Adaptation to changing environments in artificial evolution

Andrea Maesani
LIS, STI, EPFL

Abstract—To survive continuous environmental changes, living organisms undergo an uninterrupted adaptation process: natural evolution. Indeed, a lot can be gained by a deeper understanding of the mechanisms of natural evolution. From an engineering perspective, this knowledge could be exploited to create better tools to solve real-world problems in which adaptation to changing environments is a critical requirement. From a biological perspective, many questions on the dynamics of natural evolution still need to be answered. Both perspectives will be considered in this research proposal.

Index Terms—artificial evolution, changing environments, dynamic optimization, evolvability, viability evolution

I. INTRODUCTION

ENGINEERS often develop real-world applications which require solving optimization problems. The environment where these applications are operated may change over time, altering the problems that are being optimized. In dynamic optimization problems (DOPs), a change in the environment can affect the objective functions, the problem instance or the constraints. In recent years, there has been an increasing interest in applying evolutionary algorithms (EAs) to DOPs. EAs are inspired by the same principles that drive natural evolution and could therefore display robust adaptation capabilities to changing environments, as nature does [1]. EAs have been applied to dynamic problems including scheduling [2], [3], optimal control in industry processes [4], [5], forecast of economic environments [6], activity prediction of document streams [7], mission planning [8], resource management [9] and dynamic routing [10]. Even though EAs have proven to be a feasible tool to solve DOPs, further research is needed to overcome some limitations that affects them when used on DOPs.

A first issue of EAs is the loss of genetic diversity during the optimization process [11]. EAs are driven by the principle of survival-of-the-fittest. Individuals are selected and reproduced according to a performance criterion, called fitness. The fitter the individual, the higher are its chances to produce offspring. Clearly, the process is biased towards the regions of the solution space which enclose the fittest individuals. Hence, the evolving population gradually loses diversity and converges. The declining diversity in the population negatively affects the performance of EAs, especially when applied to DOPs. In fact, in a changing environment the fitness landscape varies and the optima move as the landscape is altered [12], [13]. After a population has converged, it is very difficult to readapt to a new optimum when the environment changes [14], [15], [16]. Despite several approaches to limit the impact of diversity loss have been proposed in literature [1], the problem is unavoidable with algorithms based on the principle of selection-of-the-fittest.

A second issue of EAs is that they are unconstrained optimization meta-heuristics [17]. In reality, many dynamic optimization problems are constrained. Therefore, it is necessary to find ways of incorporating the constraints in the algorithm. A plethora of methods to handle constraints in EAs has been proposed [18]. Generally adopted constraint-handling techniques include penalty-functions, repair algorithms, and multi-objective optimization techniques. Penalty-function methods transform a constrained optimization problem into an unconstrained one, adding or subtracting to/from the objective function a penalty for each constraint violation [19], [20], [21], [22], [23]. Repair algorithms try to fix infeasible individuals and make them feasible [24], [25]. Lastly, multi-objective optimization techniques can solve constrained problems transforming the constraints into objectives [26], [27], [28]. Furthermore, when solving constrained DOPs, special care must be taken to avoid interferences between constraint-handling techniques and diversity preserving methods, that could eventually hamper the search process [28].
A novel evolutionary meta-heuristic, introduced by Mattiussi and Floreano [29, p. 95-96], called Viability Evolution (ViE), is currently under investigation on different scenarios at the Laboratory of Intelligent Systems. In ViE there is a shift from the principle of survival-of-the-fittest to the non-survival of the non-viable. As in nature, individuals that satisfy certain constraints can survive and reproduce, whereas non-viable individuals are eliminated from the population. In the first part of my research, I will tune and validate ViE as a tool to solve DOPs. Viability evolution is expected to maintain a high degree of diversity and can naturally handle constraints, potentially displaying high adaptability to changing environments.

