Classification of Adaptive Memetic Algorithms: A Comparative Study

Yew-Soon Ong, Meng-Hiot Lim, Ning Zhu, and Kok-Wai Wong

***Abstract*—Adaptation of parameters and operators represents one of the recent most important and promising areas of research in evolutionary computations; it is a form of designing self-configuring algorithms that acclimatize to suit the problem in hand. Here, our interests are on a recent breed of hybrid evolutionary algorithms typically known as adaptive memetic algorithms (MAs). One unique feature of adaptive MAs is the choice of local search methods or memes and recent studies have shown that this choice significantly affects the performances of problem searches. In this paper, we present a classification of memes adaptation in adaptive MAs on the basis of the mechanism used and the level of historical knowledge on the memes employed. Then the asymptotic convergence properties of the adaptive MAs considered are analyzed according to the classification. Subsequently, empirical studies on representatives of adaptive MAs for different type-level meme adaptations using continuous benchmark problems indicate that global-level adaptive MAs exhibit better search performances. Finally we conclude with some promising research directions in the area.**

***Index Terms*—Adaptation, evolutionary algorithm, memetic algorithm, optimization.**

I. Introduction

In problems characterized by many local optima, traditional local optimization techniques tend to fail in locating the global optimum. In these cases, modern stochastic techniques such as the genetic algorithm (GA) can be considered as an efficient and interesting option [1], [2]. As with most search and optimization techniques, the GA includes a number of operational parameters whose values significantly influence the behavior of the algorithm on a given problem, and usually in unpredictable ways. Often, one would need to tune the parameters of the GA to enhance its performance. Over the last 20 years, a great deal of research effort focused on adapting GA parameters automatically [3]–[5]. These include the mutation rate, crossover, and reproduction techniques where promising results have been demonstrated. Surveys and classifications of adaptations in evolutionary computation are available in Hinterding *et al.* [6] and Eiben *et al.* [7].

Nevertheless, traditional GAs generally suffer from excessively slow convergence to locate a precise enough solution because of their failure to exploit local information. This often limits the practicality of GAs on many large-scale real world problems where the computational time is a crucial consideration. Memetic algorithms (MAs) are population-based metaheuristic search approaches that have been receiving increasing attention in the recent years. They are inspired by Neo-Darwinian's principles of natural evolution and Dawkins' notion of a meme defined as a unit of cultural evolution that is capable of local refinements. Generally, MA may be regarded as a marriage between a population-based global search and local improvement procedures. It has shown to be successful and popular for solving optimization problems in many contexts [8]–[21]. Particularly, once the technique has been properly developed, higher quality solutions can be attained much more efficiently. Nevertheless, one drawback of the MA is that in order for it to be useful on a problem instance, one often needs to carry out extensive tuning of the control parameters, for example, the selection of a problem-specific meme that suit the problem of interest [9]. The influence of the memes employed has been shown extensively in [10]–[21] to have a major impact on the search performance of MAs. These studies demonstrated that the search performance obtained by MAs is often better than that obtained by the GA alone, especially when prior knowledge on suitable problem-specific memes is available.

In discrete combinatorial optimization research, Cowling *et al.* [11] coined the term "hyperheuristic" to describe the idea of fusing a number of different memes together, so that the actual meme applied may differ at each decision point, i.e., often at the chromosome/individual level. They describe the hyperheuristic as a heuristic to choose memes. The idea of using multimemes and adaptive choice of memes at each decision point was also proposed by Krasnogor *et al.* in [13], [14]. At about the same time, Smith also introduced the co-evolution of multiple memes in [15], [20]. These are some of the research groups that have been heavily involved in work relating to memes adaptation in MAs for combinatorial optimization problems. In the area of continuous optimization, Hart in his dissertation work [10] as well as some of our earlier research work in [16]–[18] have demonstrated that the choice of memes affects the performance of MAs significantly on a variety of benchmark problems of diverse properties. Ong and Keane [17] coined the term "meta-Lamarckian learning" to introduce the idea of adaptively choosing multiple memes during a MA search in the spirit of Lamarckian learning.

From our survey[1], it is noted there has been a lack of studies analyzing and comparing different adaptive MAs from the perspective of choosing memes. Our objective in this paper is to

Manuscript received August 31, 2004; revised February 21, 2005. This work was supported by Singapore Technologies Dynamics and Nanyang Technological University. This paper was recommended by Editor E. Santos.

Y.-S. Ong and N. Zhu are with the School of Computer Engineering, Nanyang Technological University, Singapore 639798 (e-mail: asysong@ntu.edu.sg; pg04169798@ntu.edu.sg).

M.-H. Lim is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: emhlim@ntu.edu.sg).

K.-W. Wong is with the School of Information Technology, Murdoch University, Perth, WA 6150, Australia (e-mail: k.wong@murdoch.edu.au).

Digital Object Identifier 10.1109/TSMCB.2005.856143

[1]It is worth noting here that our focus is the class of adaptive MAs where the choice of memes is adapted during the evolutionary search.

```
Procedure::Canonical_MA
Begin
Initialize: Generate an initial GA population.
  While (Stopping conditions are not satisfied)
        Evaluate all individuals in the population
        For each individual in the population
            • Proceed with local improvement and replace the genotype and/or phenotype in
              the population with the improved solution depending on Lamarckian or
              Baldwinian Learning.
        End For
        Apply standard GA operators to create a new population; i.e., crossover, mutation and
        selection.
  End While
End
```

Fig. 1. Canonical MA pseudocode.

summarize the state-of-art in the adaptation of choice of memes in general nonlinear optimization. In particular, we conduct a classification on adapting the choice of memes in MAs based on the mechanism and the level of historical knowledge used. It is worth noting that such a classification would be informative to the evolutionary computation community since researchers use the terms "meta-Lamarckian learning", "hyperheuristic", and "multimemes" arbitrarily when referring to memes adaptation in adaptive MAs. Based on the resultant classification, we conduct a systematic study on the adaptive MAs for different type-level adaptations in the taxonomy using continuous benchmark problems of diverse properties. Last but not least, we hope that the classification and results presented here will help promote greater research in adaptive MAs and assist in identifying new research directions.

This paper is organized in the following manner. Section II presents the recent development of adaptive MAs in general nonlinear optimization. The proposed taxonomy or classification for adaptive MAs is then presented in Section III. Section IV presents the analyses on global convergence properties of adaptive MAs by means of Finite Markov chain. Section V summarizes the empirical studies on the assortment of adaptations using a variety of continuous parametric benchmark test functions. Finally, Section VI concludes this paper with some future research directions.

