation; the influence of the development process [46], of physical constraints, of self-organization phenomena, and of the structure of the cellular context where the genome is expressed [59], [83]; the existence of interactions allowing the emergence of new levels of organization [21], [77], [81]. Adhering to our announced theme, we will not consider here these further aspects, although we are persuaded of the importance of considering also their possible role while building computational models of the evolutionary process.

II. THE STATUS OF ELIMINATION

Today, almost all EC algorithms are built around the idea of selection [8], [43], [47]. According to this idea the individuals of the population prevailing at a certain time step are allowed to enter the composition of the population for the next step in a way that is related to the value of the so-called fitness function. Usually, this function associates a real value with each individual and, on average, individuals with higher associated values contribute with more genetic material to the definition of the new population than those having lower values. This process is supposedly modeled on the workings of natural selection. However, Ernst Mayr notes [78, pp.117-118]:

What Darwin called natural selection is actually a process of elimination. . . . Do selection and elimination differ in their evolutionary consequences? This question never seems to have been raised in the evolutionary literature.

Emphasis on the distinction between selection and elimination is indeed rare in the life sciences literature. One notable exception are the works of P.J. den Boer [34]–[37] articulating the idea that natural selection is better seen in terms of ‘non-survival of the non-fit’ rather than in terms of ‘survival of the fittest’.

Browsing the EC literature one is readily convinced that the elimination concept is seldom considered also

within the EC community. Some exceptions are Atmar’s broad methodological discussion on the simulation of evolution [3]; the Culling Method of Baum, Boneh, and Garrett [14], where a culling threshold is established and the individuals having a value of fitness below the threshold are systematically substituted with above threshold newly generated ones; the series of papers by Greenwood, Fogel, *et al.* [45], [51], where a model inspired to mass extinction is applied to function optimization; the Macroevolutionary Algorithms of Marín and Solé [74], based on the Solé-Manrubia extinction model [101], where a matrix of species interactions is built and used to determine species extinction; and Krink and Thomsen’s approach [63] inspired to the Bak-Sneppen Self-Organized Criticality model of evolution and extinction [97]. None of these approaches, however, adopts the multi-level elimination perspective nor the idea of a fluctuating and possibly extinct population that we deem instrumental in modeling correctly elimination and that we are going to champion in the present work. Besides the explicitly EC-oriented elimination approaches listed above, there is – with a tacit EC spirit – the seminal suggestion constituted by Braitenberg’s vehicles that “live” on the surface of a table, are reproduced with possible mutations, and are eliminated if they fall from the table [19].

The virtual absence of the elimination concept from the mainstream EC thinking is probably a consequence of its absence from the evolutionary literature lamented by Mayr, since the founders of EC were undoubtedly inspired by the biological formulations of evolutionary theory in laying the basis of their algorithmic counterparts. To understand the reasons of the scarce explicit presence of elimination in the evolutionary debate we must delve into the historical development that led to our current view of the evolutionary process.

III. NATURAL SELECTION

In the first four chapters of *On the Origin of Species* [30] Darwin develops an argument that from the analysis of the process of artificial selection leads to the attribution of an adaptive and creative role to the process of natural selection. Darwin starts by illustrating the existence of a great deal of inheritable and continuously renewed variation within the populations of domestic animals and plants and proceeds to explain how this fact was exploited by breeders to produce the existing domestic races. Darwin identifies the source of this accomplishment in “man’s power of accumulative selection”, embodied in the breeders’ uncanny ability to identify minimal individual variations in the desired

direction within the raw material constituted by the variation existing in the population subjected to selection.

Then, Darwin's efforts are devoted to the illustration of the fact that a large amount of inheritable and continuously renewed variation is also a characteristic of "organic beings in a state of nature". In other words, Darwin recognizes in the "state of nature" the presence of all the elements that preside over the successful creation of domestic varieties except that of the actively selecting agent represented by the breeder. With this in mind, Darwin proceeds to analyze the effects of the interaction of two observed facts: the potentially exponential rate of growth of a population and the long term stability of most populations in nature. The compresence of these two factors implies that a great part of a natural population offspring is necessarily doomed to destruction and that the individuals of those populations must therefore be engaged in a perpetual struggle for existence. If some inheritable¹ characteristics of the individuals exist that are apt to increase their probability of survival and reproduction, a selection based on the degree of ownership of those characteristics will take place. Darwin succeeds in this way to prove that an agentless selection process² exists and operates in nature.

Note that, as remarked by several authors [24], [50], [108], many thinkers before Darwin had recognized the existence of the process of natural selection. However, most of them attributed to natural selection a purely conservative and stabilizing role of elimination of the badly formed or unfit individuals. According to Gould [50] the novelty of Darwin's approach was his attributing to natural selection an adaptive and creative role, analogous to that played by the breeders in the production of existing domestic races. In other words, Darwin revolution shifted the focus from the conservative role played by elimination of the unfit to the potentially creative role played by selection of the fittest.

IV. THE FITNESS SHORTCUT

Although since Darwin's times many elements have been added to our understanding of the evolutionary process, the general picture remains substantially unchanged. Note that at first sight there is an almost complete match between the processes of artificial selection and that of natural selection. The degree of possession of the character desired by the breeder in artificial selection apparently corresponds to the degree

of possession of the characters favourable to reproductive survival in the state of nature. However, we see that the correspondence breaks down when we consider the possibility of ordering the individuals according to those criteria. In the case of artificial selection we can ideally assume that the breeder proceeds by first ranking the individuals and then allowing only the best to reproduce. In the case of natural selection, however, the variety of circumstances that lead to reproductive survival is usually immense and it appears generally impossible to estimate *a priori* (i.e., by looking only at their phenotype or genotype) the reproductive success of individuals. At most we can think of ranking the individuals *a posteriori*, that is, according to their observed reproductive success and assuming that it reflects the degree of possession of favourable characters (possibly, as we will discuss later on, at multiple levels of organization).

The idea of focusing on a measure of the reproductive success is actually at the heart of the concept of *fitness* as usually defined in evolutionary biology [28], [76], [81], [91]. For example, the fitness of a genotype can be defined as the expected number of offspring contributed by an individual of that type to the next generation [76, p. 36]. This definition suits perfectly the analysis of the dynamics of genotype frequencies within gene pools. On the other hand the problem of associating a value of fitness with a genotype, apart in very rare circumstances or as an *a posteriori* observation³, is not addressed. Recapitulating, the biologist's fitness is a value that is assumed to represent the collective effect of the myriad of factors that determine the survival and reproductive success of an individual.

EC has adopted from the start this convenient shortcut by basing the workings of its algorithms on the concept of fitness function. As mentioned above, this is a function that associates a numerical value with each individual of each generation considered in the course of an EC experiment. This value then determines (possibly probabilistically) the reproductive success of the individual. This fact justifies calling the numerical value the fitness of the individual. Note, however, that the logic⁴ is reversed relatively to the biologist's definition. There, the reproductive success was assumed as a measured numerical value and called fitness⁵; here, a numerical value is given and steps are taken to ensure that the expected reproductive success assumes that value. A moment of reflection suffices to convince that this reversal does not alter the dynamics of the genotype frequencies.

¹Inheritability is required for the selection process to lead to an evolutionary process, not for the existence of the selection process in itself.

²Darwin considered also at great length the effects of the agent-mediated selection process of *sexual selection* [31].

³And even in that case, with some nontrivial open questions [27].

⁴Or, better, the direction of the chain of causation.

⁵In fact, the observed value is an *estimate* of the fitness defined as the expected value of reproductive success [76], [98].

The central question for EC, however, is if the reversal is justified from the point of view of the outcomes that we can expect from the resulting process in terms of adaptation and evolution. In other words, can we assign directly what in nature is the result of a complicated ensemble of factors and still maintain the essential characteristics of the process? To answer this question we need to examine the nature of the factors that eventually determine the survival and reproductive success of individuals in nature to ascertain if in the current EC approach the structure of the process is modeled with sufficient fidelity, at the correct level of organization, and focusing on the right subjects, for the adaptive and creative potential of the original process to be preserved. This means that we need to examine not only the outcome of the evolutionary process but also how the process works [93], [94], [98].

V. ELIMINATION AND SELECTION

Semantically, given a finite set of elements, the elimination of the elements of a subset corresponds to the selection of the elements of the complementary subset, and vice versa. Therefore one may wonder about the meaning of Mayr's abovementioned distinction of a processes of selection from one of elimination. Once again we can turn to Darwin for inspiration, and in particular to his clarifications concerning the meaning of the expression "struggle for existence" [30, p. 116]:

I should premise that I use the term Struggle for Existence in a large and metaphorical sense Two canine animals in a time of dearth, may be truly said to struggle with each other which shall get food and live. But a plant on the edge of a desert is said to struggle for life against the drought, though more properly it should be said to depend on the moisture.

These two examples of struggle for existence differ in one important point. In the first case there is a direct comparison of the aptitudes of the two animals in contending for the available food, whereas the resistance of the plant to drought is not compared with that of another plant to determine if it must survive or not.

We can reformulate and generalize this distinction as follows. There are cases where the individuals of a population or of a group are ordered according to some criterion (for example, fecundity, or the outcome of direct fights) and their relative reproductive success is determined by that ordering. In other cases, however, no ranking of individuals is implied, since each individual faces and must survive a challenge that is "absolute" in the sense that the outcome does not depend on how

well the individual fares relatively to other individuals confronted with the same challenge. Following Mayr's usage and to avoid clumsy circumlocutions we will call the first kind of events *selections* (or *selection events*) and the second kind *eliminations* (or *elimination events*). The term *evolutionary event* will be used to denote both kinds of events⁶.

The cases where the ability of one individual in accessing a resource influences indirectly its availability to other individuals appear at first sight ambiguous relatively to the classification just introduced. However, if there is no direct comparison of individuals, each one is actually facing the "absolute" problem of the scarcity of a resource. We could, it is true, conceive a laboratory experiment aimed at determining the minimal amount of the resource required by each particular individual to stay alive and this would indeed allow the ranking of the individuals according to this characteristic. Nonetheless, the actual living conditions experienced by the individuals and the actual amount of the resource available to each of them depends on many contingent conditions and, therefore, the actual outcome of the struggle for existence does not correspond to that hypothetical ranking.

In this perspective, each individual of a population faces during its lifetime a sequence of evolutionary events, each of which can be an elimination or a selection. The rate of the two kinds of evolutionary events encountered by the individuals of a certain population will probably depend on the kind of social organization of the population. Nonetheless, a certain number of elimination events will be in any case experienced by the individuals of every population. In particular, we can assume that there is a continuous flow of elimination events that an organism must face, that constituted by the ongoing implicit assessment of its actual viability given the prevailing environmental conditions.

It is clear that the fitness function approach commonly adopted in EC, with its total ordering of the individuals, is appropriate only for the modeling of what we have called selection events. Therefore, the question raised above about the EC approach to the modeling of the evolutionary process can be reformulated as follows: can the (lack of a) separate representation of the elimination events make a difference on the setting and possible outcome of the modelled evolutionary process? We will

⁶The outcome of elimination events contributes to the differential reproductive success of individuals. Hence, elimination events and selection events are both parts of an overall selection process, whose ordering criterion is the measured differential reproductive success. Note that this usage of the term "selection" implies an ordering of individuals (but see [102] for a commentary on the conflation of the concepts of sorting and selection).

see that there are indeed multiple reasons suggesting the explicit presence of elimination events in EC models. Some of these reasons, however, can be appreciated only after realizing that an evolutionary process can take place at various levels of organization.