In the second part of my research, I plan to use viability evolution to perform evolutionary studies. Artificial evolution has proven to be a feasible scientific tool to study evolutionary dynamics [30]. Frequently, the time scales at which natural evolution acts do not allow to perform direct experiments or observations using reasonable resources and time. Artificial evolution can overcome these problems and provide a controlled framework to perform experiments. Moreover, ViE is expected to produce evolutionary dynamics very different from EAs, that could better abstract natural dynamics. By means of artificial evolution, I will investigate the effects produced by environmental changes on the evolution of genomes. My research could eventually provide useful hints to engineers, to design better representations for EAs, and to biologist to better understand natural mechanics. Previous work showed that evolution in changing environments can lead to modular solutions [31], and can enable a significant increase in adaptation [32]. Parter et al. [33] reported a detailed study on the evolvability of genomes in changing environments. However, their results are based on evolution performed with fixed-length genomes. To better model natural dynamics, it would be favourable to use variable-length genomes that display characteristics similar to natural genomes. I plan to use viability evolution, coupled with a variable-length representation to study the evolvability of genomes in changing environments. A suitable representation is the Analog Genetic Encoding (AGE), which shares many similarities with biological genomes. First, it is organized in genes that interact regulating their expression. Second, single point mutations in the genome can trigger large phenotypic changes, simulating the complex interactions between the genome and the regulatory and developmental system of living organisms. Third, it allows for neutral walks through a many-to-many genotype-to-phenotype mapping. Fourth, it supports non-coding regions.

In this research proposal, I start presenting viability evolution in section II. Then, I discuss three papers related with the proposed research. In section III, I survey the evolutionary approaches proposed in literature to deal with changing environments. In section IV, I describe the work of Parter et al. on the evolvability of fixed-length genomes in changing environments. In section V, I present the AGE representation. Finally, I detail my research plan in section VI.

II. VIABILITY EVOLUTION

Traditional evolutionary computation methods abstract the concept of survival of the fittest. Individuals are sorted according to a performance measure, the fitness function, and the ones that possess higher fitness are selected and reproduced. There are, however, some remarkable differences between the natural process of evolution and the one modelled by evolutionary algorithms. P.J. den Boer [34] argues that natural evolution acts by means of non-survival of the non-fit, rather than survival of the fittest. In nature, individuals that are not viable in a certain environment, do not survive. On the contrary, in evolutionary algorithms survival is only determined by the fitness of an individual. Secondly, there is a substantial difference between biological fitness and the fitness functions used in EAs. Biological fitness expresses the reproductive success of an individual and can be measured only at the end of the life of an individual. On the contrary, the fitness used in EAs is computed before reproduction, and determines the reproductive success of an individual, biasing the evolutionary process.

Mattiussi and Floreano [29, p. 95-96] propose a novel meta-heuristic that incorporates the concept of non-survival of the non-viable, Viability Evolution (ViE). They identify different evolutionary events which affect the evolutionary process. On one hand, individuals can be ordered by some criterion and selected consequently (selection events). On the other hand, individuals may have to satisfy some conditions to survive. The individuals that do not satisfy the viability constraints are eliminated from the population (elimination events). Therefore, we say that an individual is considered viable if it satisfies the viability constraints. The intersection of the constraints defines the viability space, where individuals can live and reproduce. Offspring that fall outside the viability space are eliminated. Moreover, contingency elimination events can eliminate from the population a number of randomly chosen individuals. In viability evolution, the reproductive success of an individual is no more determined before reproduction, using the fitness function, but depends uniquely on its viability. The viability constraints are dynamic and change over the evolutionary process. The temporal trajectory of the viability space describes the viability tube, depicted in Figure 1. The viability constraints are relaxed at the beginning of the evolutionary process, allowing a large number of viable individuals. While evolution proceeds, the viability tube is shrunk enforcing tighter constraints on the population, until the target constraints are reached.

Fig. 1. The viability tube represents the evolution of the viability space $V(t)$ over the evolutionary time $T$. An individual $i$ that remains viable over the whole time interval $T$, defines a viability trajectory $v(t)$ in the viability tube.
The ViE paradigm could display a number of advantages over classical evolutionary algorithms. First, it is expected to produce and maintain a higher degree of diversity in the evolving populations. The increased diversity could be exploited actively by the evolutionary process to adapt faster to changing environments, leading to better performances when solving DOPs. Second, it allows for a natural way of handling constraints, without recurring to techniques to transform an unconstrained into a constrained optimization process. Third, because it abstracts the principle of non-survival of the non-viable it may be a better tool to perform computational biology experiments, rather than traditional EAs. Finally, dynamic environments may be modelled directly by the dynamics of the viability tube.