II. Adaptive MAs for General Nonlinear Optimization

Optimization theory [22] is the study of the mathematical properties of optimization problems and the analysis of algorithms for their solutions. It deals with the problem of minimizing or maximizing a mathematical model of an objective function such as cost, fuel consumption, etc., subject to a set of constraints. In particular, we consider the general nonlinear programming problem of the form:

- a target objective function $f(x)$ to be minimized or maximized;
- a set of variables, x, $x_{\mathrm{low}} \leq x \leq x_{\mathrm{up}}$, which affect the value of the objective function, where x_{low} and x_{up} are the lower and upper bounds, respectively;
- a set of equality/inequality constraints $g_w(x)$ that allows the unknowns to take on certain values but exclude others. For example, the constraints may take the form of $g_w(x) \leq 0$, for $w = 1, \ldots, b$, where b is the number of constraints.

A. Memetic Algorithms (MAs)

A GA is a computational model that mimics the biological evolution, whereas a MA, in contrast mimics culture evolution [23]. It can be thought of as units of information that are replicated while people exchange ideas. In a MA, a population consists solely of local optimum solutions. The basic steps of a canonical MA for general nonlinear optimization based on the GA can be outlined in Fig. 1.

B. Adaptive MAs

One unique feature of the adaptive MAs we consider here is the use of multiple memes in the search and the decision on which meme to apply on an individual is made dynamically. This form of adaptive MAs promotes both cooperation and competition among various problem-specific memes and favors neighborhood structures containing high quality solutions that may be arrived at low computational efforts. The adaptive MAs can be outlined in Fig. 2. In the first step, the GA population may be initialized either randomly or using design of experiments technique such as Latin hypercube sampling [24]. Subsequently, for each individual in the population, a meme is selected from a pool of memes considered in the search to conduct the local improvements. Different strategies may be employed to facilitate the decision making process [11]–[20]. For example, one may reward a meme based on its ability to perform local improvement and use this as a metric in the selection process [11], [12], [16]–[19]. After local improvement, the genotypes and/or phenotypes in the original population are replaced with the improved solution depending on the learning mechanism, i.e., Lamarckian or Baldwinian learning. Standard GA operators are then used to form the next population.

1) Hyperheuristic Adaptive MAs: In the context of combinatorial optimization, Cowling *et al.* [11] introduced the term "hyperheuristic" as a strategy that manages the choice of which meme should be applied at any given time, depending upon the characteristics of the memes and the region of the solution space currently under exploration. With hyperheuristic, multiple memes were considered in the evolution search. In their work, three different categories of hyperheuristics have been demonstrated for scheduling problems, namely: 1) random; 2) greedy; and 3) choice-function [11], [12], [19].

Under the random category [11], the first is *Simplerandom*. Here, a meme is selected randomly at each decision point. It is

Procedure::Memes_adaptation_in_adaptive_MA
Begin
Initialize: Generate an initial GA population
While (Stopping conditions are not satisfied)
Evaluate all individuals in the population
For each individual in the population
- *Select suitable memes.*
- Proceed with local improvement and replace the genotype and/or phenotype in the population with the improved solution based on Lamarckian or Baldwinian learning.

End For
Apply standard GA operators to create a new population; i.e., crossover, mutation and selection.
End While
End

Fig. 2. General framework of memes adaptation in adaptive MAs.

purely stochastic in nature, and the probability of choosing each meme is kept constant throughout the search. This strategy may be regarded as a datum with which other selection strategies may be compared. In *Randomdescent*, initially the choice of memes is decided randomly. Subsequently, this same meme is used repeatedly until no further local improvements can be found. This same process then repeats to consider all the other memes. *Randompermdescent* is similar to the *Randomdescent* strategy except that a random permutation of memes $M_1, M_2, \ldots, M_\eta$ is fixed in advance, and when the application of a meme does not result in any improvement, the next meme in the permutation is used.

Unlike the random category, the greedy category [11] resembles a brute force technique that experiments with every meme on each individual and chooses the meme that result in the biggest improvement. Since it is a brute-force method, the drawback of greedy hyperheuristic is clearly the high computational cost.

In their choice-function category [11], [12], choice function incorporating multiple metrics of goodness is used to assess how effective a meme is, based upon the current state of knowledge about the region of the solution space under exploration. The choice function F proposed in [11], [12] is composed of three components. The first component represented by f_1 reflects the recent improvement made by a meme and expresses the idea that if a meme has recently performed well, it is likely to continue to be effective. The second component f_2 describes the improvement contributed by the consecutive pairs of memes, and the last component f_3 records the period elapsed since a meme was last used. Five strategies were introduced for the hyperheuristics, choice-function category [11], [12]. In the *Straightchoice* strategy, the meme that yields the best F value is chosen at each decision point. In the second strategy, *Rankedchoice*, memes are ranked according to F, and the top ranking memes are experimented individually, and only the meme that yields the largest improvement proceeds with Lamarckian learning. In the *Roulettechoice* strategy, a meme M_e is chosen with probability relative to the overall improvement, i.e., $(F(M_e)/\sum_{e=1}^{\eta} F(M_e))$, η is the total memes considered. The *Decompchoice* strategy considers each component in F, i.e., f_1, f_2 and f_3, individually. In particular, the strategy experiments with each of the meme and records the best local improvement based on f_1, f_2, f_3 and F individually. Subsequently, the meme that results in the best improvement among those identified is used. This implies that up to four memes will be individually tested in the case when all the highest ranked performing memes are different for f_1, f_2, f_3 and F. Alternatively, using the choice function, a tabu-list created may also be used to narrow down the choice of memes at each decision point [19]. This is labeled here as the *Tabu-search* strategy.

2) Multimemes and Co-Evolving MAs: Krasnogor also proposed a simple inheritance mechanism for discrete combinatorial search [13], [14]. Each individual is represented and composed by its genetic material and memetic material. The memetic material encoded into its genetic part specifies the meme that will be used to perform local search in the neighborhood of the solution. Smith also worked on co-evolving memetic algorithms that use similar mechanisms to govern the choice of memes represented in the form of rules [15], [20]. These are forms of self-adaptive MA that evolves simultaneously the genetic material and the choice of memes during the search. A simple vertical inheritance mechanism, as used in general self-adaptive GAs and evolutionary strategies is shown to provide a robust adaptation of behavior [13], [14]. The multimemes algorithm with simple inheritance mechanism is outlined in Fig. 3.

3) Meta-Lamarckian Learning: In continuous nonlinear function optimization, Ong and Keane [17] studied the meta-Lamarckian learning on a range of benchmark problems of diverse properties. Since the study on using multiple memes in a MA search concentrated on Lamarckian learning, it was termed as meta-Lamarckian learning [17]. The main motivation of the work was to facilitate competition and cooperation among the multiple memes employed in the memetic search so as to solve a problem with greater effectiveness and efficiency. A *Basic meta-Lamarckian learning* strategy was proposed as the baseline algorithm that forms a datum that other meta-Lamarckian learning strategies may be compared. This is similar to the *Simplerandom* proposed in hyperheuristic where no adaptation has been used. It has the advantage of at least giving all the available memes being considered a chance to improve each chromosome throughout the MA search.