VI. EVOLUTIONARY ENTITIES

The evolutionary process was originally conceived as applying to individual organisms in a population. However, once the structure of the process is laid out it becomes apparent that it applies in general to entities that satisfy a small set of requirements. As expressed by Lewontin [69]:

The generality of the principles of natural selection means that any entity in nature that have variation, reproduction, and heritability may evolve. ... the principles can be applied equally to genes, organisms, populations, species, and at the opposite ends of the scale, prebiotic molecules and ecosystems.

We can say, therefore, that the evolutionary process applies in general to *evolutionary entities* [28]. We can broadly call those entities *individuals* no matter at what level of organization they are situated, and *populations* or *groups* the ensembles formed by individuals. This reveals the existence of a multi-level structure where what is a population or a group at a given level can be seen as individual at the next higher one. In this sense we can speak in general of *group selection* for evolutionary events taking place at the next higher level relatively to that of the entities assumed as individuals.

The issue of the natural evolutionary entities that can legitimately assumed to be the subjects of evolutionary events is an old and far from settled one. Many words have been spent on the controversy about the levels of selection and the relevance of group selection within the biological evolutionary process [37], [49], [60], [61], [69], [91], [98], [105], [106]. In particular, it has been argued that given a selection process operating at a certain level, a selection process operating at higher levels – although logically conceivable – has a much weaker effect and can be hardly considered a driving force of evolution [105] (but, for example, see [50] for a critique of this position). Despite this and other reservations, a certain consensus is emerging on the fact that evolutionary events at higher levels can influence the final outcome of the process [50], [60], [69], [72], [90], [100], [106]. Consequently, in modeling the evolutionary process we must be sure to allow the possibility of higher organization levels to exist, act as evolutionary entities, and make their presence felt at the

“bookkeeping” level [50]. We will come back later on this topic to motivate our claim that the presence of elimination in an EA is instrumental to the existence of that possibility (although, in itself, this presence does not automatically guarantee that this possibility is exploited by the implementations of the algorithm).

VII. MODELING ELIMINATION AND SELECTION

So far we have discussed in qualitative terms the characteristics of selection and elimination events. Our final goal is the definition of an EA implementing both kinds of event; the first step consists in building a mathematical model of them. The implementation of a selection event requires the definition of a function whose values determine the ranking of the individuals of the population. Let us call this function the *selection function* $\sigma : \mathbb{I} \rightarrow \mathbb{S}$, where \mathbb{I} is the set of possible individuals (phenotypes) and \mathbb{S} is a totally ordered set [66], for example, the set \mathbb{R} of real numbers. We assume that each phenotype $\iota \in \mathbb{I}$ derives from a genotype γ belonging to a suitable space Γ . The ranking of the individuals of the population determined by σ is used by the EA to reproduce (possibly) differentially their genotype. Note that the selection function alone does not determine what biologists call the fitness of the individuals, which will be instead determined jointly by selection – modeled through the selection function – and elimination – modeled through the viability function introduced below. For this reason we do not call σ the “fitness function”. Apart from this difference in terminology, the selection function plays the role of what is called fitness function in traditional EA [18], [47], [48].

A. Elimination through Viability

The case of elimination is less familiar when it comes to modeling it. Since we are dealing here with the survival of organisms, we can get some inspiration from the concept of homeostasis, introduced by Cannon [25] and widely used by Ashby [2]. The central idea is that to survive every organism must always maintain a set of physiological parameters within certain limits. We can imagine that for each parameter there is a *viability range* (or, more generally, a *viability set*) within which it must be maintained [4], [6]. Note incidentally that the maintenance of the physiological parameters within the viability set is not in itself the organism’s goal, this being rather its own self-preservation⁷ (i.e., the

⁷Laborit [65] presents an instructive account of how not many years ago the accepted medical practice was in some cases aimed at restoring some physiological parameter at its “normal” value even if, as it was later realized, the effect of this intervention was the exacerbation of a pre-existing physiological stress.

control of a relationship pertaining to a higher level of organization [13]). To model the idea of viability range we define a *viability function* $v : \mathbb{I} \rightarrow \mathbb{V}$ where \mathbb{V} is a *viability space* that, contrary to the case of \mathbb{S} , need not have an order relation defined in it. Then, we assume as given a *viability set* $V \subset \mathbb{V}$ and define as viable an individual $\iota \in \mathbb{I}$ if and only if its viability value lies within the viability set⁸, that is, if and only if $v(\iota) \in V$. This determines a set I of viable individuals within the space \mathbb{I} of individuals (Figure 1). At each time step of the EC experiment we can evaluate the viability of the individuals of the population and proceed to the elimination of non-viable individuals⁹. This will be an instance of an elimination event according to this model.

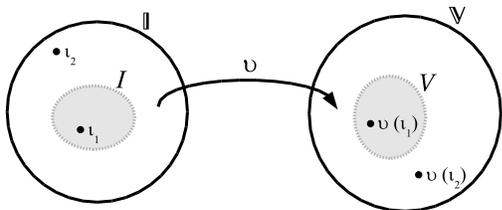


Figure 1. A diagrammatic representation of the viability space \mathbb{V} enclosing the viability set V , and of the viability function v assigning viability values to the elements ι of the space of individuals \mathbb{I} . In the case represented here individual ι_1 is viable since $v(\iota_1) \in V$ whereas individual ι_2 is not viable. Each individual (phenotype) ι derives from a genotype γ belonging to a space Γ of genotypes (not represented here).

To expose the rationale behind the definition of this model let us illustrate how it can be applied in practice. Let us consider a problem that could be the target of an algorithm implementing the model, for example, the design of an electronic circuit, say, an amplifier. In the context of such a problem a genotype γ will represent a symbolic description of the circuit and the individuals ι will correspond to the actual circuit described by γ [62]. The goal of the design is the determination of at least one circuit having an ensemble of characteristics (cost, amplification, noise, efficiency, component count, ...) within specified ranges, that is, a circuit with given specifications. For each characteristic c in the ensemble we can define a function $v_c : \mathbb{I} \rightarrow \mathbb{V}_c$

⁸Biologists often adopt a softer definition of viability (for example, the proportion of surviving individuals of a given type [99]). This could be modeled substituting fuzzy sets [80] to the crisp sets used in our definition, but at the risk of obtaining a model formally equivalent to the one based on the selection function defined previously.

⁹Note that this mechanism complies with the definition of *extinctive selection* within existing EA [8]. However, in current EA, extinctive selection – when applied – is based on the preliminary ordering of the individuals of the population, whereas the viability elimination just defined is not.

that associates with a circuit ι its value $v_c(\iota)$ of the characteristic (for example, the amplifier gain), where \mathbb{V}_c is a suitable space. The ensemble of admissible values for the characteristic is assigned as a set $V_c \subset \mathbb{V}_c$. It follows that a circuit ι is acceptable – i.e., viable – if $v_c(\iota) \in V_c$ for each characteristic c . This corresponds to the model introduced above, with the viability space \mathbb{V} defined as the Cartesian product of the spaces \mathbb{V}_c , the viability set V defined as the product of the sets V_c , and the functions v_c interpreted as the components of the viability function v ¹⁰. The set I of viable individuals then corresponds to the set of circuits complying with the given specifications¹¹.

B. Dynamic Viability

The model presented above is adequate for static scenarios, where the viability conditions do not change with time. An EA applying this model to our circuit design problem would be required to find a set of viable circuits by generating tentative circuits and eliminating those that do not comply with the given ensemble of static specifications. In practice, however – and irrespective of the details of the operation of the algorithm – it would be unwise to target directly a set of stringent specifications. It would be instead preferable to start with a suitably relaxed set of specifications and proceed to tighten them in the course of the algorithm’s execution. This view is corroborated by the example of technological evolution, where the accumulation of knowledge about past achievements is one of the factors leading to the improvement of attainable performances.

To model this variable viability set we can define a time parameter $t \in T$, where T is the set of time instants that enter the evolutionary experiment. Then, we can generalize the notation introduced above, moving from a static viability set to a time-dependent viability set $V(t) \subset \mathbb{V}$. The ensemble of all these sets constitute a *viability tube*¹² $\Theta = \{V(t)\}_{t \in T}$ defined within the *viability space-time*, $\mathbb{V} \times T$. An individual will then trace a trajectory $t \mapsto v(\iota)$ in $\mathbb{V} \times T$, and will remain viable as long as its trajectory stays within the viability tube [5] (Figure 2). Analogous structures can be defined in the space $\mathbb{I} \times T$.

This time-dependent definition enlarges substantially the range of problems to which the viability elimination

¹⁰More generally, the viability set can be determined by membership conditions on combinations of the functions v_c . Thus, V is not necessarily assigned as the Cartesian product of separately defined sets V_c .

¹¹From this perspective, the data sheet of a device contains the manufacturer’s description of its viability set.

¹²We can represent Θ also as a set-valued map $t \rightsquigarrow V(t)$ [5], [7].

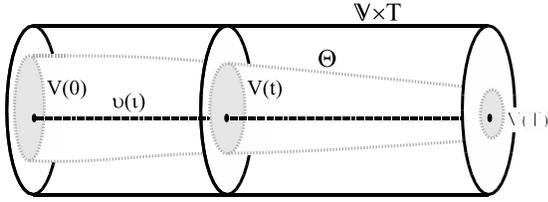


Figure 2. A schematic representation of the space $\mathbb{V} \times T$ enclosing the viability tube $\Theta = \{V(t)\}_{t \in T}$ and the viability trajectory $v(t)$ of an individual ι that remains viable during the whole time interval T .

model can be applied. However, there are still many applications of EC interest that do not fit well into it. An example is the design of a control system – say, a neural network – for a robot whose task is the realization of a behavior such as obstacle avoidance in the environment where it is placed. It would be most natural to declare non viable the robots that, for example, crash into obstacles or let their batteries discharge while away from the recharging station. It is clear that in this case the viability of an individual is determined not so much by the phenotype ι in itself as by the behavior that it brings about. This, in turn, depends also on the interaction of the robot with its environment.

In terms of our model, this means that the viability of the robot cannot be a function of its sole phenotype ι . We can assume that all the information that is relevant to the assessment of its behavior (structure of the phenotype included) is represented by a variable $\chi_\iota(t)$ that represents its state [15], [92]. The internal dynamics of the robot and its interaction with the environment will be reflected in the dynamics of the state¹³. Therefore, the robot’s viability can be represented as a function $v(\chi_\iota(t))$ of the state. This means that we now have a viability function of the kind $v : \mathbb{X} \rightarrow \mathbb{V}$, where \mathbb{X} is the state space. The behavior of the robot will translate into a trajectory $t \mapsto v(\chi_\iota(t))$ in $\mathbb{V} \times T$. As in the previous case, an individual will remain viable as long as its viability trajectory stays within the viability tube (Figure 3) and will be eliminated as soon as it abandons it. Note that this more general model of viability elimination can be thought of as including the two previously defined ones as special cases. A redefinition of the selection function as a function $\sigma : \mathbb{X} \rightarrow \mathbb{S}$ of the state of an individual is also possible and straightforward.

¹³For example, the state dynamics could be described by a dynamical system composed by the equations $\dot{\chi}_\iota(t) = f(\chi_\iota(t), E(t), t)$ and $\dot{E}(t) = g(\chi_\iota(t), E(t), t)$, where $E(t)$ represents the state of the environment [2], [15], [71]. This models in particular the fact that through its behavior an agent can actively select and modify its environment.

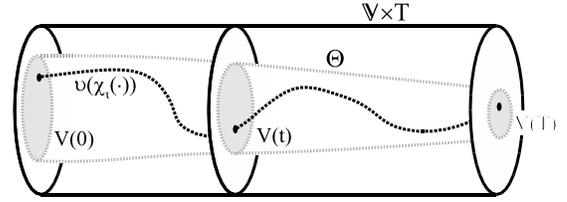


Figure 3. The viability trajectory $v(\chi_\iota(\cdot))$ of an individual whose viability is defined as a function of its state $\chi_\iota(t)$. An individual is eliminated if and when its viability trajectory leaves the viability tube Θ .