III. EVOLUTIONARY OPTIMIZATION IN DYNAMIC ENVIRONMENTS

In the first surveyed paper of my research proposal, Jin and Branke [1] classify the different types of uncertainties that evolutionary algorithms might have to take into account when solving optimization problems. First, the evaluation of the fitness function may be subject to noise. Second, the design variables may change slightly after the optimal solution is found. Therefore it is favourable to find solutions that are robust to these changes. Third, if the computation of the fitness function is very expensive, it could be approximated by a less computational intensive fitness function. Fourth, the fitness function, the problem instance or the constraints may change over time and thus the optimum of the problem may change as well.

The first three categories of uncertainties are out of the scope of this thesis proposal and will not be treated in more details. Here, I analyse and update their review on evolutionary optimization in dynamic environments, considering more recently published research. In fact, to compare and improve the performance of ViE on DOPs, it will be fundamental to understand the strong points and the weaknesses of the different published approaches. The discussion is restricted to the approaches involving evolutionary algorithms. A description of other meta-heuristics to solve DOPs, and of the standard adopted benchmarks and metrics to measure the performance of EAs in dynamic environments, can be found in the original paper. Jin and Branke start their discussion acknowledging that to solve DOPs, it may be beneficial to retain information accumulated in the evolutionary process, rather than restarting the whole process from scratch. They observe that since EAs have much in common with natural evolution and since natural evolution has coped with continuous adaptation to changing environments, they are a suitable candidate to solve DOPs efficiently. Unfortunately, EAs have a tendency to converge, at least when the environment has been static for some time. Jin and Branke classify the main approaches to address the convergence issue in dynamic environments into the following categories.

A. Approaches that generate diversity after a change

The approaches that belong to this category take explicit actions to increase diversity in the population, as soon as a change in the environment is detected. Cobb [35] proposes an approach called triggered hypermutation. Whenever a change is detected, the mutation rate is increased for a number of generations to produce additional diversity in the population. Morrison and De Jong [36] performed further experiments on triggered hypermutation showing that the choice of a good mutation rate is dependent on the period between environmental changes. Also, some environmental change may fail to trigger hypermutation [11], e.g. when the environment changes abruptly [14]. In addition, notable performance degradation can be expected if the mutation rate exceeds the required amount or if it is not sufficient [37]. A different mechanism to increase diversity is implemented in the variable local search approach [38]. The method performs an adaptive local search around the solutions in the population whenever a change is detected. If unsatisfactory solutions are found, the range of the local search is extended. In presence of relatively small changes it can produce good solutions also at high frequencies of change [37]. Finally, Maresky et al. [39] proposed another approach to introduce diversity, called selectively destructive restart. When environmental changes are detected, the EA is restarted. During the restart, only a number of genes of the individuals are reinitialized. The performance of the approach has been found to be highly dependent on the gene reinitialization probability [40].

The main weakness of the approaches that introduce diversity is the detection of environmental change. If environmental changes are not detected, the approaches fail to trigger the mechanisms to introduce diversity, leading to poor performances. Moreover the right amount of diversity to be introduced is difficult to define. If it is too high, it can potentially disrupt previously accumulated information in the population. If it is too low, it can have no effect.

B. Approaches that maintain diversity throughout the run

The approaches that fall in this category try to avoid convergence in the population. It is hoped that a population spread over the solution space could better withstand environmental changes and adapt faster. A classic approach is random immigrants [11], i.e. replace at each generation a fraction of the population with randomly generated individuals. Random immigrants are beneficial for non-deceptive functions, but may deteriorate performance on deceptive ones [41]. The approach is reported to work well with occasional, large changes in the optimum position [14]. In a slowly changing environment, random immigrants may introduce too much diversity and therefore degrade performance [42]. In fact, the diversity introduced by random immigrants is randomly generated and may interfere with the search process. A self-organizing version of random immigrants is proposed in [43]. The less fit individual and a number of individuals around it are replaced with random individuals. Selection for reproduction takes place separately for the new random immigrants and for the individuals that have not been replaced. In fact, using a normal regime of selection, the random immigrants are soon eliminated from the population, because the probability that the fitness of a new random individual is higher than the fitness
of current individuals is generally low. Another approach to maintain diversity is represented by the thermodynamic genetic algorithm [44]. The individuals that maximize the diversity in the population are selected to be retained in the next generation. The so-called temperature parameter defines the emphasis on diversity in the selection procedure. The approach shows satisfactory performances for environments that change slowly. An adaptive mechanism to modify the temperature parameter produces better results [45]. A different way to maintain diversity in the population is exploited through multi-objective optimization. A multi-objective EA is tested with different objectives and reported to perform better than random immigrants when explicit objectives to maintain diversity are used [16]. Other methods to avoid population convergence, include the multi-niche crowding genetic algorithm [40] and the primal-dual genetic algorithm [46], [47].