Further, two adaptive strategies were investigated in [16], [17]. The rationale behind the *Sub-problem Decomposition* strategy was to decompose the original search problem cost

Procedure::Simple_Inheritance_Mechanism
Begin
Replace the meme of individual with a randomly selected meme according to the specified innovation rate.
Select 2 parent chromosomes:
    If (both parents have the same meme)
        Pass common meme to the offspring.
    Else-if (parents1.fitness == parents2.fitness)
        Randomly select one of the two attached memes to the offspring.
    Else-if (parents1.fitness > parents2.fitness)
        Pass parent1 meme to the offspring.
    Else /* (parents2.fitness < parents1.fitness) */
        Pass parent2 meme to the offspring.
End

Fig. 3. Simple inheritance mechanism pseudocode.

Procedure::Sub-problem_Decomposition
Begin
If (GA Generation < g)

- Generate a random number between 1 and meme pool size.
- Select the meme corresponding to the number.
- Create/update database.

Else

- Locate k chromosomes nearest to $\hat{p}$ in database P using simple Euclidean measures, i.e., $P_k \Rightarrow \left\{ k \min_{p_j \in P} \left\| \hat{p} - p_j \right\| \right\}$.
- Find the average fitness of each member of the reduced meme pool based on P_k.
- Select the meme with the maximum average fitness.
- Update database.

End

Fig. 4. Outline of subproblem decomposition strategy.

Procedure::Biased_Roulette_Wheel
Begin
If (Training Stage)

- Ensure each meme is given one chance to participate in a random order.
- Update meme's global fitness.

Else

- Sum the fitness of each member of the meme pool.
- Determine the normalized relative fitness of each member of the meme pool.
- Assign space on roulette wheel proportional to meme's fitness.
- Generate a random number between 0 and 1 and select the corresponding meme.
- Update meme's global fitness.

End

Fig. 5. Outline of *Biased Roulette Wheel* strategy.

surface, which is often large and complex, into many sub-partitions dynamically, and attempts to choose the most competitive meme for each sub-partition. To choose a suitable meme at each decision point, the strategy gathers knowledge about the ability of the memes to search on a particular region of the search space from a database of past experiences archived during the initial EA search. The memes identified then form the candidate memes that will compete, based on their rewards, to decide on which meme will proceed with the local improvement. In this manner, it was shown that the strategy proposed creates opportunities for joint operations between different memes in solving the problem as a whole, because the diverse memes help to improve the overall population based on their areas of specialization. Hence, *Sub-problem Decomposition* promotes both cooperation and competition among the memes in the memetic search. On the other hand, the *Biased Roulette Wheel* strategy [17] is similar to the *Roulettechoice* [11]. Nevertheless, they do differ in the choice functions used. The pseudo-codes for these two strategies are outlined in Figs. 4 and 5.

III. Classification of Adaptive MAs

In this section, a classification of memes adaptation in adaptive MAs is presented. This classification is based on the mechanism of adaptation (adaptation type) [6], [7] and on which level the historical knowledge of the memes is used (adaptation level) in adapting the choice of memes in adaptive MAs. It is orthogonal and encompasses diverse forms of memes adaptation in adaptive MAs. The taxonomy on existing strategies of adaptive MAs based on adaptation type and level is depicted in Table I.

A. Adaptation Type

The classification of the adaptation type is made on the basis of the mechanism of the adaptation used in the choice of memes.

TABLE I
A CLASSIFICATION OF MEMES ADAPTATION IN ADAPTIVE MAs

Adaptive Type		Adaptive Level: External	Adaptive Level: Local	Adaptive Level: Global
Static		Basic meta-Lamarckian learning / Simplerandom		
Adaptive	*Qualitative Adaptation*		Randomdescent / Randompermdescent	Tabu-search
Adaptive	*Quantitative Adaptation*		Sub-Problem Decomposition/ Greedy	Straightchoice/ Rankedchoice/ Roulettechoice/ Decompchoice/ Biased Roulette Wheel
Self-Adaptive			Multi-memes/ Co-evolution MA	

Shaded region indicates that the type/level adaptive MAs do not exist in our classification.

In particular, attention is paid to the issue of whether or not a feedback from the adaptive MAs is used and how it is used. Here, feedback is defined as the improvement attained by the chosen meme on the chromosome searched.

1) Static: When no form of feedback is used during the evolutionary search, it is considered as a static type adaptation. The *Basic meta-Lamarckian learning* strategy for meme selection [17], or *Simplerandom* strategy [11], are simple random walk over the available memes every time a chromosome is to be locally improved. Since it does not make use of any feedback from the search, it is a form of static adaptation strategy in adaptive MAs.

2) Adaptive: Adaptive dynamic adaptation takes place when feedback from the MA search influences the choice of memes at each decision point. Here, we divided adaptive dynamic adaptation into qualitative or quantitative adaptation.

In qualitative adaptation, the exact value of the feedback is of little importance. Instead, the quality of a meme is sufficient. As long as the present meme generates improvement in the local learning process, it remains employed in the next decision point. Otherwise, a new meme is chosen and the process repeats until the stopping criteria are met. The *Randomdescent, Randompermdescent* and *Tabu-search* strategies [11], [19] are forms of qualitative adaptation.

On the other hand, the *Greedy, Straightchoice, Rankedchoice, Roulettechoice, Decompchoice, Biased Roulette Wheel*, and *Sub-Problem Decomposition* strategies [11], [17] rely on the quantitative value of the feedback obtained on each individual's culture evolution to decide on the choice of memes. They are thus considered as forms of quantitative adaptation.

3) Self-Adaptive: Self-adaptive type adaptation employs the idea of evolution to implement the self-adaptation of memes. Naturally, both *multimemes* [13], [14] and *co-evolution MAs*[15], [20] are forms of self-adaptive adaptation as the memetic representation of the memes is coded as part of the individual and undergoes standard evolution.

B. Adaptation Level

The adaptation level refers to the level of historical knowledge of the memes that are employed in the choice of memes. Here, the level of historical knowledge means the extent of past knowledge about the memes. The adaptive level is further divided into external, local, and global.

1) External: External-level adaptation refers to the case where no online knowledge about the memes is involved in the choice of memes. In many real world applications, the pool of problem-specific memes is usually selected from past experiences of human experts. This is a formalization of the knowledge that domain experts possess about the behavior of the memes and the optimization problem in general. *Basic meta-Lamarckian learning* or *Simplerandom* strategies [11], [17] are classified as forms of external adaptive level since they make use of external knowledge from past experiences.

2) Local: In local-level adaptation, the decision making process on choice of memes involves simply on parts of the historical knowledge. The *Greedy, Randomdescent*, and *Randompermdescent* strategies [11] make decisions based on the improvement obtained in the present or immediate preceding culture evolution; hence, they are strategies categorized under local adaptive level. Nevertheless, it is worth noting that global-level adaptation may be easily derived when one considers all previously searched chromosomes.

The *Sub-Problem Decomposition* strategy, *multimemes*, and *Co-evolution MA*[13], [17] selects a meme based on the knowledge gained from only the k nearest individuals or parents among all that were searched previously. Hence, they are also considered as strategies that practice local-level adaptations.