C. Discussion

As anticipated in the introduction, the viability elimination model entails a number of advantages relatively to the usual selection-based approaches. A first obvious observation is the fact that inequality constraints that define an infeasible region for the search space [33] can be directly interpreted and enforced as viability conditions. Other interesting properties ensue from the fact that, contrary to the case of selection – where the space \mathbb{S} must be ordered and, in fact, usually corresponds to \mathbb{R} – with this model we deal with a function v that takes its values in a generic set \mathbb{V} . This means, among other things, that v can easily represent multiple independent viability conditions, as illustrated by the amplifier design example mentioned above. This has significant consequences on how evolutionary experiments can be set up and run.

First, an individual can be subjected to multiple viability criteria that have a different nature and are therefore not directly comparable. With the traditional selection approach, instead, we would be forced to aggregate these incommensurable criteria in a single real-valued function. Furthermore, we can easily add a new viability criterion or drop an existing one independently from the other criteria, whereas with an aggregated selection function the modification of a selective criterion requires the redefinition of the whole function. This fact allows in particular the setup of EC experiments with a steadily increasing number and severity of viability criteria, which can be useful in the perspective of incremental evolution [85]. Another important consequence of this property is that we can prune more efficiently the set of trials generated by the evolutionary process. As Campbell has put it, referring to the traditional example of the group of chimpanzees typing at random in the experimenter’s hope of producing by chance all of the books in the British Museum [22, p. 106] (see also [106, p. 127])

In biological evolution and in thought, the number of variation explored is greatly reduced by having *selective criteria imposed at*

every step. ... It is this strategy of cumulating selected outcomes from blind variation, and then exploring further blind variations only for this highly selected stem, that, as R.A. Fisher has pointed out ..., makes the improbable inevitable in organic evolution. ... In constructing our “universal library” we stop work on any volume as soon as it is clear that it is gibberish.

Obviously, this step by step selective process can be easily implemented in terms of elimination but hardly if at all in terms of selection, since preliminary outcomes at different stages are unlikely to be comparable. Finally, thanks to the independence of the various viability conditions and to the viability being a function of the state, different individuals can be subjected to different elimination events, thus greatly increasing the variety of evolutionary events available for the individuals of the population.

The viability elimination approach entails also another, less apparent consequence. We allude to the fact that the assignment of a viability set implies the existence of some knowledge about the results that can be expected from the evolutionary process. In EC it is instead usually assumed that no information is available about the range of results achievable. However, this stance – due probably to the prevalence of the selection-centered perspective combined with an earlier focus of EC on function optimization – appears questionable. For example, in the amplifier design example introduced above a lot of information is usually available on the combined ranges of performances (cost, amplification, noise, efficiency, ...) that can be attained. It would be clearly unreasonable to ignore that prior information in modeling the problem and setting up the experiment¹⁴.

In sum, we see that the reduced requirements in terms of mathematical structure of the codomain¹⁵ of the viability function relatively to those of the selection function have a beneficial impact on the nature and variety of evolutionary events that can be modeled. Moreover, as observed by Mayr [78, pp.118], the action of what is represented in our model by viability elimination in place of selection can have a favourable effect in terms of variance of the resulting population:

A process of selection would have a concrete objective, the determination of the “best” or

¹⁴Note, however, that even if no prior information is assumed, an initial population can be usually randomly generated and used to assign a (possibly very permissive) initial viability set to be then shrunk during the experiment.

¹⁵The codomain of a function $f : D \rightarrow C$ is the set C where the function takes its values [66].

“fittest” phenotype. Only relatively few individuals in a given generation would qualify and survive the selection procedure. This small sample would be able to preserve only a small amount of the whole variance of the parental population. Such survival selection would be highly restrained. By contrast, a mere elimination of the less fit might permit the survival of a rather large number of individuals because they have no obvious deficiencies in fitness.

In fact, according to Mayr, the appreciation of the presence of abundant variance within natural populations and of the related concept of *population thinking* [78] are essential for a correct understanding of the workings of the evolutionary process.

We are now in the position to define the structure of an evolutionary algorithm (EA) that implements both the selection and the elimination processes. We will call such algorithm a *Selection-Elimination EA*¹⁶ (SE-EA). The traditional EA algorithm – that we will call *Selection-Based EA* (SB-EA) – will be a particular case of it, obtained by skipping the elimination part of the SE-EA, whereas by skipping the selection part (e.g., using a constant function as selection function σ) we will obtain an *Elimination-Based EA* (EB-EA).

VIII. SELECTION-ELIMINATION EA

An SE-EA has many elements in common with an SB-EA. Therefore its presentation and symbolism can be loosely based on those used to define the General Evolutionary Algorithm [8]. The discrete time steps (or generations) are indicated by integer numbers $t \in T$, starting with $t = 0$, where T is the time domain of the experiment. The population at time t is denoted by $P(t)$ and is composed by elements ι of the space of individuals (or phenotypes) \mathbb{I} . As anticipated above, each phenotype ι derives from a genotype γ in the space of genotypes Γ and is endowed with a state $\chi_\iota(t)$ that takes its values in the state space \mathbb{X} . Note that $\chi_\iota(t)$ must not be considered a pre-assigned function of time but rather a quantity that unfolds in time during the experiment according to the structure of the phenotype and of the environment. However, for many simple problems $\chi_\iota(t)$ ends up corresponding simply to the phenotype ι .

With some abuse of notation, when no confusion is to be feared we will use the symbol P to denote both the set of genotypes and that of phenotypes of the population, and will suppose that access to the phenotype implies the

¹⁶Alternative names could be *Homeostatic EA*, and *Viability EA*, although, as explained below, the viability elimination is not the only kind of elimination present in the algorithm.

access to the corresponding genotype and to the state of the individual. The number of individuals in $P(t)$ is denoted by $\mu(t)$. We will not adopt here the λ, μ symbolism typical of Evolutionary Strategies [8], [9] to distinguish the number of offspring and parent individuals. We will use instead a sequence $P'(t), P''(t) \dots$ to denote the successive temporary populations produced by the algorithm and the symbol $\mu(P(t))$ to denote their cardinality. As described above, we define a *viability function* $v : \mathbb{X} \rightarrow \mathbb{V}$ and a *selection function* $\sigma : \mathbb{X} \rightarrow \mathbb{S}$. To simplify the notation, the sets of viability and selection values $\{v(\chi_i(t)), \iota \in P(t)\}$ and $\{\sigma(\chi_i(t)), \iota \in P(t)\}$ will be denoted by $\mathcal{V}(P(t))$ and $\mathcal{S}(P(t))$, respectively.

We define then a *selection and reproduction operator* $r(P(t), \mathcal{S}(P(t)))$ that takes as arguments a population $P(t)$ and the corresponding set of selection values $\mathcal{S}(P(t))$ and produces a new set of individuals¹⁷. The action of the operator r corresponds to that of the selection and reproduction operators used in current EA [8], [18], [32]. To this familiar operator the SE-EA adds two *elimination operators*. The first is the *viability elimination operator* $e^v(P(t), \mathcal{V}(P(t)), V(t))$. It takes as arguments a population $P(t)$, the corresponding set of viability values $\mathcal{V}(P(t))$, and the current viability set $V(t)$, and returns a population constituted by the viable individuals of $P(t)$. It corresponds obviously to the action of the viability elimination events discussed above. The other elimination operator is the *contingency elimination operator* $e^c(P(t))$. It takes as argument a population and discards from it some “unlucky” individuals, to prevent population explosion when the joint action of selection and of viability elimination does not lead to a sufficient reduction of the population. The role and interpretation of e^c will be detailed later on. Note that all these operators are usually characterized by a set of parameters (for example, the mutation and crossover probability of a reproduction operator r [47], or the maximum number of individuals that the contingency elimination operator e^c must return), although we do not explicitly mention them in our symbolism.

Since, as noted above, the viability set can change at each time step (Figure2) we define a *viability set update operator* $u(V(t))$ that takes as argument the current viability set and returns the viability set that will be used

¹⁷We could imagine reproduction taking place at time instants belonging to a subset of the time domain of the experiment. This would be probably more plausible in the cases where the individuals are endowed with a state unfolding in time with a consequent ongoing viability assessment. To keep simple the description of algorithm we will not consider further this otherwise reasonable distinction.

in the next time step¹⁸. Finally, $\omega(P(t), \dots, P(0); t)$ denotes a *termination criterion*, which can in general depend on the whole history of the population and on the number of time steps executed. With these provisions, Algorithm 1 describes the basic structure of the SE-EA.

Algorithm 1 : The Selection-Elimination EA

```

t := 0;
initialize population P'(0)
assign viability set V(0);
eliminate for viability:
    P(0) := e^v(P'(0), V(P'(0)), V(0));
while (not (omega(P(t), ..., P(0); t))) do
    select and reproduce:
        P'(t) := r(P(t), S(P(t)));
    eliminate for viability:
        P''(t) := e^v(P'(t) union P(t), V(P'(t) union P(t)), V(t));
    eliminate for contingency:
        P(t+1) := e^c(P''(t));
    update viability set:
        V(t+1) = u(V(t));
    t := t + 1;
od

```

Note that according to Algorithm 1 the viability elimination operator e^v acts on the newly generated individuals and on those of the previous generation. There is however no loss of generality in always writing the old population as an argument of e^v since we can still define e^v as eliminating all the old individuals. With this assumption the formalism will permit both the implementation of steady-state and of generational EA [8], [9]. Moreover, for the sake of generality, the viability of the old population is assumed as being re-evaluated at each time step, even if in some cases the viability values and the viability set may not have changed from the previous time step.

This view of the evolutionary process, with its separate modeling of selection and elimination processes, seems to be close in spirit to that entertained by Darwin. As mentioned earlier, besides natural selection Darwin considered at length *sexual selection*. Introducing the subject in *The Descent of Man* Darwin gives the following definition [31, p. 256]

¹⁸As will become clear later on in discussing the issue of population explosion, it is often useful to pass as additional parameters to the viability set update operator u the current population and the tentatively generated new ones.

We are ... here concerned only with that kind of selection, which I have called sexual selection. This depends on the advantage which certain individuals have over other individuals of the same sex and species, in exclusive relation to reproduction

In their introduction to the reprint of *The Descent of Man* [31, p. xxviii], Bonner and May wrote:

Darwin did ... see “sexual selection” as a mechanism somewhat distinct from “natural selection” (which he often tended to treat as pertaining to survival) ...

If Bonner and May’s interpretation of Darwin’s attitude is correct (see also Michod [81, p. 13]), we can attribute to Darwin a propensity to distinguish quite clearly natural selection from sexual selection, and - as suggested by Mayr in the opening quotation - to see the former as an elimination process. It is sometimes argued [96] that present day EA are focused on natural selection and do not provide an implementation of sexual selection. If we adhere to Darwin’s definition, however, we must recognize that what is implemented in present day EA, being centered on selection aimed at reproduction, corresponds actually to sexual selection and it is instead Darwin’s concept of natural selection that is poorly represented in current EC practice.

IX. EXTINCTION

One of the consequences of the presence of an elimination step within an EA is the possibility that at a given time instant no viable individuals exist within a population, which therefore becomes extinct. At first sight, the idea of an EC experiment contemplating the extinction of the population appears ludicrous. However, consider the implications of using instead an algorithm whose structure excludes *a priori* a similar event, for example one where - as in most current EC experiments - the sole selection is at work. The tacit stance in this case is the willingness to accept in any case the outcome of the execution of the algorithm, no matter how “good” the resulting population is in absolute terms; in other words, the experiment becomes a self-confirming process.

