# Chapter 10
# Diversity Management in Memetic Algorithms

Ferrante Neri

## 10.1 Introduction

In Evolutionary Computing, Swarm Intelligence, and more generally, population-based algorithms diversity plays a crucial role in the success of the optimization. Diversity is a property of a group of individuals which indicates how much these individuals are alike. Clearly, a group composed of individuals similar to each other is said to have a low diversity whilst a group of individuals dissimilar to each other is said to have a high diversity. In computer science, in the context of population-based algorithms the concept of diversity is more specific: the diversity of a population is a measure of the number of different solutions present, see [239].

Ideally, a population-based algorithm is supposed to work in high diversity conditions during the early stages of the optimization, then progressively lose diversity in proximity to the global basin of attraction, and eventually lose all diversity when the global optimum is detected. The latter condition means that the entire population is characterized by a unique genotype, i.e. the global optimum. The described functioning happens very rarely in practice since the algorithm tends either to prematurely converge to a suboptimal solution or to stagnate. Premature convergence is an undesired condition, which very often jeopardizes the functioning of Evolutionary Algorithms (EAs), consisting of a diversity loss in the presence of a sub-optimal (and unsatisfactory) candidate solution, see [246]. Stagnation is typical of Swarm Intelligence Algorithms (SIAs) but is present also in some EA structures. An algorithm stagnates when it does not succeed at enhancing upon its individual with the best performance while the diversity is still high. In other words, the algorithm repeatedly explores less promising areas of the decision space and thus does not manage to register improvements.

Due to their different structures, EAs and SIAs require different and complementary techniques for handling diversity. More specifically, in EAs a mechanism which

Ferrante Neri
Department of Mathematical Information Technology, P.O. Box 35 (Agora),
40014 University of Jyväskylä, Finland
e-mail: ferrante.neri@jyu.fi

F. Neri et al. (Eds.): Handbook of Memetic Algorithms, SCI 379, pp. 153–165.
springerlink.com 

preserves diversity and thus inhibits premature convergence is beneficial, while such an approach in SIAs can be detrimental and turn into stagnation behavior.

In Memetic Algorithms (MAs), since their earliest definition in [621] and early original works in Memetic Computing (MC), see [615] and [622], the problem of diversity is taken into account and implicitly analyzed. Since MAs perform the search by employing multiple search logics, diversity is preserved by studying the decision space under complementary perspectives, see [489]. This means that if the search logic within the evolutionary framework fails at detecting new promising solutions, the local search components give an extra chance to the algorithm to detect fresh and promising genotypes. This is probably one of the main reasons contributing to the success of MAs.

However, as remarked in [239], MAs by themselves are not a "magic solution" to optimization problems, and the employment of multiple search logics does not guarantee the prevention of premature convergence or stagnation. For example, a MA based on an evolutionary framework and employing local search components can naturally lose diversity since the application of the local search to a set of points belonging to the same (sub-optimal) basin attraction would produce the convergence of a part of the population to the corresponding local optimum.

In order to prevent MAs from premature convergence and stagnation, several approaches attempting to handle population diversity in MAs have been proposed during recent years. This chapter deals with diversity in MAs and presents a survey of techniques recently proposed in literature for handling diversity and coordinating the various algorithmic components contained within MAs. Section 10.2 gives a short survey on the topic. Section 10.3 focuses on Fitness Diversity adaptation and, presents various diversity metrics and the related adaptation techniques.

## 10.2 Handling the Diversity of Memetic Algorithms: A Short Survey

Most of the MAs proposed in the literature employ an evolutionary framework (and not a swarm intelligence framework). Thus, most of the work on diversity attempts to preserve diversity and prevent premature convergence.

A classical and straightforward approach has been proposed in [246] where a generational Genetic Algorithm (GA) employing truncation selection is proposed. The algorithm randomly pairs parents; but only those string pairs which differ from each other by some number of bits (i.e., a mating threshold) are allowed to reproduce. In this way, diversity is preserved by inhibiting the presence of duplicates. A similar approach has been proposed with reference to an engineering problem in [863] and [205].

In [640] the problem of diversity is handled by employing a structured population. A distributed GA and a local search algorithm process the entire population. The sub-population evolves independently and thus preserves the diversity of the entire population.