The multi-niche crowding genetic algorithm preserves niches of similar individuals to maintain diversity. Individuals can mate only with similar ones. The primal-dual genetic algorithm evaluates the dual individual, i.e. the bitwise complement, of a subset of the individuals in the population. If the fitness of the dual is better than the primal individual than the primal is replaced by its dual. The number of evaluated duals depends on the number of primal-dual replacements in the previous generation. The approach seems robust especially when the environment changes significantly. Finally, other methods try to maintain diversity modifying the fitness evaluation phase or the selection mechanism in an EA. For example, Ghosh et al. [48] introduced a mechanism to age individuals. Middle-aged individuals are considered more fit than new or older ones. The idea is to avoid any individual to dominate for a long number of generations. Instead, [49] presented an adaptive selection pressure technique that was tested on different approaches. Its performance was reported to be highly dependent on the problem.

The main pitfall of this class of approaches is that they are creating variation that is not useful for the evolutionary process to maintain diversity. Moreover, this increased non-useful diversity can interfere with the search process degrading the performances. Some of the approaches are highly dependent on the type of problem and cannot be used reliably on DOPs.

C. Memory-based approaches

EAs coupled with memory-based approaches are equipped with implicit or explicit mechanisms to store information from previous generations. Implicit memory approaches store the information at the level of the genome. On the contrary, explicit memory approaches implement some policy to store a number of individuals from previous generations.

The most popular examples of implicit memory approaches is diploidy. Goldberg and Smith [50] use a triallelic scheme in which dominance can be evolved, i.e. the recessive allele can become dominant and vice versa. The dominant allele is expected to shield the recessive one, allowing for long-term memory. The scheme seems suitable only to account for environments that oscillates between two states [37], [40]. Goldberg and Smith argue that the diploid triallelic scheme outperforms a haploid one. Successively, Ng and Wong [51] show that actually the haploid scheme can work better than the triallelic scheme after steady state has been reached by the genetic algorithm. In the same paper, they propose a novel diploid scheme that performs better than the triallelic scheme. Ryan [52] proposes to use an additive diploid scheme. Lewis et al. [53] compared the three approaches concluding that none of them can deal with environments that switch between more than two environmental states. More recently the diploid dominance genetic algorithm is shown to outperform all the three previously cited diploid schemes [54]. The allele dominance is computed at each locus maintaining a dominance vector which is updated with a policy similar to the one used by the population based incremental learning meta-heuristic [55].

The structured genetic algorithm [56] is another attempt to use a redundant representation to act as a memory and favour adaptation to environmental changes. Again, the approach is demonstrated to work only when the environment switches between only two configuration.

Explicit approaches include the extension of previously described algorithms with a memory to store individuals at certain generations, like the memory-based random immigrants [41], [42] and the memory-based thermodynamic genetic algorithm [57] which both have better performance on cyclic environments than their versions without memory. Branke [58] introduces the memory search genetic algorithm, where two populations are maintained. The search population is restarted for every environmental change, while the memory population retains solutions that had high fitness in previous environments. The transformation-based genetic algorithm [15] keeps a pool of genome fragments and uses them to transform individuals after mating.

Memory approaches are useful only if the environment is somehow cyclic and returns periodically to a previous state. Implicit memory methods seem able to store only information on a restricted number of previous state, and are unable to deal with environments that change between many different states.