In some cases - especially in the realm of engineering parameter optimization - some preliminary conditions are given that assure the acceptability of the results. In many other cases, however, it is often the case that the final results are judged unsatisfying and brushed off. In this sense most EC experiments to date can be said to have ended up with the tacit extinction of the final population. From this point of view, the systematic occurrence of extinction in an EC experiment

implementing elimination should not be considered a fault but merely a sign that something is wrong either in our expectations or in the process that we have concocted to fulfill them. Moreover, an EC experiment where by definition the population cannot go extinct must be considered analogous to a scientific theory that cannot be falsified [87] and to that experiment we can apply, *mutatis mutandis*, Rushton’s remark (as quoted by Platt [86]): “A theory which cannot be mortally endangered cannot be alive”.

More concretely, in an evolutionary perspective the extinction of a population can constitute an instance of elimination for an higher level evolutionary entity. Conversely, in the absence of extinction the very existence evolutionary entities of higher level is jeopardized. In other words, the possibility of extinction is a central component for any model of the evolutionary process considered as a potentially multi-level process. Note, however, that this is true only if extinction derives from the elimination of lower level evolutionary entities. For example, the cases of EA cited in the Introduction that implement the concept of extinction define this concept directly at the species level and do not contemplate the existence of other evolutionary levels. Therefore, these approaches cannot result in a multi-level evolutionary process.

X. POPULATION EXPLOSION

The size of unchecked populations in nature is potentially subject to an exponential rate of growth. This means that, besides extinction, EC modeling must deal with another critical issue: population explosion. Current EA usually solve the problem by defining the selection and reproduction operators in ways that assure the stability or even the constancy of population size [8], [48]. This prevents simultaneously both the extinction and the unbounded growth of the population. We can of course adopt the same straightforward approach to settle the case of population explosion in SE-EA. The presence in SE-EA of the elimination operators, however, opens the way to fresh approaches to the problem. To understand what algorithmic models could be fruitfully implemented it is useful to consider briefly the real-world counterpart of this issue and its biological and ecological models.

How the delicate balance is obtained that keeps most of the time a natural population away from explosive growth is still matter of debate between population ecologists [37], [73], [104]. Simplifying at the extreme the state of the debate, we can single out two major attitudes towards the problem. The first, originally championed by Nicholson [84], sees the population as a regulated system controlled by density-dependent limiting factors. This

approach is questioned by some scholars (for example, see the discussion and references in [37]) who observe that one cannot assume the stability of the populations as a datum, since, for example, field observations show that populations can experience large density fluctuations and multiple episodes of local extinction followed by repopulation, while no density-dependent regulating factor can be explicitly recognized. According to some of this critics the limitation of population size, apart from isolated episodes of resource exhaustion, seems to be mostly due to density independent factors, usually represented by contingent environmental factors [34], [37], [95]. Without taking side in this debate, we will briefly illustrate how both these mechanisms can be easily modeled within the SE-EA framework using the viability and contingency elimination operators.

Let us start by considering the case of density-dependent limiting factors. If the EC experiment implements the interaction of each individual with the environment and with other individuals, a more crowded environment can result automatically in a greater difficulty for individuals in maintaining viability. Thus, the existence of density-dependent factors can ensue even with a fixed viability set. With simpler implementations that do not model these interactions we can still represent the presence of density dependent factors by linking the “size” of the viability set V to the cardinality of the population (or to the local density of the population, if the experiment implements a spatially distributed population and local viability sets are used). This manipulation of the viability set can be performed by the viability set update operator u . To this effect, u would take as additional parameter the current population and, possibly, the tentatively generated new ones. Density-independent contingent limiting factors can be instead modeled using the expressly defined contingent elimination operator e^c . For example, e^c could operate by eliminating randomly individuals of the tentatively newly generated population if its size exceeds a given limit.

XI. ERROR-DRIVEN PROCESSES

The *modus operandi* of natural selection has been often compared to that of a control system [13], [38]. In fact, the analogy with a regulator was already evoked by Wallace in the essay he sent to Darwin reporting his discovery of natural selection [103]¹⁹:

The action of this principle is exactly like that of the steam engine [governor], which checks and corrects any irregularities almost before they become evident; and in like manner no

unbalanced deficiency in the animal kingdom can ever reach any conspicuous magnitude, because it would make itself felt at the very first step, by rendering existence difficult and extinction almost sure to follow

An engineering control system such as that described by Wallace, however, is designed with a definite purpose in mind: that of obtaining a prescribed behavior from a system. No such a purpose can be recognized in the workings of natural selection. The elimination perspective permits to show that Wallace’s intuition was not unfounded by revealing that the two processes share an important structural characteristic: that of being both *error-driven*.

A traditional control system [17], [64], [67] is composed by two subsystems: the controlled system, or “plant”, and a controller that generates the signals that are used to control the plant, that is, to assure that its behavior is the desired one. We speak of closed loop control if the controller generates the control signals by comparing the output signals of the plant with some given reference signals representing its desired behavior. The comparison gives rise to an error signal, and the controller is structured so as to transform the error signal in controlling signals that act to reduce the measured error. In other words, the error produces an effect that promotes its own dissolution [16], [64]. In this sense, a closed loop control system is error-driven [67].

For an engineered control system, it is the responsibility of the designer to structure the system so as to render it error-driven, that is, to assure that any deviation from the desired behavior automatically generates a correcting action that tends to force the system to resume the desired behavior [16]. In the case of an evolutionary entity, we don’t have a predefined desired behavior nor a designer that takes care of “wiring” the system so as to make it error-driven. However, the unique circumstance holds that the absence of viability of an individual leads automatically to the elimination of the individual itself. Hence, if we interpret the absence of viability as the error, we have intrinsically in place a structure where – as required for a system to be error-driven – the error promotes its own dissolution. In this sense we can interpret the process of natural selection as giving birth to a self-structuring error-driven system. Note that we have here another reversal of the original logic. In the case of a closed loop control system we start with an error and proceed to enforce its elimination thus creating an error-driven system, whereas in the case of natural selection we start with an elimination and interpret it as that of an error thus interpreting the process as error-driven. It is thanks to this concurrence of the error with

¹⁹The term “governor” in the quote is due to Bennet [17].

the process of its elimination that the process can self-activate itself in nature. On the contrary, usually in EC experiments the error refers to a “viability” condition that does not entail the automatic elimination of individuals, which must be instead enforced by the experimenter. If this separating the error from its elimination can leave intact the evolutionary potential of the process is a far from trivial matter. We will discuss it in the next section.

The fact of being error-driven has interesting consequences on the behavior of a system. Particularly noteworthy is the fact that, since the effect of the perturbations is represented by the error signal, a correcting action is produced independently of the nature of the disturbing factor (provided the amplitude and the rate of generation of the disturbance are not excessive). Therefore, such a system can face perturbations whose characteristics were not known in detail (in fact, whose existence was possibly not even conceived) by the designer. The other side of this property is the fact that no controlling action is exerted as soon as the error vanishes or stays within prescribed bounds²⁰. From our evolutionary perspective this means that, once the population is viable and produces a viable progeny, no error is present and none will be as long as the viability conditions and the capacity of individuals to maintain viability remain the same. Hence, no corrective action will be enforced. In other words in the absence of perturbations such as a changing environment or the emergence of novel strategies in the ecological environment, natural selection interpreted as an error-driven process appears primarily as a conservative and stabilizing mechanism. However, – going back to the announced intention of thinking in terms of how the evolutionary process works – in the end it is environmental interaction that causes differential replication [20], [57]. Therefore we must not be surprised to discover a mechanism that tends to produce a stable response in the presence of a stable environment²¹. Moreover, as observed by Bateson [12, p. 429], “[natural selection] may act at higher levels to keep constant that complex variable which we call “survival””. Finally, let us remark that the evolutionary process results from the iteration of a two-step cycle [70], [78] where first a “cloud” of genotypes is produced according to the rules

²⁰This property holds for a number of recently proposed approaches to the understanding and modeling of biologically related phenomena. For example, Matzinger’s Danger Model for the operation of human immune system [75] and Bak and Chialvo’s Extremal Dynamics approach to learning in neural networks [11], [26]. The concept evokes also to the Inertia Principle advocated by Aubin in the context of Viability Theory [4].

²¹We must also remember that, as pointed out previously, the environment can change as a consequence of the presence of the organisms.

of genetic recombination and mutation (and, possibly, of those of sexual selection) applied to the existing population, and then the corresponding cloud of phenotypes is “shaped” by the effect of natural selection. The natural selection “shaping force” acts on phenotypes, but its effects are pulled back on the space of genotypes. From this point of view, the viability elimination shaping effect seen in the error-driven perspective can result (if the state dynamics is absent or elementary) in a very simple, dichotomic mechanism operating in the phenotype space. Yet, as will be argued in the next section, it is often the only one that within an EC context can be justified on the basis of the information actually available about the problem at hand.

XII. DISCUSSION

We have so far advocated the explicit presence of an elimination process in models of the evolutionary process. Correspondingly, we have defined and detailed the structure of the SE-EA. It is now time to ask what advantages we can expect from the adoption of this approach in EC. We will consider first the case of engineering applications in general, and then the more particular cases of Artificial Life (ALife) [1], [82] and Evolutionary Robotics (ER) [42], [53], [54], [85].

A. Engineering, Optimality and Prior Information

EA are usually considered and presented as optimization tools. It is easily seen that we cannot consider an EB-EA as a search or optimizing algorithm in the conventional sense [80, p. 39], since at each time step the result of the execution of the EB-EA is the extant population (if any), whose paramount property is that of being viable according to the current viability criterion. On the one hand, however, the relevance of the interpretation of evolution as optimization has been questioned many times (see for example, [39] and [52]). Moreover, although engineering is often assumed to be interested mainly in optimal solutions, it is more appropriate to say that engineering is the art of finding “good enough”, i.e., viable solutions to problems.

In a typical design scenario a range of tentative solutions are first (ideally or concretely) considered and the “not good enough” ones, i.e., the non viable ones according to the given specifications, are discarded. It is only when one or several eligible system structures and parameters values have been found, that an optimization process is possibly enforced on them. This optimization process can be represented in terms of selection provided that the problem can be formulated in terms of a single scalar function to be maximized or minimized. Most

real world problems, however, are multi-objective ones and it usually difficult to aggregate them into a single scalar function [33], [40]. If no obvious criterion for the comparison and combination of the goals can be derived from the statement of the problem, the imposition of the aggregation of objectives can change substantially the nature of the problem. This follows from the fact that this forceful aggregation corresponds actually to assuming as given information that is actually not available. For example, considering again our familiar amplifier design problem, it is usually difficult to quantify the relative importance of, say, an increase in gain and an increase in power supply noise rejection. There is usually an obvious ordering of pairs of tentative solutions based on the fact that one of them is better than the other relatively to some criteria and not worse relatively to the remaining ones (in the language of multi-criteria optimization it is said that one solution dominates the other [33], [40]). This, however, produces in general only a partial ordering of the tentative solutions whereas if we specify a fitness function we are assuming that this ordering is actually available for all pairs of solutions. In the amplifier design example this corresponds to the ability to order in a linear sequence all possible amplifiers; a feat that would impress the most knowledgeable of engineers. It is true that in the process of natural evolution individuals are ranked according to their reproductive success (or fitness). We must not forget, however, that in nature the sorting constituted by the relative reproductive success of individuals emerges only *a posteriori*, as a consequence of the interaction of the individuals with their environment. The potential complexity of this interaction and the ensuing difficulties in modeling it in general, should make us wary of the assumption that the information to perform this sorting is available from the start in the form of a fitness function.