In [648] a local search crossover is integrated within the evolutionary framework. The basic idea of this local search crossover is to remove and replace genes in a selected parent solution on the basis of its common and different edges with the other parent solution. As a result, the offspring is genotypically different from the parents and diversity is preserved.

In [581] a specific crossover for preserving the diversity is proposed. This crossover keeps constant the Hamming distance (i.e. the number of genes in a candidate solution at which the corresponding symbols are different) between parents and offspring. Moreover, in [581] a restarting mechanism is proposed. This simple (and sometimes efficient) mechanism consists of resampling the individuals of the population in the presence of diversity loss and possible premature convergence.

In [491] a MA composed by a GA and an adaptive local search algorithm is proposed. This adaptive local search is inspired by Simulated Annealing, see [468] and [122], and is supposed to improve upon the available genotypes when the population is diverse and to increase the diversity when the population is approaching the convergence condition. The diversity preservation logic proposed in [491] can be summarized in the following way: a solution which is slightly worse than the starting one can be accepted under the condition that it increases the diversity in the population. More formally, for a given minimization problem and for a given candidate solution $x$, a newly generated solution $x'$ replaces $x$ according to the following probability:

$$P = \begin{cases} 1 & \text{if } f(x') \leqslant f(x) \\ e^{\frac{k|f(x')-f(x)|}{|f_{min}-f_{avg}|}} & \text{otherwise} \end{cases} \tag{10.1}$$

where $f_{min}$ and $f_{avg}$ are, respectively, minimum and average fitness values among the population individuals and $k$ is a normalization constant. This technique measures the diversity by means of the fitness value and is strongly related to the fitness diversity adaptation which will be extensively discussed in Section 10.3.

In [492] the encoding of memetic information (in the mentioned paper mutations for some problems and local search algorithms for another problem) is performed within the solutions. A probabilistic criterion manages the transmission of the memes and thus search strategies from parents to offspring. In [492], multiple local search algorithms are employed, de facto composing a multimeme algorithm, see [496] and [489]. The resulting algorithmic structure is supposed to prevent premature convergence by offering multiple search perspectives of the decision space. The main algorithmic philosophy is that the combination and coordination of a set of various search logics enhances the chance of obtaining a high performance or, more modestly, at least overcome the bottlenecks resulting from the specific limitations of a certain search structure. For example the employment of a local search algorithm employing a steepest descent pivot rule can be efficient in the proximity of the global optimum when it is important to finalize the search by exploiting the neighborhood while a random walk algorithm can support the evolutionary framework to detect new promising directions when the search still has not detected a promising direction. If a MA employs both these local searches, it might be able to handle both the situations. In addition, the adaptation is supposed to allow the algorithm to decide itself the most proper local search on the basis of the situation. The employment and thus coordination of multiple local search algorithms within a MA is a crucially important topic in Memetic Computing and is somehow the "hearth" and the reason for success/unsuccess of a MA. Some examples of studies on this specific topic are reported in [411], [493], [683], [830] and references therein.

In [806] a MA for clustering is proposed. Two modified selection schemes based on fitness assignment concur at global and local levels to preserve diversity and to prevent premature convergence. In [715], a MA for solving multimodal problems is presented. The concept of fitness sharing is extended to the local search algorithms, thus defining Baldwinian sharing. In practice, the algorithm employs a sharing technique (i.e. a normalization of the fitness values based on the Euclidean distances to affect the sorting/selection and thus prefer a population composed by spread out points) in order to guarantee that diversity is preserved.

In [536] a real-coded MA is proposed. Within this MA two mechanisms for preserving the diversity are employed. The first mechanism, namely negative assortative mating, consists of selecting genotypically distant parents in order to obtain an offspring which does not look similar to either generating parent. The second mechanism, namely Breeder Genetic Algorithm (BGA) mutation [639], is a mutation operator which promotes the generation of distant genes within the solutions by employing an ad-hoc probability distribution function.

In [873] the problem of diversity is handled by using multiple search logics and a structured population. Two adaptive systems for preserving diversity are also presented. Both mechanisms rely on the fact that the frequency of the local search helps to preserve diversity. According to the first adaptive system, at the beginning of the optimization process the sub-populations already contain enough diversity and therefore do not need additional search moves coming from the local search; hence the local search algorithms are activated with a low frequency. Subsequently, since the population naturally tends to progressively lose diversity, the local search is activated with a higher frequency. More specifically, the frequency $\gamma$ of local search activation is given by the following heuristic rule:

$$\gamma = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{gen-\mu}{\sigma}\right)^2\right)\eta \tag{10.2}$$

where $\mu$ and $\sigma$ are mean value and standard deviation of a Gaussian distribution, *gen* is the generation number, and $\eta$ is a scaling factor.