D. Multipopulation approaches

Multipopulation approaches employ several populations to search independently in different regions of the solution space. Examples include the multinational, the shifting balance and the self-organizing scouts genetic algorithms. The multinational genetic algorithm [59] groups the individuals at the end of each generation in different sub population, separated by valleys in the fitness landscape. The shifting balance algorithm [60] maintains a core population that tracks the current optimum and several colony populations whose purpose is to explore the search space. When the colonies are too close to the core population, their objective is changed to maximize the distance from the core. The self-organizing scouts [61] approach forks the population continuously in sub-populations when a certain convergence criterion is achieved. Forked sub-populations are forced to search in a specific region of the search space.

Multipopulation approaches seem to be the more efficient respect to the other categories [62], because they effectively
try to maintain an useful level of diversity, intelligently exploring the search space. However, this comes at the cost of very high computational requirements. Moreover they cannot be used with representations that allows for many-to-many genotype-to-phenotype mappings, avoiding their employment in problems that require the use of indirect encodings of the solutions.

Understanding the real performances of the methods discussed in the four aforementioned categories is complicated by a general lack of published results on standard benchmarks. Anyway, every method face the unavoidable issue of convergence and loss of diversity caused by their driving evolutionary principle, selection of the fittest. On the contrary, viability evolution appears a better candidate for solving DOPs, because of its different search dynamics, which preserves helpful diversity in the population. Moreover, the ViE approach does not depend on the representation used.

IV. EFFECTS OF VARYING ENVIRONMENTS ON GENOMES

In the second reviewed paper, Parter et. al [33] analyse the evolvability of fixed-length genomes in changing environments. In their study, they consider the theory of facilitated variation (FV). The theory addresses the capacity of organisms to generate novel phenotypes which are potentially useful. FV is made possible by weak regulatory linkage and modularity, among other factors described in details in the original work of Gerhart [63]. Weak regulatory linkage is defined as the capability to trigger large pre-prepared phenotypic variation using weak signals. For example, single mutations in the genome of an organism can have huge effects on the phenotype, if the changes affect some genes that regulate the development of the organism. Modularity helps to avoid interference of mutations between separate functional components of an organism. Parter et al. restricted their study on the emergence of facilitated variation to a particular class of changing environments. Their assumption is that environmental goals faced by organism are combinations of sub-goals. Even though the environment may change, the organisms encounter a new goal that still contains a subset of the same sub-goals. The environments used in their experiments, called modularly varying goals (MVG), can change due to the modification of one or more of these sub-goals. They use two model systems: combinatorial logic circuits composed of NAND gates and RNA sequences that are folded into their secondary structure. Parter et al. compared the outcome of the evolutionary process under stationary and MVG. In the case of combinatorial circuits, Parter et al. compared the outcome of the evolutionary process under stationary and MVG. In the case of RNA secondary structures the different goals share the same library of structural modules. They find that genotypes evolved in modular varying environments can adapt more rapidly when the goal is changed. Moreover, they conclude that in modularly changing environments the phenotypic neighbourhood 1 of the individuals have higher fitness when exposed to previously seen goals than what observed in stationary goal evolution. This could indicate that the evolved organisms effectively store information about past goals migrating their genomes towards the borders of the neutral network 2 of a specific phenotype. Parter et al. can also identify weak regulatory linkage induced by specific mutations in the genomes of the evolved individuals in varying environments, which they call genetic triggers. Furthermore, they analyse the ability of the evolved populations to adapt to unseen goals, showing that populations evolved under MVG can not only store information about past goals but also generalize for future goals. Interestingly, they find that the genetic variance of individuals evolved under stationary environments is higher than the diversity reached by the population under varying environments. Even though the outcome of their experiment seems to contradict what happens in nature, this behaviour can be expected when fixed-length genomes are used. In fact, because the evolutionary process can only use a genome of fixed length, genomes evolved under stationary environments are more robust to mutations, allowing more diversity, while genomes evolved under MVG are more sensitive to mutations.

Finally they propose a measure for facilitated variation. According to the proposed measure organisms with high FV are likely to form potentially useful variation and have a relatively low probability of varying towards non useful phenotypic direction. In line with the rest of the discussion, individuals evolved under MVG display higher facilitated variation, measured according to their proposed formula.

Parter et al. performed an evolvability study on fixed-length genome. However, natural evolution acts on variable-length genomes. Some of the results obtained by Parter et al. seem strongly biased by the restrictions posed by the fixed-length genome, i.e. the results on the diversity of the populations. Therefore, to perform an evolvability study in conditions closer to natural evolution, we need a representation that could grow in size and complexity.