3) Global: Global-level adaptation takes place if the complete historical knowledge is used to decide on the choice of memes. *Straightchoice, Rankedchoice, Roulettechoice,* and *Decompchoice*, *Biased Roulette Wheel,* and *Tabu-search* [11], [17], [19] strategies are classified as forms of adaptations at the global adaptive level since they make complete use of historical knowledge on the memes when deciding which memes to opt for.

IV. CONVERGENCE ANALYSIS OF ADAPTIVE MAs

In this section, we analyze the global convergence properties of adaptive MAs according to their level of adaptations described in Section III, i.e., External, Local or Global, using the theory of Markov chain and extending from previous efforts on convergence analysis of genetic algorithms [25]–[29].

A. Finite Markov Chain

The Markov chain is a popular theory among the EA community as it offers an appropriate framework for analyzing discrete-time stochastic process. To begin, we outline some basic definitions of Markov chain that are used in the analysis on global convergence properties of adaptive MAs [30].

Definition 1: If the transition probability $p_{ij}(\cdot)$ is independent of time, i.e., $p_{ij}(t_1) = p_{ij}(t_2)$ for all $i, j \in S$ and $t_1, t_2 \in \{0, 1, 2, \ldots\}$, the Markov chain is said to be homogeneous over the finite state space S.

Definition 2: A Markov chain is called irreducible if for all pairs of states $i, j \in S$ there exists an integer t such that $p_{ij}^t > 0$. finite irreducible chains are always recurrent.

Definition 3: An irreducible chain is called aperiodic if for all pairs of states $i, j \in S$ there is an integer t_{ij} such that for all $t > t_{ij}$, the probability $p_{ij}^t > 0$.

B. Markov Chain Analysis of Adaptive MAs

To model an adaptive MA, we define the states of a Markov chain. Let S be the collection of length l binary strings. Hence the number of possible strings r is 2^l. If n is the size of the population pool, it is possible to show that N, the total number of population pools or the number of Markov states, can be defined by $N = \begin{bmatrix} n+r-1 \\ r-1 \end{bmatrix}$ in the adaptive MA.

Further, we model the adaptive MA as a discrete-time Markov chain $\{X^{(t)}\}$ $(t = 0, 1, 2, \ldots)$, with a finite state space $S = \{S_1, S_2, \ldots, S_N,\}$, and tth step transition matrix $\mathbf{P}(t)$ where

$$p_{ij}^{(t)} = \Pr\left\{X^{(t)} = S_j \middle| X^{(t-1)} = S_i\right\}, \quad i,j = 1, \ldots N \quad (1)$$

and initial probability distribution

$$v_i^{(0)} = \Pr\left\{X^{(0)} = S_i\right\},$$
$$v_i^{(0)} \geq 0, \quad \sum_{i=1}^{N} v_i^{(0)} = 1 \quad i = 1, 2, \ldots, N. \quad (2)$$

The probabilistic changes in the adaptive MA population due to the evolutionary operators may be modeled using stochastic matrices $\mathbf{P}_c$, $\mathbf{P}_m$ and $\mathbf{P}_s$ representing the crossover, mutation, and selection operations, respectively. Besides the standard evolutionary operators, the adaptive MA will also refine each individual using different memes at each decision point. We model this process of local improvement as a $N \times N$ transition matrix $\mathbf{P}_l$, which represents the state transition matrix of each individual that undergoes Lamarckian learning. This indicates that $\mathbf{P}_l$ has at least one positive entry in each row.

On the whole, the process of adaptive MAs is then modeled as a single transition matrix $\mathbf{P}$ of size $N \times N$ given by

$$\mathbf{P} = \mathbf{P}_l \mathbf{P}_c \mathbf{P}_m \mathbf{P}_s \quad (3)$$

where transition matrix $\mathbf{P}$ incorporates all adaptive MA operators which includes crossover, mutation, fitness-proportional selection and the Lamarckian learning mechanisms. In (3), transition matrices $\mathbf{P}_c, \mathbf{P}_m, \mathbf{P}_s$ are independent of time. In contrast, $\mathbf{P}_l$ may be dependent or independent of time, depending on the adaptive MA strategies considered.

We begin with local-level adaptive MAs, particularly the *Randomdescent, Randompermdescent,* and *Sub-Problem Decomposition* strategies. Since the choice of meme is made based on the immediate preceding or neighboring culture evolution, $\mathbf{P}_l$ is dependent of time and varies for each decision point. Hence, this form of adaptive MAs may not possess any global convergence guarantee.

On the other hand, since the choice of meme in the *Basic meta-Lamarckian* strategy is randowm while the *Greedy* strategy judges all options of memes experimentally and always selects the best among them, the transition matrix $\mathbf{P}_l$ is clearly time homogeneous or independent of time. In addition, strategies considered in the global-level adaptive MAs also converge with a probability of one [28]. When $t \to \infty$, it is feasible to assume the probability that the most suitable meme is selected tends to 1, i.e., $\lim_{t\to\infty} \mathbf{P}_l^t$ converges to a constant matrix.

Theorem 1: This kind of adaptive MAs is irreducible and aperiodic.

Proof: We know that $\mathbf{P}_c\mathbf{P}_m\mathbf{P}_s$ is positive as shown in [26]. Further, from the properties of $\mathbf{P}_l$, it can be easily shown that $\mathbf{P}_l\mathbf{P}_c\mathbf{P}_m\mathbf{P}_s$ is positive, i.e., since $\mathbf{P}_l$ has at least one positive entry in each row and $\mathbf{A} = \mathbf{P}_c\mathbf{P}_m\mathbf{P}_s$ is positive, $\mathbf{P} = \mathbf{P}_l\mathbf{A}$ is also strictly positive. Hence it is irreducible and aperiodic.

Theorem 2: There are only positive recurrent states in this kind of adaptive MAs.

Proof: The entire state space is a closed (ergodic) set because the Markov chain is irreducible. Hence, the Markov chain must be composed of only positive recurrent states.

Theorem 3: This kind of adaptive MAs possesses the property of global convergence.

Proof: Suppose $S_{\text{opt}} \subset S$, S_{opt} being the set of states containing the global optima. Because it is irreducible, aperiodic and positive recurrent, as $t \to \infty$ the probability that all the points in the search space will be visited at least once, approaches 1. Let i^* refers to the fittest individual in the evolutionary search process. So, whatever the initial distribution is, it must happen such that $\lim_{t\to\infty} \Pr\{i^* \in S_{\text{opt}}\} = 1$. Further, since the fittest solution is always tracked in practice, it extends from [26] that the search converges globally using an *elitist selection* mechanism.

V. Empirical Study on Benchmarking Problems

In this section, we present an empirical study on the various adaptive MA strategies. In particular, the representative strategies from each category of adaptive MAs as depicted in Table I, are compared with the canonical MAs. These adaptive MA strategies include the following.