The second adaptation system is more complex and less intuitive compared to the first one. In order to explain this mechanism, let us consider a (sub-)population $S$ of individuals. The population can be partitioned into $Q$ groups $S_1, S_2, \ldots, S_Q$ where each group contains individuals characterized by the same fitness value. With reference to the generic $j-th$ group, we can define the ratio $p_j$ as:

$$p_j = \frac{|S_j|}{\sum_{i=1}^{Q} |S_i|} \tag{10.3}$$

where with $|*|$ is indicated the cardinality of the set, i.e. how many individuals belong to a given group. On the basis of the described partitioning, Shannon's information entropy, see [775], is defined as:

$$E = -\sum_{j=1}^{Q} p_j \log(p_j). \tag{10.4}$$

For a given population the entropy can be considered as a fitness-based diversity measure. In [873] the entropy variation is used to determine the amount of local search to be employed. More specifically the diversity frequency at the generation *gen* is given by:

$$\beta(gen) = 1 + \frac{E(gen) - E(gen-k)}{E(gen-k)} \tag{10.5}$$

where $E(gen)$ and $E(gen-k)$ (where $gen \geqslant k$) are the population entropy measure at the $gen-th$ and $(gen-k)-th$ generation, respectively.

## 10.3 Fitness Diversity Adaptation

Fitness Diversity Adaptive MAs are a class of algorithms which, like other works e.g. [491] and [873], measure fitness diversity in order to estimate the population diversity. This choice is done considering that for multi-variate problems the measure of genotypical distance can be excessively time and memory consuming and thus the adaptation might require an unacceptable computational overhead. Obviously, fitness diversity could not give an efficient estimation of population diversity, since it can happen that very different points take the same fitness values, e.g. if the points lay in a plateau. However, this fact does not effect the decision mechanism of the adaptive system for the following reasons.

The MAs employing Fitness Diversity Adaptation (FDA) are composed of an evolutionary framework and a list of local searchers. The coordination of the local search is carried out by the fitness diversity. More specifically, when the diversity is low one or more explorative local searchers, e.g. Nelder-Mead Simplex [653], are activated in order to offer an alternative search logic, and possibly to detect new promising search directions and increase the diversity. If this mechanism fails and the algorithm keeps losing diversity and converging to some areas of the decision space an exploitative local search algorithm, e.g. Rosenbrock Algorithm [776], attempts to quickly perform the exploitation of the most promising basin of attraction and thus quickly complete the search. If the fitness diversity is low, the candidate solutions in the population have a similar performance. This fact can mean either that the solutions are concentrated within a small region of the decision space, or that the solutions are distributed over one or more plateaus or over two or more basins of attraction having a similar performance. It can easily be visualized that all the listed situations are undesirable and that the activation of an alternative search move can increase the chances to detect "fresh" genotypes. In other words, although the FDA does not guarantee a proper estimation of the population diversity, it is an efficient index to estimate the correct moment of the evolution which would benefit from a local search application.

Although the fitness diversity mechanism sounds reliable at first, it hides two practical issues when the algorithmic design is performed. The first issue is how to measure the diversity while the second is how to use the diversity information for coordinating the local and global search. The following subsections address these two problems.

### *10.3.1 Fitness Diversity Metrics*

Before analyzing the various metrics presented in the literature for measuring diversity a comment on the approach is necessary. As highlighted in [657], there is no "best" metric in general but there is a "most suitable" metric dependent not only on the problem (i.e. the fitness landscape) but also on the nature of the evolutionary framework. For example, an efficient diversity metric for Evolution Strategy (ES) would likely be inadequate to measure the diversity of Differential Evolution (DE). This consideration can be seen as a consequence of the No Free Lunch Theorem [940].