V. ANALOG GENETIC ENCODING

The third reviewed paper introduces the representation that I will use for evolutionary studies in changing environment. The design of networks that represent real-world systems is a challenging process. Often, the optimal size and the connectivity of the network for a specific task is not known in advance and its manual design may encounter several difficulties. Mattiussi and Floreano [64] propose the Analog Genetic Encoding (AGE) representation to evolve the topology and the sizing of networks. AGE realizes an implicit encoding of an analog network, i.e. an hyper-graph composed by devices (nodes) interconnected by links (edges) that have an

1The phenotypes reachable applying a single point mutation to the genotype of the individual.

2A neutral network of a phenotype is composed of all the genotypes that express that specific phenotype.
associated numerical value. The devices may be of different kind and have several ingoing and outgoing ports, to which other devices can connect, called device terminals. The AGE genome is constituted by one or more chromosomes, strings of characters selected from a finite genetic alphabet. The actual devices encoded in the genome are separated from the non-coding regions by special sequences, the device and terminal tokens. Between the device token and the terminal token, one or more sub-sequences of characters encode the values of the device parameters. Moreover, it contains a sub-sequence for each device terminal. The genome is decoded into an actual device and a terminal token are decoded. A decoded device is associated with a list of sequences that represents its terminals and parameters.

Devices are decoded by means of a decoding algorithm. This algorithm requires to decode a network from its genome. It marks, it suffers of certain drawbacks. First, the scalability of the approach may be hampered by the time complexity of the algorithm required to decode a network from its genome. In fact, the terminals of every device in the network must be aligned with all the other terminals to compute the interactions between the devices. Second, the evolution of large sparse networks may be difficult due to the innate propensity of AGE to generate interactions between all the devices. Nonetheless, it is particularly well suited for my purpose of studying the evolvability of artificial genomes in changing environments, because of its support for non-coding regions and its ability to display weak regulatory linkage.

VI. Research Proposal

The aim of my research is twofold. First, I want to investigate the adaptation capabilities of viability evolution in changing environments. I strongly believe that viability evolution can contribute significantly to engineering as a tool for solving DOPs. Second, I want to understand the impact of environmental changes on evolved genomes. If successful, my research could eventually lead to cues on the dynamics of the effects of natural evolution on real genomes. The temporal plan of the proposed research is showed in Figure 3.

In the first work package (WP1), I will validate viability evolution as a tool to solve DOPs. ViE will be evaluated using a set of standard benchmark problems, namely the moving peak benchmark [70], the dynamic problem generator proposed by Yian [71] and the dynamic constrained benchmarks introduced by Nguyen [28]. The chosen benchmarks allow to compare the capabilities of ViE with state-of-the-art EAs. The moving peak benchmark [70] allows to study the capability of a meta-heuristic to track a continuously changing optimum. The dynamic problem generator proposed by Yian [71] allows to check the behaviour of a meta-heuristic on different problem properties, e.g. deception, multi-modality, etc. Finally the benchmarks proposed by Nguyen [28] allow to test ViE on problems with changing constraints. The algorithms will be compared using many metrics, including accuracy, stability and reactivity. A review of the metrics used in dynamic optimization that can be employed for the comparison, can be found in [72]. The validation of ViE on such a broad set of benchmarks will enable the refinement of the viability framework. Possible improvements include investigating novel
methodologies to update the viability tube, mechanisms to control the amount of selection/elimination events and the number of generated offspring per individual depending on the dynamics of the environment. Ideally the first work package will be coupled with one or more case-studies of real-world DOPs that will be solved using ViE, possibly through projects assigned to students.

In the second part of my research, I plan to use the ViE framework as a scientific tool to perform evolutionary studies. As already discussed in this document, viability evolution permits to better model natural evolutionary dynamics rather than standard EAs. Evolutionary studies are always complicated or made impossible by the fact that natural evolution acts on incredibly long time scales, and long-term evolution experiments in complex environments require unreasonable investments of time and resources [73]. Although experiments performed by means of artificial evolution are forced to abstract many of the details of the modelled systems, they can be a feasible tool to investigate many evolutionary dynamics. Moreover, artificial evolution allows to record the whole genetic patrimony generated by an evolutionary process, enabling detailed analysis that would been impossible with in vivo or in vitro experiments.