1) External-Static: Basic meta-Lamarckian learning.
2) Local-Qualitative: Randomdescent, Randompermdescent.
3) Global-Qualitative: Tabu-search.
4) Global-Quantitative: Straightchoice, Roulettechoice.
5) Local-Quantitative: Sub-Problem Decomposition.
6) Local-Self-adaptive: multimemes.

For the sake of brevity, these strategies are abbreviated in this section as S-E, QL1-L, QL2-L, QL3-G, QN1-G, QN2-G, QN3-L, S-L, respectively. Further, considering that most existing efforts on adaptive MAs have been on combinatorial optimization problems, the emphasis here is placed on continuous benchmark optimization problems.

A. Benchmark Problems for Function Optimization

Five commonly used continuous benchmark test functions are employed in this study [17], [31], [32]. They have diverse properties in term of epistasis, multimodality, discontinuity, and constraint, as summarized in Table II.

B. Memes for Function Optimization

Various memes from the OPTIONS optimization package [33] were employed in the empirical studies. They consist of a variety of optimization methods from the Schwefel libraries [34] and a few others in the literature [35]–[40]. The eight memes used here are representatives of second, first, and

TABLE II
CLASSES OF BENCHMARK TEST FUNCTIONS CONSIDERED. *1: EPISTASIS, *2: MULTIMODALITY, *3: DISCONTINUITY, *4: CONSTRAINT

Benchmark Test Functions	Range of x_a	Characteristics				Global Optimum
		Epi*1	Mul*2	Disc*3	Con*4	
$F_{Bump} = \frac{abs\left[\sum_{a=1}^{d}\cos^4(x_a) - 2\prod_{a=1}^{d}\cos^2(x_a)\right]}{\sqrt{\sum_{a=1}^{d} a x_a^2}}$ $\prod_{a=1}^{d} x_a > 0.75$ **and** $\sum_{a=1}^{d} x_a < 15d/2$	$[0, 10]^{20}$	high	high	none	yes	~ 0.81
$F_{Griewank} = 1 + \sum_{a=1}^{d}\frac{x_a^2}{4000} - \prod_{a=1}^{d}\cos(x_a/\sqrt{a})$	$[-600, 600]^{10}$	weak	high	none	no	0.0
$F_{Rastrigin} = 20d + \sum_{a=1}^{d}\left(x_a^2 - 20\times\cos(2\pi x_a)\right)$	$[-5.12, 5.12]^{20}$	none	high	none	no	0.0
$F_{Sphere} = \sum_{a=1}^{d} x_a^2$	$[-5.12, 5.12]^{30}$	none	none	none	no	0.0
$F_{Step} = 6d + \sum_{a=1}^{d}\lfloor x_a \rfloor$	$[-5.12, 5.12]^{5}$	none	weak	high	no	0.0

TABLE III
LIST OF MEMES OR LOCAL SEARCH METHODS CONSIDERED

Abbreviations	**Memes or Local Search Methods**
BL	Bit climbing algorithm [35]
DP	Davis, Swan and Campey with Palmer orthogonalizational by Schwefel [34,36]
FB	Schwefel library Fibonacci search [34].
FL	Fletcher's 1972 method by Siddall [37]
GL	Repeated one-dimensional Golden section search by Schwefel [34]
SX	Powell's strategy of conjugate directions by Schwefel [38]
PS	A direct search using the conjugate direction approach with quadratic convergence [39].
SK	A series of exploratory moves that consider the behavior of the objective function at a pattern of points, all of which lie on a rational lattice [40].

zeroth-order local search methods and are listed in Table III together with their respective abbreviations used later in the paper.

C. Choice Function

The choice function employed in this study is based on that proposed in [11], which appears to be one of the most sophisticated that exists. It is used to select a meme M_g and is defined by the effective choice function $F(\cdot)$ given by (4) and (5), shown at the bottom of the next page. Equation (4) records for each meme the feedback f_1 on the effectiveness of a meme, feedback in regard to the effectiveness on consecutive pairs of (M_e, M_g) memes represented by f_2, and f_3 to facilitate diversity in memes selection (see [11] for details on the formulations of f_1, f_2 and f_3). In our empirical studies, the parameters of (4) and (5) are configured as follows: $\alpha = 0.9$, $\beta = 0.1$, $\delta = 1.5$, $\varepsilon = 1$ and $\rho = 1.5$, which corresponds to the suggested values in the literature [17], [18], [22]. Since the number of memes employed totals to eight, hence $\eta = 8$.

D. Results for Benchmark Test Problems

To see how the choice of the memes affects the performance and efficiency of the search, the eight different memes used to form the canonical MAs were employed on the benchmark problems. All results presented are averages of ten independent runs. Each run continues until the global optimum was found or a maximum of 40 000 trials (function evaluation calls) was reached, except for the Bump function where a maximum of up to 100 000 trials was used. In each run, the control parameters used in solving the benchmark problems were set as follows: population size of 50, mutation rate of 0.1%, 2-point crossover, 10-bit binary encoding, maximum local search length of 100

$$F(M_g) = \max\left\{-\alpha f_1(M_g) - \beta f_2(M_e, M_g) + \delta f_3(M_g), Q\rho^{\alpha f_1(M_g) + \beta f_2 e,(M_g) - \delta f_3(M_g)}\right\} \quad (4)$$

$$Q = \frac{\sum_g \max\{0, -\alpha f_1(M_g) - \beta f_2(M_e, M_g) + \delta f_3(M_g) + \varepsilon\}}{10\eta} \quad (5)$$

TABLE IV
RESULTS FOR BENCHMARK TEST PROBLEM

Level-Type		Bump Function (Maximum)		Griewank Function (Minimum)		Rastrigin Function (Minimum)		Sphere Function (Minimum)		Step Function (Minimum)	
		Mean at 100,000	Rank	Mean at 40,000	Rank	Mean at 40,000	Rank	Eval. Count when Global Optimum is found	Rank	Eval. Count when Global Optimum is found	Rank
External-Static	S-E	0.5641	9	0.005250	7	16.876	8	12593	9	23433	9
Local - Qualitative	QL1-L	0.6867	7	0.525366	12	84.97718	13	> 40000	11	19504	8
	QL2-L	0.6840	8	0.010610	9	18.62152	6	8599	3	8942	4
Global-Qualitative	QL3-G	**0.7444**	**1**	0.000450	2	18.05298	5	8599	3	**8056**	**1**
Global-Quantitative	QN1-G	0.7358	3	**0.000062**	**1**	**9.607814**	**1**	**8193**	**2**	9653	5
	QN2-G	0.7160	5	0.006106	8	14.52411	4	9196	6	14329	7
Local-Quantitative	QN3-L	0.7378	2	0.000558	4	33.49291	7	10194	7	12007	6
Local-Self-adaptive	S-L	0.6985	6	0.002863	5	14.16887	2	11792	8	28100	10
Canonical MAs	GA- BL	0.5275	12	0.6137	13	92.334	14	> 40000	14	8588	2
	GA- DP	0.7278	4	0.000516	3	14.448	3	9098	5	8931	3
	GA- FB	0.5415	11	19.096	15	144.25	15	> 40000	16	25706	11
	GA- FL	0.5183	13	0.00707	10	69.863	9	6666	1	> 40000	14
	GA- GL	0.5494	10	22.646	16	155.11	16	> 40000	15	25706	11
	GA- PS	0.4990	18	0.003378	6	74.106	11	12292	10	> 40000	15
	GA- SK	0.5062	14	0.33862	11	81.118	12	40000	13	> 40000	16
	GA- SX	0.3642	16	0.7861	14	73.79	10	> 40000	12	> 40000	13

evaluations and the probability of applying local search on a parent chromosome is set to unity.