The first fitness diversity metric has been introduced in [104] and then used in [659]. This metric is given by:

$$\xi = \min\left\{\left|\frac{f_{best} - f_{avg}}{f_{best}}\right|, 1\right\}, \tag{10.6}$$

where $f_{best}$ and $f_{avg}$ are respectively best and average fitness values over the individuals of the population. Measurement $\xi$ can be seen as the answer to the question "How close is the average fitness to the best one?". Thus, if the average fitness value is as good as the best, the diversity is low and $\xi \approx 0$. On the contrary, if the fitness values are very distant the diversity metric is saturated to 1 and the diversity can be considered to be high. In this way, the metric $\xi$ can say whether the local search activation is suitable ($\xi \approx 0$) or unnecessary ($\xi = 1$). This metric proved to lead to a high algorithmic performance in some cases but suffers from robustness, as shown in [657]. The main drawback of this metric is that it is dependent on the codomain width: adding a constant value to the fitness function would lead to an important variation of the diversity values. However, this diversity metric is very efficient in the specific cases it has been used: for multivariate and complex fitness landscapes having a limited range of variability in the fitness values (e.g. $[0, 10]$) and the minimum around zero (e.g. for error minimization in engineering problems).

The second fitness diversity metric has been introduced in [888] and used also in [889]. The metric is:

$$\nu = \min\left\{1, \frac{\sigma_f}{\left|f_{avg}\right|}\right\}, \tag{10.7}$$

where $\left|f_{avg}\right|$ and $\sigma_f$ are respectively the average value and standard deviation over the fitness values of individuals of the population. Also the parameter $\nu$ can vary between 0 and 1 and can be seen as a measurement of the fitness diversity and distribution of the fitness values within the population. In other words, $\nu$ is the answer to the question "How sparse are the fitness values within the population?". As well as $\xi$, $\nu$ is codomain dependent and works with a limited range of variability. Unlike $\xi$, $\nu$ depends on the standard deviation and thus on the fitness distribution over all individuals of the population. In addition, $\nu$ is less sensitive than $\xi$ to fitness diversity variations and would not consider high diversity a situation where one individual has a performance significantly better than the others. For this feature if $\xi$ is efficient on an ES framework employing the plus strategy, $\nu$ can be employed for SIAs and DE i.e. for those algorithms which normally work in high diversity conditions, see [889].

The third fitness diversity metric has been introduced in [658] for a specific medical application. This metric consists of the following:

$$\psi = 1 - \left| \frac{f_{avg} - f_{best}}{f_{worst} - f_{best}} \right| \tag{10.8}$$

where $f_{best}$, $f_{avg}$ and $f_{worst}$ are respectively best, average and worst fitness over the individuals of the population. The parameter $\psi$ can be seen as the answer to the question "If we sort all fitness values over a line, which position is occupied by the average fitness?". The metric $\psi$ also varies between 0 and 1. It can be noticed that, unlike the two metrics previously presented, $\psi$ is not codomain dependant, i.e. its value does not depend on the range of variability of the fitness values. Due to its structure, this metric is very sensitive to small variations and thus is especially suitable for fitness landscapes containing plateaus and low gradient areas. Parameter $\psi$ has been successfully employed within memetic frameworks which employ plus strategy in the spirit of the ES.

In [106] the following parameter is used:

$$\chi = \frac{\left| f_{best} - f_{avg} \right|}{max \left| f_{best} - f_{avg} \right|_k} \tag{10.9}$$

where $f_{best}$ and $f_{avg}$ are the fitness values of, respectively, the best and average individuals of the population. $max \left| f_{best} - f_{avg} \right|_k$ is the maximum difference observed (e.g. at the $k^{th}$ generation), beginning from the start of the optimization process. It is clear that $\chi$ varies between 0 and 1; it scores 1 when the difference between the best and average fitness is the largest observed, and scores 0 when $f_{best} = f_{avg}$ i.e. the entire population is characterized by a unique fitness value. Thus, $\chi$ is the answer to the question "How much better is the best individual than the average fitness of the population with respect to the history of the optimization process?".