In addition to this risk of assuming the availability of a knowledge that is not actually there, there is the complementary risk of disregarding knowledge about the problem and its possible solutions that is available at the beginning of the process. We have previously mentioned the fact that in most real-world problems we have some prior information on the results that one can reasonably expect to achieve. We have observed that adopting the viability elimination approach this information can be utilized to define a target viability set. On the contrary, it is easily seen that this information can be hardly if at all utilized within an EA based on the fitness function approach. Summing up, we can say that adopting the conventional fitness function approach we can be forced by the structure of the algorithm to disregard some information that we actually have and to assume the

possession of information that we do not actually have. With the addition of the viability elimination event to the EA structure we acquire the possibility to avoid both kinds of problem.

Considering the issue from a more abstract point of view, we can remember that – as a consequence of the No Free Lunch (NFL) theorems [29], [107] and contrary to earlier expectations – no algorithm, be it evolutionary or not, and be it based on selection or on elimination, can aspire to be a “universal” search and optimization procedure. In relation to this remark, it is interesting to consider English’s analysis of the NFL theorems from the point of view of information theory [41]. According to this analysis, the NFL results stem from the fact that for a generic optimization problem, the outcome of a point evaluation is not informative with respect of other possible evaluations and, therefore, in particular, relatively to the localization of the optimal points, which constitutes the aim of the process. Contrary to this, both the realization that an individual is (at a given instant) viable and the discovery that it is not, constitutes a piece of relevant information for a process whose goal is the production and preservation of a viable population. These observations can be reformulated using the concepts of System Theory [92] interpreting the genetic information as representing the state of the evolutionary system [55]. In this perspective, the population²² is the memory of the challenges and environmental interactions faced by the past and present populations at all the organization levels. This memory is updated by the evolutionary process. From this point of view, the main advantage of an SE-EA over a conventional SB-EA lies in its allowing the collection information that reflects more closely the information actually available thanks to its capability of modeling (and let individuals experience) a greater variety of evolutionary events.

B. ER, ALife, and Open-Ended Evolution

The previous observations relative to the engineering design process apply in part also to the field of Artificial Life (ALife) [1], [82] and to that of Evolutionary Robotics (ER) [42], [53], [54], [85], whose problems can be considered as being halfway between those of engineering and those of ALife. For example, the separate applicability of different viability criteria related to performances such as obstacle avoidance, navigation, task completion, and so on, has the potential to change substantially the way ER experiments are

²²Or, better, the structures that function as “bookkeeping entities”, or “replicators” in the population [50], [57].

set up and run. In the fields of ER and ALife, however, the problems constituted by the creativity and the open-endedness of the process assume a new relevance. These are particular cases of the the general problem of the *growth of knowledge* considered at length by evolutionary epistemologists (see for example [22], [23] and references therein). Popper [88] identified the source of this possible growth in a process of *conjecture and refutation*, where we can easily to recognize in the refutation a process of elimination for viability. From this perspective, besides the problem constituted by the generation of the conjectures - that we will provisionally imagine as solved with some random process endowed with sufficient generative potential - the problem remains of how to choose the refutation criterion.

It has been argued that a major obstacle to the creativity of the evolutionary process in EC is constituted by the constraint imposed by the explicit definition of the fitness function (see, for example, the discussion on the fitness space in [85, p. 65] and that on the endogenous vs. exogenous fitness function in [82]). The elimination perspective represents undoubtedly a progress in this sense, since a negative injunction is less prescriptive than a positive one (similarly to the proverbial country where all that is not explicitly forbidden is allowed, which enjoys more freedom than that where everything which is not explicitly allowed is prohibited). Yet, it is not clear if this progress can be considered sufficient. A really creative solution, in fact, has the potential to change not only how viability is achieved but also what must be considered a viability condition²³, whereas in the SE-EA the viability conditions can be fully endogenously generated only by populations of completely autonomous individuals.

Some authors have indeed conjectured that in the context of EC a truly creative evolutionary process can be obtained only in the presence of autonomous self-reproducing entities [79], [89]. It is clear that this provision solves at once the problem of providing a compelling and intrinsic causal foundation to the differential reproductive success that animates the evolutionary process. However, until this provision can be actually met, our goal remains that of ascertaining if an effective evolutionary process, be it creative or "merely" adaptive, can be based on a partially extrinsic definition of that causal foundation. In this respect, the agent-based

example constituted by the growth of cultural knowledge and the role attributed in it to the process of confutation seems to speak for the centrality of elimination in any foreseeable realization of that so far elusive goal. It can well be the case, however, that a truly open-ended evolutionary process mimicking closely what happens in nature can be realized only at the price of relinquishing the control over the process and its outcomes. As pointed out by Lewontin [70, p. 10]

If genetic algorithms are to be used as a way of solving engineering problems by analogy to the supposed success of natural selection in producing adaptation, then they must be constructed for the limited domain on which the analogy holds. The alternative is to evolve machines and later to find uses for them to which they are preadapted, a process not unknown in human invention.

XIII. EXAMPLE: MULTI-OBJECTIVE PROBLEMS

We proceed now to illustrate the SE-EA approach with an example. We consider the case of an elementary multi-objective problem [33], [40] that we assume as representative of the class of engineering problems described in the previous section, namely, those concerned with the search of a set of viable solutions relatively to multiple criteria. As discussed above, we will adopt at first the hypothesis that a target viability set is known. Note that this assumption does not trivialize the problem since once a population complying with the target viability requirements has evolved we still have the problem of how to further tighten the viability requirements to improve the system performance. Moreover, there is in general the possibility that new ways to realize the target performance are discovered by the algorithm.

To ease the visualization of the algorithm's workings we consider a design problem where only two real-valued system parameters x and y must be assigned. We assume that some pre-defined system constraint limits the set of admissible parameters is the two-dimensional interval $\mathbb{I} = [x_{min}, x_{max}] \times [y_{min}, y_{max}] \subset \mathbb{R}^2$. A potential solution is therefore constituted by a pair $\iota = (x, y) \in \mathbb{I}$. To be acceptable, a solution must result in an acceptable performance of the system. We represent this fact assuming as given p real-valued functions $v_c(x, y) : \mathbb{I} \rightarrow \mathbb{R}$, $c = 1, \dots, p$. Each function gives a measure of a specific characteristic of the system for the pair of parameters passed as argument. We assume that a system with parameters (x, y) is acceptable only if the value $v_c(x, y)$ of each function belongs to a specific set $V_c \subset \mathbb{R}$, $c = 1, \dots, p$. In the actual implementation we

²³In 1888 the mathematician David Hilbert gave a solution to the central problem of invariant theory. However, his solution was an existence proof while only computational solutions had been previously considered acceptable by the leading experts in the field. Thus, Hilbert's revolution in invariant theory implied the creative act of redefinition of the concept of "viable" solution.

can obviously adopt a shortcut evaluation of viability, that is, for each individual we can stop the evaluation of the ensemble of functions v_c as soon as we find a value that lies outside the corresponding set V_c .

The rendition of these specifications in the vocabulary of the SE-EA is straightforward. Each pair of parameters is considered as the phenotype of an individual $\iota = (x, y) \in \mathbb{I} \subset \mathbb{R}^2$. The performance functions $v_c(\iota) = v_c(x, y) : \mathbb{I} \rightarrow \mathbb{R}$, $c = 1, \dots, p$ are assumed as the p components of a viability function $v(\iota) = (v_1(\iota), \dots, v_p(\iota)) : \mathbb{I} \rightarrow \mathbb{R}^p$. We see that here the viability space \mathbb{V} corresponds to \mathbb{R}^p . An individual is viable if each component of the viability function belongs to the corresponding set V_c . This condition corresponds to the requirement that the viability value $v(\iota) = v(x, y)$ belongs to the viability set $V = V_1 \times \dots \times V_p \subset \mathbb{R}^p$ obtained as the Cartesian product of the V_c . Usually the sets V_c will change (independently) during the execution of the algorithm. Hence, we will have a time-varying viability set $V(t) = V_1(t) \times \dots \times V_p(t)$. The collection of all these sets constitutes the viability tube. For the phenotypes ι we adopt a genetic encoding where the genotype γ is the concatenation of the binary representation (with linear scaling) of length b bits of the two parameters x and y of ι .

Coming now to the SE-EA operators, we assume that the reproduction operator r picks pairs of distinct individuals of the existing population and performs on their genotype two point crossover with probability p_c , and uniform mutation with probability p_m [47]. This means that the algorithm ignores the value of the selection function $\sigma(\iota)$ (or, alternatively, assumes a constant selection function) and implements actually the sole elimination. The reproduction operator is defined so as to assure that for each individual of the original population n_{os} new individuals are produced, where n_{os} is a positive integer. To facilitate the implementation of the case $n_{os} = 1$, and contrary to established GA practice, of the two new individuals potentially produced by a crossover operation, the reproduction operator retains only one. In practice, a random permutation of the individuals of the parent population is generated, and pairs of adjacent individuals in the permutation (with wraparound at the end of the list) are used to generate a single offspring by crossover. The operation is repeated n_{os} times. As dictated by its definition, the elimination operator e^v returns a population that contains only the viable individuals of the argument population, constituted by the union of the old population and of the population newly generated by r . The contingency elimination operator e^c trims to μ_{max} individuals the argument population performing a random elimination on it. The termination

criterion is constituted by the time step counter reaching a predefined value t_{max} or by the population going extinct.

We can make this model a bit more interesting, and illustrate an elementary case of state-dependent viability, by implementing a structure of age-classes in the population. To each individual we associate an age tag $\alpha_\iota(t)$ that is set to zero when the individual is generated and incremented (possibly stochastically) at each generation. We then associate to the individuals a state variable $\chi_\iota(t) = (\iota, \alpha_\iota(t))$ that is composed by the (fixed) phenotype ι and by the age tag $\alpha_\iota(t)$. Correspondingly, we add a viability condition that eliminates individuals reaching a maximum age a_{max} by defining an additional set $V_\alpha = [0, a_{max}]$ and redefining the viability set as $V(t) = V_1(t) \times \dots \times V_p(t) \times V_\alpha$ and the viability function as $v(\chi_\iota(t)) = (v_1(\iota), \dots, v_p(\iota), \alpha_\iota(t))$.

Finally, we must define the viability set update operator u and the criterion for the assignment of the initial population and of the initial viability set. The update of the viability set transforms progressively the initial viability set into the target viability set. We can consider several policies for this transition. Some of them do not depend on the size of the population. For example, if the sets $V_i(t)$ are real intervals $[l_i(t), L_i(t)] \subset \mathbb{R}$ we can define a linear transition $l_i(t) = l_i(0) + (\bar{l}_i - l_i(0)) t/t_{max}$ from $l_i(0)$ to $\bar{l}_i = l_i(t_{max})$ or, more generally a power law transition $l_i(t) = l_i(0) + (\bar{l}_i - l_i(0)) (t/t_{max})^\beta$, $\beta > 0$, with an analog rule for the upper limit. We will adopt this approach in the first example below. To these deterministic policies a small noise term $n(t)$ can be possibly added. Obviously, one can also consider policies that depend on the population and on its dynamics. These policies can be applied in particular when no target viability set is given or to probe the possibilities of further evolution of the population when the target is met. Typically, in this case we will shrink a little the viability set, observe the effect on the population and readjust the viability set accordingly. Alternatively, one can estimate the sensitivity of the population size to the variation of each viability limit and use that information to specify an updating policy for the viability set. This is the approach that we will adopt in the second example presented below.