The results obtained from our studies on the benchmark test problems are presented in Table IV. In the case where an algorithm manages to locate the global optimum of a benchmark problem, the number of evaluation count presented indicates the effort taken to reach this optimum solution. Otherwise, the best fitness averaged over ten runs is presented. Further, the canonical MAs and adaptive MAs are ranked according to their ability to produce high-quality solutions on the benchmark problems under the specified computational budget. The search traces of the best performing canonical MA and adaptive MA, together with the worst performing canonical MA on each benchmark function are also revealed in Figs. 6–10. Note that in all the figures, results are plotted against the total number of function calls made by the combined genetic and local searches. These numerical results obtained are analyzed according to the following aspects.

- *Robustness and Search Quality*—the capability of the strategy to generate search performances that are competitive or superior to the best canonical MA (from among the pool considered), on different problems and the capability of the strategy to provide high quality solutions.
- *Computational Cost*—the computational effort needed by the different adaptive MAs.

1) Robustness and Search Quality: From the results (see Table IV and Figs. 6–10), it is worth noting that the ranking of the adaptive MA strategies are relatively higher than most canonical MAs for the majority of the benchmark problems, implying that adaptive MAs generally outperform canonical MAs for these benchmark problems. This demonstrates the effectiveness of adaptive MAs in providing robust search performances, suggesting the basis for the increasing research interests in developing new adaptive MA strategies.

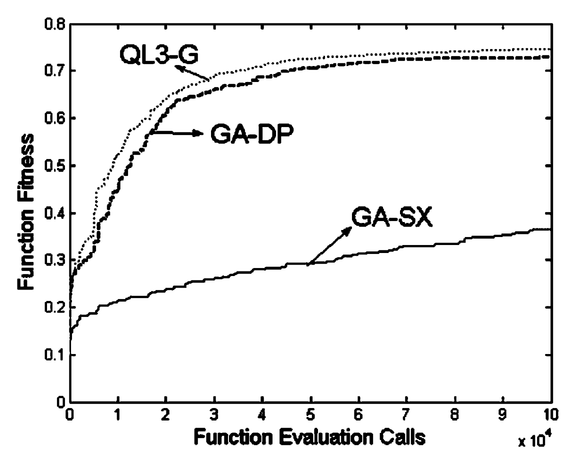


Fig. 6. Search traces for maximizing 20-dimensional Bump function.

From the search traces, it is worth noting that the adaptive strategies gradually "learn" the characteristics of the problem, i.e., the strengths and weaknesses of the different memes to tackle the problem in hand, suggesting why it is inferior to some canonical MAs during the early stages of the search. This is referred to as the learning stage in [17]. After this initial training stage, most of the best adaptive MAs were observed to outperform even the best canonical MAs on each benchmark function (see Figs. 6–8). The only exception is on the Sphere test function. While any adaptive MA takes effort to learn about

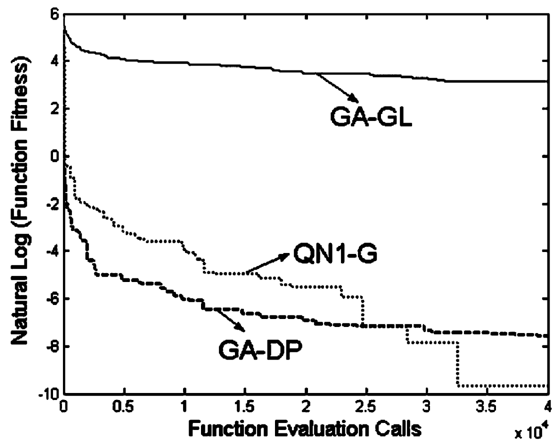


Fig. 7. Search traces for minimizing 10-dimensional Griewank function.

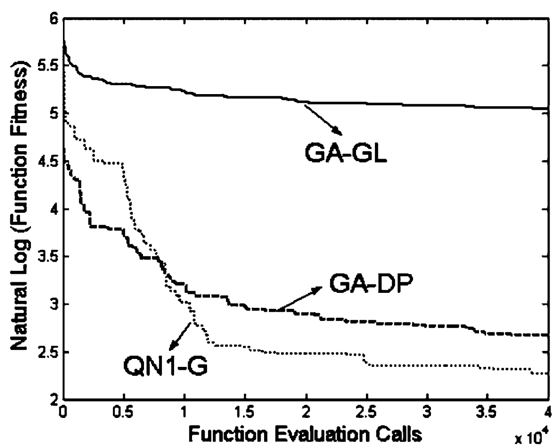


Fig. 8. Search traces for minimizing 20-dimensional Rastrigin function.

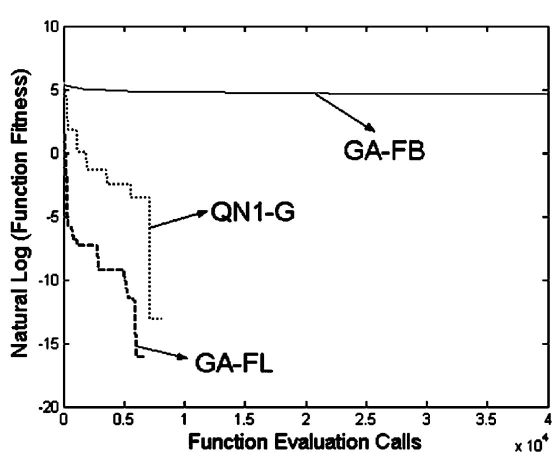


Fig. 9. Search traces for minimizing 30-dimensional Sphere function.

the different memes employed, the canonical MA using FL had already converged to the optimum of the unimodal quadratic sphere function since as a quasi-Newton method, it is capable of locating the global/local optimum of the unimodal quadratic Sphere function rapidly. The reader is referred to [17] for greater details on such observations where it was demonstrated that an increase in the size of memes employed in the adaptive MAs results in greater learning time.