Besides considering it as a measurement of the fitness diversity, $\chi$ is an estimation of the best individual performance with respect to the other individuals. In other words, $\chi$ measures how much the super-fit outperforms the remaining part of the population. More specifically, the condition $\chi \approx 1$ means that one individual has a performance far above the average, thus one super-fit individual is leading the search. Conversely, the condition $\chi \approx 0$ means that performance of the individuals is comparable and there is not a super-fit. Due to its nature, $\chi$ is suitable for guessing the state of convergence in a population of a SIA or a DE. In [106], $\chi$ has been defined for coordinating the search components of a MA based on a DE framework. This choice was carried out by taking into account the fact that a DE structure works well when one individual is better than the others since it has the role of guiding the search. However, its performance should not be excessively good with respect to the others; otherwise, it would be unlikely for another individual to succeed at outperforming the leading individual. As a general guideline, a DE population containing a super-fit individual needs to exploit the direction offered by the super-fit in order to eventually generate a new individual that outperforms the super-fit. Conversely, a DE population made up of individuals with comparable fitness values requires that one individual that clearly outperforms the others be generated in order to have a good search lead. A similar analysis can be carried out for Particle Swarm Optimization (PSO) and other SIAs.

In [887] another fitness diversity metric has been introduced. This metric is given by:

$$\phi = \frac{\sigma_f}{|f_{worst} - f_{best}|} \tag{10.10}$$

where $\sigma_f$ is the standard deviation of fitness values over individuals of the populations, and $f_{worst}$ and $f_{best}$ are the worst and best fitness values, respectively, of the population individuals.

Analogous to the other fitness diversity indexes listed above, $\phi$ varies between 0 and 1. When the fitness diversity is high, $\phi \approx 1$; on the contrary when the fitness diversity is low, $\phi \approx 0$. The index $\phi$ can be seen as a combination of $\nu$ in formula (10.7) and $\psi$ in formula (10.8) because it represents the distribution of fitness values over individuals of the population with respect to its range of variability. In other words, $\phi$ is the answer to the question "How sparse are the fitness values with respect to the range of fitness variability at the current generation?". The index $\psi$ was also designed for DE based algorithms. Employment of the standard deviation in the numerator in formula (10.10) is due to the fact that a DE framework tends to generate an individual with performance significantly above the average (as mentioned for the metric $\chi$) and efficiently continues optimization for several generations. In this sense, an estimation of the fitness diversity of a DE population by means of the difference between best and average fitness values can return a misleading result and each value must be taken into account. Regarding the denominator in formula (10.10), a normalization to the range of variability of the current population makes the index co-domain invariant (unlike $\nu$ in formula (10.7) ) and its estimation is not affected, for example by adding an offset to the fitness function. Thus, the index $\phi$ can be successfully employed, within a DE framework, on problems of various kinds.

Finally, another fitness diversity index inspired also by the entropy study in [873] has been proposed in [481]. The population is sorted according to the fitness values. Thus an interval $[f_{min}, f_{max}]]$ having width $d$ can be detected. Let us indicate with $n_1$ the number of individuals falling within $\left[f_{min}, f_{min} + \frac{d}{3}\right]$ and with $n_3$ the number of individuals falling within $\left[f_{max} - \frac{d}{3}, f_{max}\right]$. Indicating with $N_p$ the number of individuals of the population and assuming that we want to solve a minimization problem, the diversity is then estimated as:

$$\tau_3 = 0.5 + \frac{n_1 - n_3}{2N_p}. \tag{10.11}$$

In other words, this metric subdivides the population into three quality classes and measures the diversity as a difference of the cardinality of the classes. Metric $\tau_3$ has been used for an ES framework but it might be suitable also for different frameworks. It has successfully been applied to a chemical engineering problem characterized by a highly multi-variate function but likely not a very multi-modal fitness landscape. It must be remarked that although $\tau_3$ also varies between 0 and 1, the interpretation of the parameter is different from the other diversity metrics. The maximum diversity condition occurs when $\tau_3 = 0.5$, which corresponds to maximum distribution of the performance over the individuals of the population. The conditions $\tau_3 \approx 0$ and $\tau_3 \approx 1$ mean that a few individuals have a very high performance with respect to the others and that a few individuals have a very low performance with respect to the others, respectively. In order to visualize this approach, it may be useful to imagine a ring where value 0 and 1 are contiguous. In this sense, this metric measures the balance among the three performance regions. This sophisticated way to measure diversity has the drawback that the metric can suffer from abrupt changes in proximity to 0 and 1 and very slow changes in proximity to 0.5, in correspondence of the same variations within the population. This can make the adaptation rather complicated to handle.

In order to summarize the features of the diversity metrics listed in this section, a synoptical scheme is shown in Table 10.1.