When we decide to stop we are left with the individuals of the populations that did not go extinct. From this set we can extract a subset of solutions according to some additional criterion. For example, if the problem can be interpreted as a multi-objective optimization one, we could extract the non dominated solutions [33], [40]. Note, however, that even in this case the algorithm does not aim at locating or representing faithfully the

Pareto front [33], [40] which, in general, will lie partially outside the viability set. We could also imagine a multi-level algorithm where several populations are evolved in parallel using different sets of algorithm parameters for each population and the additional rule that if a population goes extinct it is substituted by a copy of the most florid extant one, with slightly mutated algorithm parameters.

A. Example 1

We consider the case $p = 2$, i.e., the viability of the system that is being designed depends on two characteristics (say, the cost and the gain, in our amplifier design example). We encode with $b=10$ bits each parameter in the bi-dimensional interval $\mathbb{I} = [-2, 2] \times [-2, 2] \subset \mathbb{R}^2$. The viability function is $v(\chi_i(t)) = (v_1(t), v_2(t), \alpha_i(t)) = (|\sin(\pi(x+2y^2))|, \cos(9x - (y-1)^2), \alpha_i(t))$ where $\alpha_i(t)$ is the age of the individual, which is set to zero at creation and incremented by one at each generation. The viability set has the form $V(t) = V_1(t) \times V_2(t) \times V_\alpha$. The first characteristic $v_1(t)$ will be considered a cost, hence the corresponding set $V_1(t)$ will have the form $[0, L_1(t)]$. The second characteristic $v_2(t)$ will be considered improved with increased value, hence the corresponding set $V_2(t)$ will have the form $[\bar{l}_2(t), 1]$. The knowledge of the upper limit for $V_2(t)$, that in this case derives from the viability function being given in closed form, might appear artificial relatively to the case of real-world problems. However, it is almost always the case that in practice finite estimates of the limits of the viability ranges can be deduced from the information available about the problem. The age related viability condition is specified by the set $V_\alpha = [0, a_{max}]$. The target viability set is $\bar{V} = \bar{V}_1 \times \bar{V}_2 \times V_\alpha$, with $\bar{V}_1 = [0, \bar{L}_1] = [0, 0.25]$, $\bar{V}_2 = [\bar{l}_2, 1] = [0.5, 1]$, and $V_\alpha = [0, a_{max}] = [0, 3]$. Note that in our examples V_α does not change with time.

We generate a tentative initial population $P'(0)$ of μ_o individuals randomly in \mathbb{I} . Then, we assign the initial viability set so that all the individuals of the tentative initial population $P'(0)$ result viable by putting $L_1(0) = \max_{i \in P'(0)} \{v_1(t)\}$ and $\bar{l}_2(0) = \min_{i \in P'(0)} \{v_2(t)\}$. We chose to enforce a linear transition from $L_1(0)$ to \bar{L}_1 and from $\bar{l}_2(0)$ to \bar{l}_2 so that $L_1(t) = L_1(0) + (\bar{L}_1 - L_1(0))t/t_{max}$ and $\bar{l}_2(t) = \bar{l}_2(0) + (\bar{l}_2 - \bar{l}_2(0))t/t_{max}$. We use a small number $t_{max} = 7$ of steps to perform the transition from the initial to the target viability set, in order to be able to display the whole evolution. The termination function $\omega(t, P(t))$ is true when $t = t_{max}$ or the population size $\mu(P(t)) = 0$, and false otherwise. We start with a randomly generated population of $\mu_o = 100$ individuals and set the maximum number of individuals to $\mu_{max} = 1000$. The probability of crossover is

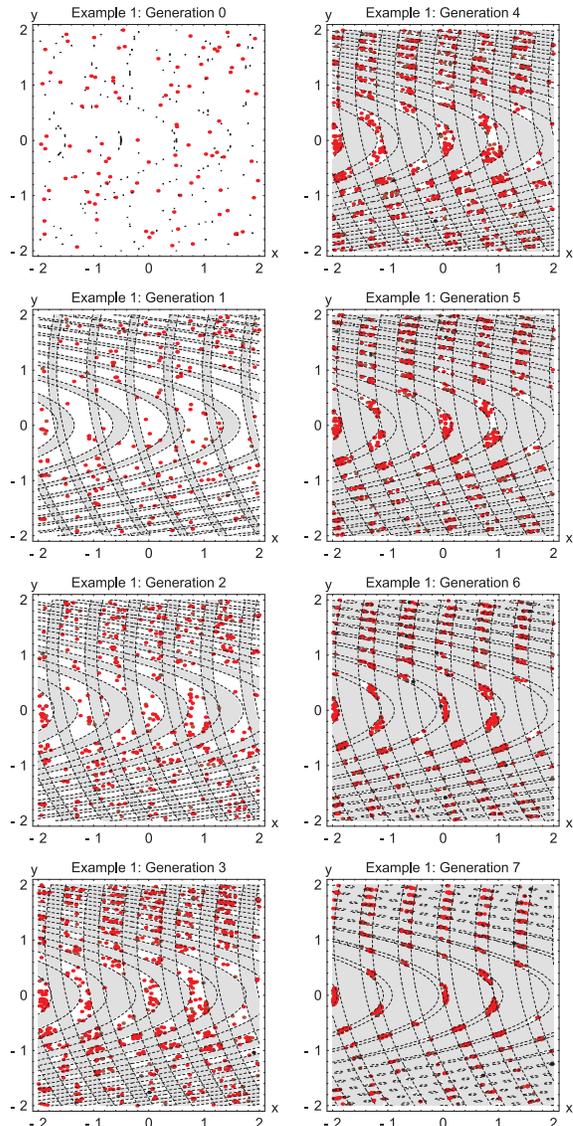


Figure 4. The evolution of the population while the viability tube goes shrinking, as observed in the phenotype space I . The shaded regions correspond to the non viable subset. Newly generated individuals are represented by light points, whereas old individuals are represented by darker points. Note that the initial viability set is defined so as to make viable all randomly generated individuals. The viability set shrinks progressively during the evolution, from the initial set to a predefined target set, where the extant individuals (if any) at the end of the evolution can be found.

$p_c = 1.0$ and that of mutation is $p_m = 0.03$. Finally, the number n_{os} of newly generated offspring for each individual in the parent population is $n_{os} = 3$, so that $P'(t) = r(P(t), \mathcal{S}(P(t)))$ has three times the size of $P(t)$. Figure 4 shows an example of evolution with these parameters, observed in the space I of phenotypes. Given the linearity of the updating of the viability set, the viability tube corresponding to this evolution is a truncated pyramid in the viability space-time $\mathbb{V} \times T =$

$\mathbb{R}^3 \times [0, t_{max}]$, with bases $V(0)$ and $V(t_{max}) = \bar{V}$.

B. Example 2

Now let us imagine that, given the final population obtained in Example 1, we want to further tighten the viability conditions. If our objective is just to modify a single limit, say L_2 , we can proceed by simply tentatively choosing a value for the variation that must be applied at each further step to that limit, in our case ΔL_2 , and proceed with the algorithm applying at each step the selected variation to the limit. Imagine, however, that we are instead interested in modifying the viability conditions relatively to several performances. This means that we must proceed to vary the viability limits for all the characteristics that we intend to influence. The problem in this case is that, in the absence of a target viability set, we do not know *a priori* how to select an updating policy for parameters that are referred to characteristics which, as pointed out previously, have in general an heterogeneous nature. To find a common measure for the effect of the variation of the viability limits, we can refer to the extension of the set I of viable individuals in \mathbb{I} (see Figure 1). For example, let us assume that in our case we are interested in improving the solution relatively to both L_1 and l_2 . The extension of I (which, with some abuse of notation, we will represent with the same symbol I instead of, for example, $|I|$) will depend on both parameters. Thus, we have $I = I(L_1, l_2)$ and, therefore, $\Delta I(\Delta L_1, \Delta l_2) \approx \partial I / \partial L_1 \Delta L_1 + \partial I / \partial l_2 \Delta l_2$. We now assume that variations of parameters that produce the same shrinking of I are “equally difficult” to achieve from the point of view of the EA, since I corresponds to the space available to the population. Hence, we will try to vary the limits L_1 and l_2 so that the relation $\partial I / \partial L_1 \Delta L_1 = \partial I / \partial l_2 \Delta l_2$ holds.

Now, in general, we don’t know the function $I(l_1, L_2)$, nor its partial derivatives. We can however estimate ratio of the two partial derivatives using the information constituted by the population that occupies I , which in our case is the final population P obtained in Example 1. If we assume that the population size $\mu(P)$ is sufficiently large, and that the population is distributed uniformly within I , then, by varying the parameters L_1 and l_2 and observing the change in size of the population we can obtain two values that correspond to an estimate of the two partial derivatives multiplied by a same constant k . In formulas, this corresponds to assuming that for small ΔL_1 and Δl_2 we have $\Delta \mu(P) / \Delta L_1 \approx k \partial I / \partial L_1$ and $\Delta \mu(P) / \Delta l_2 \approx k \partial I / \partial l_2$. In practice, we proceed to evaluate the relative variation $\rho|_{\Delta L_1} = \Delta \mu(P) / \mu(P)|_{\Delta L_1}$ and $\rho|_{\Delta l_2} = \Delta \mu(P) / \mu(P)|_{\Delta l_2}$ of the

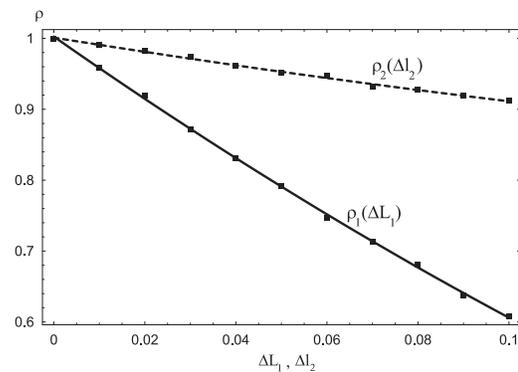


Figure 5. We can estimate the sensitivity of the problem difficulty to the variation of the viability limits observing the effects of a variation of the limits on the size of a viable population. Here, an ensemble of ten samples is obtained starting from the final population obtained in Example 1. By least-squares approximating this samples with a second order polynomial we obtain two continuous curves $\rho_1(\Delta L_1)$ (continuous line) and $\rho_2(\Delta l_2)$ (dashed line), whose slopes $e_1 \approx -4.50$ and $e_2 \approx -1.03$ evaluated for $\Delta L_1 = 0$ and $\Delta l_2 = 0$ are our estimate of the sensitivities relatively to changes of L_1 and l_2 .

population size corresponding to various values of ΔL_1 and ΔL_2 . Then we approximate with two continuous curves $\rho_1(\Delta L_1)$ and $\rho_2(\Delta l_2)$ the two sets of samples thus obtained (for example, with a least squares approximation using a low order polynomial as approximating function) and determine the slope of the curves at $\Delta L_1 = 0$ and $\Delta l_2 = 0$. These values – let us call them $e_1 = d\rho_1/d\Delta L_1|_{\Delta L_1=0}$ and $e_2 = d\rho_2/d\Delta l_2|_{\Delta l_2=0}$ – are our estimates for $k \partial I / \partial L_1$ and $k \partial I / \partial l_2$, respectively (Figure 5).