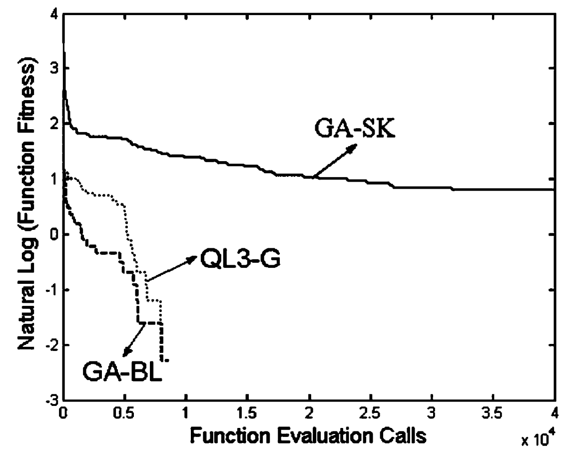


Fig. 10. Search traces for minimizing 5-dimensional Step function.

TABLE V
NORMALIZED COMPUTATIONAL COST RELATIVE TO S-E ON BENCHMARK PROBLEMS

	Bump Function	Rastrigin Function	Griewank Function	Sphere Function	Step Function
S-E	1	1	1	1	1
QL1-L	1.02	1.04	1.03	1.04	1.04
QL2-L	1.01	1.05	1.05	1.05	1.05
QL3-G	1.04	1.09	1.06	1.09	1.09
QN1-G	1.03	1.11	1.12	1.11	1.12
QN2-G	1.05	1.15	1.13	1.18	1.15
QN3-L	**1.10**	**1.20**	**1.18**	**1.23**	**1.20**
S-L	1.01	1.05	1.06	1.05	1.06

In general, all the adaptive MA strategies surveyed were capable of selecting memes that matches the problem appropriately throughout the search, thus producing search performances that are competitive or superior to the canonical MAs on the benchmark problems. It is notable that external-level adaptation fares relatively poorer than local and global-level adaptations. This makes good sense since external-level adaptation does not involve any "learning" on the characteristics of the problem during a search. Particularly, global-level adaptation results in the best search performance among all forms of adaptations considered. This is evident in Table IV where global-level adaptation was ranked the best among all other adaptive MA strategies on all the benchmark problems considered.

2) Computational Cost: Among the adaptive MA strategies, S-E MA incurs minimum extra computational cost over the canonical MA since it does not make use of any historical knowledge but selects the meme randomly. For the sake of brevity, we compare the computational costs of the adaptive MA strategies relative to S-E and tabulated them in Table V.

Generally, qualitative adaptive MAs, i.e., QL1-L, QL2-L, QL3-G, and self-adaptive MA, i.e., S-L, consume only slightly higher computational cost than the S-E. The extra effort is a result of the mechanisms used to choose a suitable meme using historical knowledge at each decision point. Note that QL3-G incurs slightly more computational cost than QL1-L, QL2-L since greater effort is required in the former's tabu-search mechanism to select memes. Overall, since quantitative adaptive MAs, i.e., QN1-G, QN2-G, QN3-L, involve the use of choice-functions in the process of selecting memes, they incur the highest computational costs in comparison. Furthermore, QN3-L is found to require more effort than the others, since it involves the additional distance measure mechanism used in its strategy. Nevertheless, it is worth noting that on the whole, the differences in computational costs between the various adaptive MA strategies may be considered as relatively insignificant.

VI. Conclusion

The limited amount of theory and *a priori* knowledge currently available for the choice of memes to best suit a problem has paved the way for research on developing adaptive MAs for tackling general optimization problems in a robust manner. In this paper, a classification of adaptive MAs based on types and levels of adaptation has been compiled and presented. An attempt to analyze the global convergence properties of adaptive MAs according to their level of adaptations was also presented. Empirical study on adaptive MAs was also presented according to their type-level adaptations. Numerical results obtained on representatives of adaptive MAs with different type-level adaptations using a range of commonly used benchmark functions of diverse properties indicate that the forms of adaptive MAs considered are capable of generating more robust search performances than their canonical MAs counterparts. More importantly, adaptive MAs are shown to be capable of arriving at solution qualities that are superior to the best canonical MAs more efficiently. In addition, among the various categories of adaptive MAs, the global-level MA adaptation appears to outperform others considered.

Clearly there is much ground for further research efforts to discover ever more successful adaptive MAs. From our investigations conducted, we believe there are strong motivations to warrant further research in the areas of memes adaptations.

- The success of global-level adaptation schemes may be attributed to its ease of implementations and the uniformity in the local and global landscapes of the test problems considered. On the other hand, there are insufficient research attempts on other forms of adaptations in MAs. This suggests a need for greater research efforts on local-level and self-adaptive MA adaptations. Some of the experiences gained from global-level adaptation may also apply to other forms of adaptive MAs. For example, besides the simple credit assignment mechanisms used in [13] and [20], more sophisticated mechanisms such as in [11] and [12] may be tailored for self-adaptive MAs. Further, statistical measure may then be used to characterize fitness landscapes or neighborhood structures [41], [42] and the success of the memes on them. Subsequently, knowledge about the diverse neighborhood structures of the problem in hand may be gathered during the evolutionary search process and choice of memes is then achieved by matching memes to neighborhood structures.
- Thus far, little progress has been made to enhance our understanding on the behavior of MAs from a theoretical point of view. It would be more meaningful to provide some transparency on the choice of memes during the adaptive MAs search, see for example [20]. Greater efforts may be expended on discovering rules to enhance our understanding on when and why a particular meme or a sequence of memes should be used constructively, given a particular problem or landscape of known properties. Knowledge derived in this form would help fulfill the human-centered criterion. Besides, domain specialists can manually validate these rules and also use them to enhance their knowledge of the problem domain. This will pave the way for further theoretical developments of MAs and the designs of successful novel memes.
- Most work on meme adaptations have concentrated on using the improvements in solution quality against how much effort incurred to express the capability of memes to search on a problem. Nevertheless, more effective credit assignment mechanisms and rewards should be considered and explored.
- Besides the issue on choice of memes in MAs, a number of other core issues affecting the performance of MAs including interval, duration and intensity of local search have been studied in recent years. Most of these are related to balancing the computational time between local and genetic search [21], [43]. While researchers often experiment with each issue separately, it would be worthwhile to explore how they may be used together to optimize the performance of MAs. For instance, given a fixed time budget, by monitoring the status of a search and the remaining time budget, one may use it as a basis to make decisions on the choice of memes and local/global search ratio. This in turn helps to define the local search interval and duration online throughout the entire evolutionary search.
- Last but not least, it would be interesting to extend the efforts on choice of memes in MAs to multicriteria, multiobjective, and constrained optimization problems, for example, developing appropriate reward measures and credit assignment mechanisms.

Acknowledgment

The authors wish to thank the anonymous referees and editors for their constructive comments on an earlier draft of this paper. They would also like to thank the Intelligent Systems Centre of Nanyang Technological University for their support in this work.