**Table 10.1.** Diversity Metrics: Synoptical Scheme

| Diversity Metric | Framework | Landscape Features | Drawbacks |
|---|---|---|---|
| $\xi$ | EAs | Highly Multi-modal Lanscape | Non scalable |
| $\nu$ | SIAs, DE | Flexible | Non scalable |
| $\psi$ | EAs | Plateaus, Flat Landscapes | Very sensitive |
| $\chi$ | SIAs, DE | Flexible | Very DE and PSO tailored |
| $\phi$ | SIAs, DE | Flexible | Very sensitive |
| $\tau_3$ | EAs | Large Scale not too Multi-modal | Abrupt and Slow Variations |

### 10.3.2 Coordination of the Search: The "Natura non Facit Saltus" Principle

At each generation, when a diversity metric is calculated the problem that follows is how to use such information in order to perform the coordination of global and local search. As mentioned before, let us consider that the MA employs an evolutionary framework and two local search algorithms, the first having explorative features, the second having exploitative features. The goal is to activate the explorative local search algorithm when the population has lost part of its diversity and to activate the exploitative local search algorithm when the population has lost most of its diversity and is approaching a convergence condition. In order to obtain this effect three adaptive schemes have been proposed in the literature.

The first scheme, used in [104], [659], and [658], employs a threshold mechanism for the application of local search. More specifically, when the control parameter surpasses a given threshold, the corresponding local search algorithm is activated. This mechanism can be seen as a probabilistic scheme where the probability of the local search activation, dependent upon the control parameter, is a step function which takes the value 1 within the threshold limits and 0 elsewhere. Although this kind of scheme has proven to be efficient for various applications (see e.g. [657]), the continuous variation of the fitness diversity in an evolutionary algorithm is not in accordance with this step function. In other words, if the fitness diversity metrics measure the necessity of the algorithm increasing/decreasing the local search within the memetic framework, the intensity of the local search is supposed to be related to the variation of the diversity metrics. On the contrary, a step function suggests that the local search is abruptly introduced within the search at its maximum intensity and can thus be too crude an approximation of the exploration/exploitation necessity of the MA.

In order to introduce smooth variation in the intensity of the local search application, two more schemes have been proposed in [106] and [889], respectively. More specifically, the step function has been replaced with a continuous function within the memetic frameworks under examination. Thus, the probability of local search activation is given by a function of the fitness diversity.

Indicating with $\lambda$ the fitness diversity metric, the first function is the beta distribution function, see [106]:

$$p(\lambda) = \frac{1}{B(s,t)} \cdot \frac{(\lambda - a)^{(s-1)} (b - \lambda)^{(t-1)}}{(b-a)^{(s+t-1)}} \tag{10.12}$$

where $a$ and $b$ are, respectively, the inferior and superior limits of the distribution; $B(s,t)$ is the beta function; $s = 2$ and $t = 2$ are the shape parameters. Parameters $a$ and $b$ must be set on the basis of the algorithm under consideration. The latter parameters play the same role as the thresholds in the previous scheme.

The second, used in [889], is the exponential distribution:

$$p(\lambda) = e^{\frac{-(\lambda-\mu_p)}{2\sigma_p^2}} \tag{10.13}$$

where $\mu_p$ and $\sigma_p$ are the parameters characterizing the intensity application range of the local search.

In order to better explain the three coordination scheme, let us consider the Fast Adaptive Memetic Algorithm (FAMA) proposed in [104]. This algorithm is based on an ES framework and two local search algorithms. The first local search, playing an explorative role, is the Nelder-Mead Algorithm (NMA) [653] and the second playing an exploitative role, is the Hooke-Jeeves Algorithm (HJA) [391]. For a proper functioning of FAMA, we desire that the NMA be activated when the diversity becomes low in order to give an alternative search logic, and that the HJA be activated in very low diversity condition. Since FAMA employs the $\xi$ metric, this statement can be rephrased as: the NMA is activated when $0.05 < \xi < 0.5$ and the HJA when $\xi < 0.2$. By keeping the same amount of local search, if the beta distribution function is employed then $a = 0$ and $b = 0.68$ for the NMA and, $a = 0$ and $b = 0.3$ for the HJA. Fig. 10.1 gives a graphical representation of the local search coordination, dependent on the diversity metrics, for the FAMA. The diagram shows the step functions (as in the original implementations) in the upper part, the related beta distribution functions in the central part, and the related exponential distributions in the lowest part.