In the second work package (WP2), I will study the effects of environmental changes on evolvability. The modes of creation of phenotypic variation are still largely unknown [63]. Indeed, previous work show that evolvability is a trait selectable by evolution. Earl and Deem [74] evolved individuals that could tune the rates of the mutational processes that act on their genome. They discovered that, depending on the frequency of changing environments, the individuals adapt their mutational processes. Parter et al. [33] performed a study on the theory of facilitated variation. They also confirm that individuals evolved in changing environments are more evolvable. Moreover, they showed that evolution shapes the genomes to allow large phenotypic variation with small changes. In this work package, I will build on the previous research employing variable-length genomes, rather than fixed-length ones as used in both papers. Therefore, I plan to couple ViE with a variable-length representation, the Analog Genetic Encoding (AGE) [64]. I am interested in understanding how the evolvability of organisms with variable-length genomes is affected by environmental changes. Furthermore, variable-length genomes enable the study of the dynamics of evolution on non-coding regions. In fact, evolution could take advantage of redundancy and neutral mutations in non-coding regions. Do non-coding regions play a strong role in evolution? Are non-coding regions actively exploited by the evolutionary process to increase evolvability? What can be inferred about the past environmental history of a population analyzing the non-coding regions in the genomes? To provide answers to these questions I plan to set up two simple experiment that will allow the investigation of these issues. To ease comparison with previously published research, I will use the same model systems used in the work of Parter et al. [33]. During the first experiment, I will evolve populations that undergo different environmental histories, i.e. sequences of environmental changes of defined severity. At the end of the evolutionary process, a study on the evolvability properties of the population will be performed. I want to measure the evolvability of the populations and correlate it with the environmental history. In a second experiment, I will look for clear evidence on the role of non-coding regions. An experimental and a control population will be subject to the same environmental history. After a certain number of generations, the non-coding regions will be totally removed from the experimental population. Then, evolution will be resumed and the successive evolutionary history of the experimental population will be compared with the control population. Finally, I will examine the evolved genomes and try to find in them indicators of the previous environmental history.

The work on non-coding sequences and changing environments could provide cues for another big issue in biology: the variability of genome sizes. In the third work package (WP3), I will investigate the effects of changing environments on the evolution of genome sizes. In nature, the genome sizes of living organisms vary greatly, spanning over different orders of magnitude, without any apparent links to organism complexity.
Even though the mechanisms that allow for genome length increase and decrease are well known, why evolution has resulted in such wide variation in genome sizes is still an open question. Is the genome size affected by the environmental history encountered by a population? Gregory [75] reviewed a number of proposed hypotheses: 1) the neutral evolution hypotheses posit random accumulation of DNA; 2) the nucleotypic hypothesis claims that DNA content affects cell size and division rate; 3) the nucleoskeletal hypothesis state that genome size is selected to meet a required cell volume to balance growth; and 4) Lynch and Conery [76] posit that lineages with small population sizes evolve larger genomes as a consequence of reduced efficacy of natural selection over small populations. The proposed hypotheses can hardly explain the observations [75], [77]. I will test the hypothesis that the environmental history of an organisms affects its genome size. I will investigate the effects of multiple parameters on the genome size, primarily considering the impact of environmental changes. The experimental setup for this work package will include the possibility to vary frequency and severity of the changes and the influence of genetic operators that increase or decrease the genome size.

VII. CONCLUSION

In this research proposal, I reviewed the various approaches to deal with dynamic environments in artificial evolution [1]. The lack of methods that could maintain satisfactory adaptability lead us to the introduction of viability evolution. Then, I described the research of Parter et al. [33] on facilitated variation. In their work they considered fixed-length genomes to study the effects of changing environments on the evolvability of variable-length genomes we needed a suitable representation. Therefore, I presented the Analog Genetic Encoding (AGE) [64]. Finally, I detailed my research plan. My research could contribute substantially to both fields of engineering and evolutionary biology. On one side, viability evolution could be demonstrated as a powerful tool to solve DOPs. On the other side the investigation of evolvability of variable-length genomes could provide hints to biologist on natural evolutionary dynamics and could serve for engineers as a basis to design better representations for evolutionary computation methods.

REFERENCES