With the information constituted by e_1 and e_2 we can now proceed to execute some further steps of the algorithm. If the slope information represented by e_1 and e_2 can be considered valid for $\Delta \mu(P) / \mu(P) < \bar{\rho}$, then we can restart our algorithm setting $t = 0$, taking P as the initial population, defining $L_1(0) = \bar{L}_1$ and $l_2(0) = \bar{l}_2$ (that is, setting the initial value of the limits to the values used to obtain the population P), and executing t'_{max} time steps with the provision $L_1(t) = L_1(0) + \bar{\rho} / e_1 t / t'_{max}$ and $l_2(t) = l_2(0) - \bar{\rho} / e_2 t / t'_{max}$, that is, applying at each time step a variation $\Delta L_1 = \bar{\rho} / e_1$ and $\Delta l_2 = \bar{\rho} / e_2$ of the limits. Figure 6 shows the initial and final population for the execution of $t'_{max} = 7$ further steps with the values of $e_1 \approx -4.50$ and $e_2 \approx -1.03$ and the value of $\bar{\rho} = 0.1$ obtained applying our method to the case illustrated in Figure 5. If a further variation of the viability conditions is required, the process just described can be repeated using the final population thus obtained to determine new estimates e'_1 and e'_2 for the sensitivities and a new value $\bar{\rho}'$ for their range of validity.

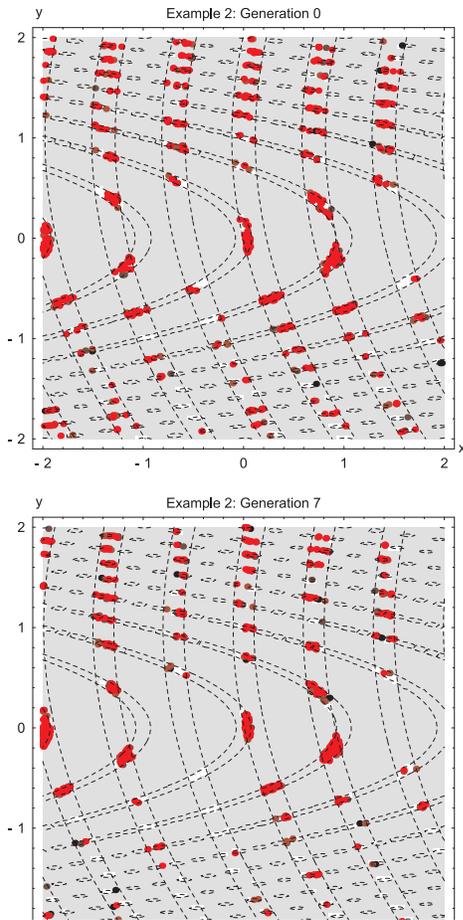


Figure 6. With the information constituted by the estimate of the sensitivity of the problem difficulty to the variation of the viability limits, we can proceed to the execution of further steps of the SE-EA. Here we show the initial and final populations for $t'_{max} = 7$ and the values of e_1 and e_2 corresponding to the approximation of Figure 5. What is here the initial population corresponds to the final population of Figure 4. Note that the different magnitude of e_1 and e_2 correspond to a more severe shrinking of the initially wider “vertical” stripes that determine the set $I(0)$, relatively to the narrower parabolic strips. The final population can be used to obtain a new estimate of the sensitivity relatively to the newly obtained viability set. As could be expected, the peculiarities of the history of the evolving population lead the final population to desert some regions of the final viable set $I(t_{max})$.

Note that this second example is not meant to maintain that this particular approach to the extraction of information about the problem must be the preferred one. The point we want to make is that once we have supplied to the algorithm (in terms of genetic coding, target viability set, and so on) all the prior information in our possession about the problem, new information can be obtained through the history of the evolving population subjected to the action of the algorithm. This includes the extrapolations that can be based on that history, such

as the elementary extrapolation based only on the final population, that we have utilized here.

XIV. CONCLUSIONS

Hull, Langman and Glenn [58, p. 53] defined selection as “repeated cycles of replication, variation, and environmental interaction so structured that environmental interaction causes replication to be different”. To model mathematically this process from the point of view of its causes it is necessary to model not only replication and variation, but also environmental interaction. The founders of population genetics choose not to pursue this approach and focused instead on the modeling of the outcome of the process: differential replication. Correspondingly, they introduced the concept of fitness to represent it. They had good reason to do so, since the alternative would have probably lead to evolutionary models of excessive complexity. But the decision to describe or define a process in terms of what is only an outcome of its workings comes at a price: the reduction of the empirical content of the model [56]. By subscribing to this decision with the adoption of the fitness function approach, EC has been forced to pay this price. We are convinced that this is not only unnecessary but unwise, given that EC experiments are supposed to implement the process that leads to differential reproduction, not model it mathematically.

The EC choice, therefore, should be instead the definition and implementation of suitable models of the interactions mentioned above. Here we have identified and modeled two kinds of events that could represent those interactions in EC experiments: selection events and elimination events. Selection events imply an ordering of the individuals and can be represented by the selection schemes based on the fitness function that are used in present-day EA. To avoid confusion we have simply suggested that what is now called “fitness function” be renamed “selection function”. Elimination events are instead based on a concept that does not belong to the current paraphernalia of EC: that of state-based viability. The addition of this kind of events greatly enlarges the complexity and variety of causes of differential reproductive success that can modeled in EA. The consequence is the possibility to make the information supplied to the evolutionary process reflect more closely the prior information actually available, in the dual sense of allowing the utilization of the existing information and of avoiding the unfounded assumption of additional spurious knowledge.

At best, the evolutionary process can use that information and keep the population viable; at worst, the population would go extinct. Nothing guarantees that

the outcome of the process would be the former. If the population goes systematically extinct then we know that either our goal is unreasonable or something in the way the evolutionary process is implemented is not up to the task; probably the way variation is produced or genetic memory is realized. But if we end up with an extant population, and since we have not prevented the possibility of extinction, then we can be confident that the extant populations have adapted to the events they have experienced. For better or worse, these are the rules of the evolutionary game.

ACKNOWLEDGMENTS

Many thanks to Jesper Blynel for reading and commenting on a preliminary version of the manuscript.

REFERENCES

- [1] C. Adami, *Introduction to Artificial Life*. Springer, Berlin, 1998
- [2] W.R. Ashby, *Design for a Brain*, 2nd ed. Wiley, New York, 1960.
- [3] W. Atmar, Notes on the Simulation of Evolution. *IEEE Trans. on Neural Networks* 5 (1), pp. 130-140, 1994.
- [4] J.-P. Aubin, *Viability Theory*. Birkhäuser, Boston, MA, 1991.
- [5] J.-P. Aubin, *Mutational and Morphological Analysis: Tools for Shape Evolution and Morphogenesis*, Birkhäuser, Boston, MA, 1999.
- [6] J.-P. Aubin and H. Frankowska, Heavy Viable Trajectories of Controlled Systems. In J.-P. Aubin, D. Saari, and K. Sigmund, Eds. *Dynamics of Macrosystems*, Springer, Berlin, 1985.
- [7] J.-P. Aubin and H. Frankowska, *Set-valued Analysis*, Birkhäuser, Boston, MA, 1990.
- [8] T. Bäck, *Evolutionary Algorithms in Theory and Practice*. Oxford University Press, Oxford, 1996.
- [9] T. Bäck, D.B. Fogel, and Z. Michalewicz, Eds., *Evolutionary Computation 1: Basic Algorithms and Operators*. Institute of Physics Publishing, Bristol, 2000.
- [10] Bäck, T., Fogel, D.B., and Michalewicz, Z., Eds., *Evolutionary Computation 2: Advanced Algorithms and Operators*. Institute of Physics Publishing, Bristol, 2000.
- [11] P. Bak and D.R. Chialvo, Adaptive learning by extremal dynamics and negative feedback. *Physical Review E* 63, 031912, pp. 1-11, 2001.
- [12] G. Bateson, Conscious Purpose Versus Nature. In *Steps to an Ecology of Mind*. Ballantine, New York, 1972, pp. 426-439.
- [13] G. Bateson, *Mind and Nature: A necessary unity*. Fontana, London, 1980.
- [14] E.B. Baum, D. Boneh, and C. Garret, Where Genetic Algorithms Excel. *Evolutionary Computation* 9 (1), pp. 93-124, 2001.
- [15] R.D. Beer, A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence* 72, pp. 173-215, 1995.
- [16] R. Bellman, *Some Vistas of Modern Mathematics*. University of Kentucky Press, [sine loco], 1968.
- [17] S. Bennett, *A history of control engineering: 1800-1930*. Peter Peregrinus, London, 1979.
- [18] T. Bickle and L. Thiele, A Comparison of Selection Schemes Used in Evolutionary Algorithms. *Evolutionary Computation* 4 (4), pp. 361-394, 1997.
- [19] V. Braitenberg, *Vehicles: Experiments in Synthetic Psychology*. MIT Press, Boston, MA, 1984.
- [20] R.N. Brandon, *Adaptation and Environment*. Princeton University Press, Princeton, NJ, 1990.
- [21] L.W. Buss, *The Evolution of Individuality*. Princeton University Press, Princeton, NJ, 1987.
- [22] D.T. Campbell, Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes. *Psychological Review* 67, pp. 380-400, 1960.
- [23] D.T. Campbell, Evolutionary Epistemology. In P.A. Schilpp Ed., *The Philosophy of Karl R. Popper* (). Open Court, La Salle, IL, 1974, pp. 412-463.
- [24] D.T. Campbell, Unjustified Variation and Selective Retention in Scientific Discovery. In F.J. Ayala, and T. Dobzhansky, Eds., *Studies in The Philosophy of Biology*. Macmillan, London, 1974, pp. 139-161.
- [25] W.R. Cannon, *The Wisdom of the Body*. W.W. Norton, New York, 1939.
- [26] D.R. Chialvo and P. Bak, Learning from Mistakes. *Neurosciences* 90 (4), pp. 1137-1148, 1999.
- [27] F.B. Christiansen and T. Prout, (). Aspects of Fitness. In R.S. Singh and C.B. Krimbas, Eds., *Evolutionary Genetics: From Molecules to Morphology*. Cambridge University Press, Cambridge, 2000, pp. 146-156.
- [28] W.S. Cooper, Expected Time to Extinction and the Concept of Fundamental Fitness. *J. Theor. Biol.* 107, pp. 603-629, 1984.
- [29] J.C. Culberson, On the Futility of Blind Search: An Algorithmic View of "No Free Lunch". *Evolutionary Computation* 6 (2), pp. 109-127, 1999.
- [30] C. Darwin, *On the Origin of Species*. Penguin Books, London, 1859 (1968 reprint).
- [31] C. Darwin, *The Descent of Man, and Selection in Relation to Sex*. Princeton University Press, Princeton, NJ, 1871 (1981 reprint. Introduction by J.T. Bonner and R.M. May).
- [32] D. Dasgupta and Z. Michalewicz, Evolutionary Algorithms - An Overview. In D. Dasgupta and Z. Michalewicz, Eds., *Evolutionary Algorithms in Engineering Applications*. Springer, Berlin, 1997, pp. 3-28.
- [33] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*. Wiley, New York, 2001.
- [34] P.J. Den Boer, Spreading of risk and the stabilization of animal numbers. *Acta Biotheoretica* 18, pp. 165-194, 1968.
- [35] P.J. Den Boer, Exclusion, competition or coexistence? A question of testing the right hypotheses. *Zeitschrift für zoologische Systematik und Evolutionsforschung* 23, pp. 259-274, 1985.
- [36] P.J. Den Boer, Natural selection or the non-survival of the non-fit. *Acta Biotheoretica* 47, pp. 83-97, 1999.
- [37] P.J. Den Boer and J. Reddingius, *Regulation and Stabilization Paradigms in Population Ecology*. Chapman & Hall, London, 1996.
- [38] T. Dobzhansky, (). Chance and Creativity in Evolution. In F.J. Ayala, and T. Dobzhansky, Eds., *Studies in the Philosophy of Biology: Reduction and Related Problems*. Macmillan, London, 1974, pp. 307-338.
- [39] J. Dupré, Ed., *The latest on the best : essays on evolution and optimality*. MIT Press, Cambridge, MA, 1987.
- [40] M. Ehrgott, *Multicriteria Optimization*. Springer, Berlin, 2000.
- [41] T.M. English, Evaluation of Evolutionary and Genetic Optimizers: No Free Lunch. In P. Angeline, T. Bäck, L. Fogel, Eds., *Evolutionary Programming V: Proceedings of the Fifth Annual Conference on Evolutionary Programming*. MIT Press, Boston, MA, 1996, pp. 163-169.
- [42] D. Floreano, S. Nolfi, and F. Mondada, Competitive Co-Evolutionary Robotics: From Theory to Practice. In R. Pfeifer, Ed., *From Animals to Animats 5. Proceedings of the Fifth*