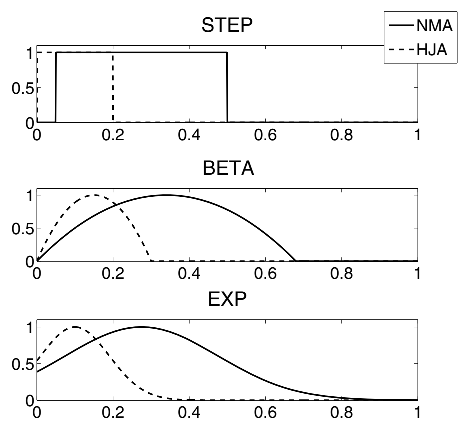


**Fig. 10.1.** Coordination of the local search for the FAMA

It should be remarked that the scaling of beta and exponential functions is done taking into account the fact that the areas below each trend are the same i.e. the overall balance between global and local search is the same for the original and proposed versions of each algorithm. For the sake of clarity, activation of a local searcher is performed by sampling (by means of a uniform distribution) a pseudo-random number $\varepsilon$ in $[0,1]$ and then comparing it with $p(\lambda)$; if $\varepsilon < p(\lambda)$ the corresponding local search is performed.

Numerical results reported in [890] show that the employment of continuous functions is beneficial and succeeds at improving upon the step scheme for a constant amount of local and global search. This fact has been expressed as the “natura non facit saltus” principle. The Latin expression “natura non facit saltus”, i.e. nature does not make (sudden) jumps, is a principle of classical physics, claimed since Aristoteles’ time until the formulation of the quantum mechanic theory, which states that variation of physical phenomena is continuous, thus not containing “jumps”. This concept has been extended to Memetic Computing and more specifically to the local search coordination, dependent on a fitness diversity index. The local search activation should not be abruptly started on the basis of some conditions but should slowly be increased and decreased around a suitable diversity condition.

## 10.4 Conclusion

This chapter analyzes the problem of diversity in Memetic Computing. The problem of diversity loss is very relevant in Evolutionary Computation since a premature diversity loss can lead to a premature algorithmic convergence into undesired areas of the decision space. Dually, some algorithms could fail at generating new genotypes despite a high diversity and thus stagnate. In Memetic Computing this problem is even more important because the local search application might cause the convergence to the same (or a very similar) point starting from a set of solutions belonging to the same basin of attraction. However, since Memetic Algorithms employ different search logics, if a proper coordination of the algorithmic components is carried out, a successful optimizer can be designed. Modern Memetic Algorithms use different local search algorithms for preserving a proper diversity which promotes the enhancements in the search, and they propose adaptive techniques for coordinating the various algorithmic components.

Several schemes for handling diversity have been illustrated. The employment of structured population has been widely used since it implicitly allows a preservation of diversity. However, distributed algorithms by themselves are not enough to prevent stagnation and premature convergence. Therefore, an adaptive system can support the memetic framework. A control mechanism based on Shannon’s entropy can be an efficient countermeasure. Fitness diversity adaptation also provides an efficient diversity control system since a diversity metric is used to coordinate the local search. Although this approach is promising it hides two problems: how to measure the diversity and how to use this information within a memetic framework. In accordance with the No Free Lunch Theorem, there is no optimal diversity metric, but rather its design should take into account the problem and the evolutionary/swarm intelligence structure under consideration. A synoptical table compares the metrics and gives some hints on how to use some diversity metrics proposed in the literature. Regarding the coordination of the algorithmic components, it has been observed that an efficient Memetic Algorithm should contain both explorative and exploitative local search algorithms. The explorative local search algorithm(s) assist the framework to detect novel promising search directions when the diversity is decreasing, while the exploitative one(s) perform an extensive search within already detected basins of attraction when the population has lost most of its diversity. To pursue this aim three control functions are illustrated in this chapter. The first function is a step function, i.e. local search is activated simply by means of threshold comparison. Although this approach is efficient, it has a wide margin of improvement if instead of a step function a continuous function is preferred. Two probability distribution functions have been considered. Previous studies observed that, while keeping constant the amount of local and global search, a Memetic Algorithm employing continuous functions outperforms on a regular basis the corresponding algorithm employing the step function. This fact was previously named the "natura non facit saltus" principle for Memetic Algorithms.

**Acknowledgements.** This work is supported by Academy of Finland, Akatemiatutkija 130600, Algorithmic Design Issues in Memetic Computing.