- International Conference on Simulation of Adaptive Behavior*. MIT Press, Cambridge, MA, 1998, pp. 512-524.
- [43] D.B. Fogel, An Introduction to Simulated Evolutionary Optimization. *IEEE Trans. on Neural Networks* 5 (1), pp. 3-14, 1994.
- [44] D.B. Fogel, *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, 2nd ed. IEEE Press, New York, 1999.
- [45] G.B. Fogel, G.W. Greenwood, and K. Chellapilla, Evolutionary Computation with Extinction: Experiments and Analysis. In *Proc. of the 2000 Congress on Evolutionary Computation: CEC00*. IEEE Press, Piscataway, NJ, 2000, pp. 1415-1420.
- [46] J. Gerhart and M. Kirschner, *Cells, Embryos, and Evolution: Toward a Cellular and Developmental Understanding of Phenotypic Variation and Evolutionary Adaptability*. Blackwell, Malden, MA, 1997.
- [47] D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA, 1989.
- [48] D.E. Goldberg and K. Deb, (). A Comparative Analysis of Selection Schemes Used in Genetic Algorithms. In G.J.E. Rawlins, Ed., *Foundations of Genetic Algorithms*. Morgan Kaufmann, San Mateo, CA, 1991, pp. 69-93.
- [49] S.J. Gould, Gulliver's further travels: the necessity and difficulty of a hierarchical theory of selection. *Phil. Trans. R. Soc. Lond. B* 353, pp. 307-314, 1998.
- [50] S.J. Gould, *The Structure Of Evolutionary Theory*. Harvard University Press, Cambridge, MA, 2002.
- [51] G.W. Greenwood, G.B. Fogel, and M. Ciobanu, (). Emphasizing Extinction in Evolutionary Programming. In Angelina, P.J., Michalewicz, Z., Schoenauer, M., Yao, X., and Zalzal, A., Eds., *Proceedings of the 1999 Congress of Evolutionary Computation: CEC99*. IEEE Press, Piscataway, NJ, 1999, pp. 666-671.
- [52] I. Harvey, Cognition is Not Computation; Evolution is Not Optimisation. In W. Gerstner *et al.*, Eds., *Proceedings of the 7th International Conference on Artificial Neural Networks, ICANN97*. Springer, Berlin, 1997, pp. 685-690.
- [53] I. Harvey, P. Husbands, and D. Cliff, Issues in Evolutionary Robotics. In J.-A. Meyer, H.L. Roitblat, and S.W. Wilson, Eds., *From Animals to Animats 2: Proceedings of the Second International Conference on Simulation of Adaptive Behavior*. MIT Press, Cambridge, MA, 1993, pp. 364-373.
- [54] I. Harvey, P. Husbands, D. Cliff, A. Thompson, and N. Jakobi, Evolutionary Robotics: the Sussex Approach. *Robotics and Autonomous Systems* 20, pp. 205-224, 1997.
- [55] J.H. Holland, *Adaptation in Natural and Artificial Systems*, 2nd edition. MIT Press, Boston, MA, 1992.
- [56] D.L. Hull, Individuality and Selection. *Ann. Rev. Ecol. Syst.* 11, pp. 311-332, 1980.
- [57] D.L. Hull, *Science and Selection: Essays on Biological Evolution and the Philosophy of Science*. Cambridge University Press, Cambridge, 2001.
- [58] D.L. Hull, R.E. Langman, and S.S. Glenn, A General Account of Selection: Biology, Immunology, and Behavior. In [57], 2001, pp. 49-93.
- [59] E. Jablonka and M.J. Lamb, *Epigenetic Inheritance in Evolution: The Lamarckian Dimension*. Oxford University Press, Oxford, 1995.
- [60] L. Keller, Ed., *Levels of Selection in Evolution*. Princeton University Press, Princeton, NJ, 1999.
- [61] B. Kerr and P. Godfrey-Smith, Individualist and Multi-level Perspectives on Selection in Structured Populations. *Biology and Philosophy* 17, pp. 477-517, 2002.
- [62] J.R. Koza, F.H. Bennett III, D. Andre, and M.A. Keane, *Genetic Programming III: Darwinian Invention and Problem Solving*. Morgan Kaufmann, San Francisco, CA, 1999.
- [63] T. Krink and R. Thomsen, Self-organized criticality and mass extinction in evolutionary algorithms. In *Proc. of the 2001 Congress on Evolutionary Computation: CEC01*. IEEE Press, Piscataway, NJ, 2001, pp. 1155-1161.
- [64] K.J. Kurman, *Feedback Control: Theory and Design*. Elsevier, Amsterdam, 1984.
- [65] H. Laborit, *La vie antérieure*. Grasset, Paris, 1989.
- [66] W. Ledermann and S. Vajda, Eds., *Algebra*. Wiley, New York, 1982.
- [67] J.R. Leigh, *Control theory: a guided tour*. Peter Peregrinus, London, 1992.
- [68] B. Lewin, *Genes VII*. Oxford University Press, Oxford, 1999.
- [69] R.C. Lewontin, The Units of Selection. *Annual Review of Ecology and Systematics* 1, pp. 1-18, 1970
- [70] R.C. Lewontin, Evolution as Engineering. In J. Collado-Vides, B. Majasanik, and T.S. Smith, Eds., *Integrative Approaches to Molecular Biology*. MIT Press, Boston, MA, 1996, pp. 1-10.
- [71] R.C. Lewontin, *The Triple Helix: Gene, Organism, and Environment*. Harvard University Press, Cambridge, MA, 2000.
- [72] R.C. Lewontin and L.C. Dunn, The evolutionary dynamics of a polymorphism in the house mouse. *Genetics* 45, pp. 705-722, 1960.
- [73] S.E. Kingsland, *Modeling Nature: Episodes in the History of Population Ecology*, 2nd ed., The University of Chicago Press, Chicago, 1995.
- [74] J. Marín and R.V. Solé, Macroevolutionary algorithms: A new optimization method on fitness landscapes. *IEEE Trans. on Evolutionary Computation* 3 (4), pp. 272-286, 1999.
- [75] P. Matzinger, An innate sense of danger. *Seminars in Immunology* 10, pp. 399-415, 1998.
- [76] J. Maynard Smith, *Evolutionary Genetics*. 2nd ed., Oxford University Press, Oxford, 1998.
- [77] J. Maynard Smith and E. Szathmáry, *The major transitions in evolution*. Oxford University Press, Oxford, 1995.
- [78] E. Mayr, *What Evolution Is*. Basic Books, New York, 2001.
- [79] B. McMullin, John von Neumann and the Evolutionary Growth of Complexity: Looking Backward, Looking Forward, *Artificial Life* 6 (4), pp. 347-361, 2000.
- [80] Z. Michalewicz and D.B. Fogel, *How to Solve It: Modern Heuristics*. Springer, Berlin, 2000.
- [81] R.E. Michod, *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality*. Princeton University Press, Princeton, NJ, 1999.
- [82] M. Mitchell and S. Forrest, Genetic Algorithms and Artificial Life. *Artificial Life* 1, pp. 267-289, 1994.
- [83] L. Moss, *What Genes Can't Do*. MIT Press, Boston, MA, 2003.
- [84] A.J. Nicholson, The self-adjustment of populations to change. *Cold-Spring Harbor Symposia on Quantitative Biology* 22, pp. 153-173, 1957.
- [85] S. Nolfi and D. Floreano, *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-organizing Machines*. MIT Press, Cambridge, MA, 2001.
- [86] J.R. Platt, Strong Inference. *Science* 146, pp. 347-353, 1964.
- [87] K.R. Popper, *The Logic of Scientific Discovery*. Hutchinson, London, 1959.
- [88] K.R. Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge, London, 1963.
- [89] T. Ray, An Approach to the Synthesis of Life. In C. Langton *et al.*, Eds., *Artificial Life II*. Addison-Wesley, Reading, MA, 1992, pp. 371-408.
- [90] K. Reeve and L. Keller, Levels of Selection: Burying the Units-of-Selection Debate and Unearthing the Crucial New Issue, In [60], 1999, pp. 3-14.
- [91] M. Ridley, *Evolution*. 2nd ed., Blackwell, Boston, MA, 1996.

- [92] A. Ruberti and A. Isidori, *Teoria dei Sistemi*. Boringhieri, Torino, 1979.
- [93] W.C. Salmon, Theoretical Explanation. In S. Korner, Ed., *Explanation*. Blackwell, Oxford, 1975.
- [94] W.C. Salmon, Statistical Explanation and Causality. In J.C. Pitt, Ed., *Theories of Explanation*. Oxford University Press, Oxford, 1988.
- [95] F. Schwerdtfeger, Is the Density of Animal Populations Regulated by Mechanisms or by Chance? In *Proceedings of the 10th International Congress on Entomology* 4, pp. 115-122, 1958.
- [96] M. Sipper, Notes on the Origin of Evolutionary Computation. *Complexity* 4 (5), pp. 15-21, 1999.
- [97] K. Sneppen, P. Bak, H. Flyvberg, and M. Jensen, Evolution as a Self-Organized Critical Phenomenon, *Proc. Natl. Acad. Sci.*, 92 (11), pp. 5209-5213, 1995.
- [98] E. Sober, *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. MIT Press, Boston, MA, 1984.
- [99] E. Sober and R.C. Lewontin, Artifact, Cause and Genic Selection. *Philosophy of Science* 49, pp. 157-180, 1982.
- [100] E. Sober and D.S. Wilson, *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Harvard University Press, Cambridge, MA, 1998.
- [101] R.V. Solé and S.C. Manrubia, Extinction and Self-Organized Criticality in a Model of Large-Scale Evolution. *Phys. Rev. E* 54, pp. R42-R45, 1996.
- [102] E.S. Vrba and S.J. Gould, The hierarchical expansion of sorting and selection: sorting and selection cannot be equated. *Paleobiology* 12 (2), pp. 217-228, 1986.
- [103] A.R. Wallace, On the Tendency of Varieties to depart indefinitely from the Original Type. *J. Proc. Linnean Society, Zoology* 3, pp. 53-62, 1858.
- [104] T.C.R. White, *The Inadequate Environment: Nitrogen and the Abundance of Animals*. Springer, Berlin, 1993.
- [105] G.C. Williams, *Adaptation and Natural Selection*. Princeton University Press, Princeton, NJ, 1966.
- [106] G.C. Williams, *Natural Selection: Domains, Levels, and Challenges*. Oxford University Press, Oxford, 1992.
- [107] D.H. Wolpert and W.G. Macready, No Free Lunch Theorems for Optimization. *IEEE Trans. on Evolutionary Computation* 1 (1), pp. 67-82, 1997.
- [108] C. Zirkle, Natural Selection before the "Origin of Species". *Proceeding of the American Philosophical Society*, 84 (1), pp. 71-123, 1941.