References

[1] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison-Wesley, 1989.
[2] L. O. Hall, B. Ozyur, and J. C. Bezdek, "The case for genetic algorithms in fuzzy clustering," in *Proc. Int. Conf. Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Jul. 1998, pp. 288–295.
[3] J. J. Grefenstette, "Optimization of control parameters for genetic algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-16, no. 1, pp. 122–128, Jan./Feb. 1986.
[4] M. Srinivas and L. M. Patnaik, "Adaptive probabilities of crossover and mutation in genetic algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 4, pp. 17–26, Apr. 1994.

[5] Z. Bingul, A. Sekmen, and S. Z. Sabatto, "Evolutionary approach to multi-objective problems using adaptive genetic algorithms," in *Proc. IEEE Int. Conf. Systems, Man, and Cybernetics*, vol. 3, Oct. 2000, pp. 1923–1927.

[6] R. Hinterding, Z. Michalewicz, and A. E. Eiben, "Adaptation in Evolutionary Computation: A Survey," in *IEEE International Conference on Evolutionary Computation*. Piscataway, NJ: IEEE Press, Apr. 1997, pp. 65–69.

[7] A. E. Eiben, R. Hinterding, and Z. Michalewicz, "Parameter control in evolutionary algorithm," *IEEE Trans. Evol. Comput.*, vol. 3, pp. 124–141, Jul. 1999.

[8] J. A. Miller, W. D. Potter, R. V. Gandham, and C. N. Lapena, "An evaluation of local improvement operators for genetic algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, pp. 1340–1351, Sep./Oct. 1993.

[9] A. Torn and A. Zilinskas, *Global Optimization*. Berlin, Germany: Springer-Verlag, 1989, vol. 350, Lecture Notes in Computer Science.

[10] W. E. Hart, "Adaptive Global Optimization With Local Search," Ph.D. dissertation, Univ. of California, San Diego, 1994.

[11] P. Cowling, G. Kendall, and E. Soubeiga, "A hyperheuristic approach to scheduling a sales summit," in *PATAT 2000*, Springer Lecture Notes in Computer Science, Konstanz, Germany, Aug. 2000, pp. 176–190.

[12] G. Kendall, P. Cowling, and E. Soubeiga, "Choice function and random hyperheuristics," in *Proc. 4th Asia-Pacific Conference on Simulated Evolution and Learning*, Singapore, Nov. 2002, pp. 667–671.

[13] N. Krasnogor, B. Blackburne, J. D. Hirst, and E. K. N. Burke, "Multimeme algorithms for the structure prediction and structure comparison of proteins," in *Parallel Problem Solving From Nature*, 2002, Lecture Notes in Computer Science.

[14] N. Krasnogor, "Studies on the Theory and Design Space of Memetic Algorithms," Ph.D., Faculty of Comput., Math., and Eng., Univ. of the West of England, Bristol, U.K., 2002.

[15] J. E. Smith *et al.*, "Co-evolution of memetic algorithms: Initial investigations," in *Parallel Problem Solving From Nature—PPSN VII*, G. Guervos *et al.*, Eds. Berlin, Germany: Springer, 2002, vol. 2439, Lecture Notes in Computer Science, pp. 537–548.

[16] Y. S. Ong, "Artificial Intelligence Technologies in Complex Engineering Design," Ph.D. dissertation, Sch. Eng. Sci., Univ. Southampton, Southampton, U.K., 2002.

[17] Y. S. Ong and A. J. Keane, "Meta-Lamarckian in memetic algorithm," *IEEE Trans. Evol. Comput.*, vol. 8, pp. 99–110, Apr. 2004.

[18] N. Zhu, Y. S. Ong, K. W. Wong, and K. T. Seow, "Using memetic algorithms for fuzzy modeling," *Austral. J. Intell. Inform. Process. (Special Issue on Intelligent Technologies)*, vol. 8, no. 3, pp. 147–154, Dec. 2004.

[19] E. K. Burke, G. Kendall, and E. Soubeiga, "A tabu search hyperheuristic for timetabling and rostering," *J. Heuristics*, vol. 9, no. 6, 2003.

[20] J. E. Smith, "Co-evolving memetic algorithms: A learning approach to robust scalable optimization," in *IEEE Congress on Evolutionary Computation*. Piscataway, NJ: IEEE Press, 2003, vol. 1, pp. 498–505.

[21] H. Ishibuchi, T. Yoshida, and T. Murata, "Balance between genetic search and local search in memetic algorithms for multiobjective permutation flowshop scheduling," *IEEE Trans. Evol. Comput.*, vol. 7, pp. 204–223, Apr. 2003.

[22] S. G. Byron and W. Joel, *Introduction to Optimization Theory*, ser. International Series in Industrial and Systems Engineering. Englewood Cliffs, NJ: Prentice-Hall, 1973.

[23] R. Dawkin, *The Selfish Gene*. Oxford, U.K.: Oxford Univ. Press, 1976.

[24] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, pp. 239–245, 1979.

[25] A. E. Eiben, E. H. L. Aarts, and H. K. M. Van, "Global convergence of genetic algorithms: A Markov chain analysis," in *Proc. 1st Int. Conf. Parallel Problem Solving from Nature (PPSN)*, 1991, pp. 4–12.

[26] G. Rudolph, "Convergence analysis of canonical genetic algorithms," *IEEE Trans. Neural Netw.*, vol. 5, pp. 96–101, 1994.

[27] J. Suzuki, "A Markov chain analysis on simple genetic algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. 25, pp. 655–659, Apr. 1995.

[28] Y. J. Cao and Q. H. Wu, "Convergence analysis of adaptive genetic algorithm," in *Genetic Algorithms in Engineering Systems Conf.: Innovations and Applications*, Sep. 1997, pp. 85–89.

[29] B. Li and W. Jiang, "A novel stochastic optimization algorithm," *IEEE Trans. Syst., Man, Cybern.*, vol. 30, pp. 193–198, Feb. 2000.

[30] M. Iosifescu, *Finite Markov Processes and Their Application*. New York: Wiley, 1980.

[31] C. A. Floudas, P. M. Pardalos, C. Adjiman, W. R. Esposito, Z. H. Gümüs, S. T. Harding, J. L. Klepeis, C. A. Meyer, and C. A. Schweiger, *Handbook of Test Problems in Local and Global Optimization* Dordrecht, The Netherlands, 1999.

[32] A. O. Griewank, "Generalized descent for global optimization," *J. Optim. Theory Applicat.*, vol. 34, pp. 11–39, 1981.

[33] A. J. Keane. (2002) The OPTIONS Design Exploration System User Guide and Reference Manual. [Online]. Available: http://www.soton.ac.uk/~ajk/options.ps

[34] H. P. Schwefel, *Evolution and Optimum Seeking*. New York: Wiley, 1995.

[35] L. Davis, "Bit climbing, representational bias, and test suite design," in *Proc. 4th Int. Conf. Genetic Algorithms*, 1991, pp. 18–23.

[36] W. H. Swann, *A Report on the Development of a New Direct Searching Method of Optimization*. Middleborough, U.K.: ICI CENTER Instrument Laboratory, 1964.

[37] R. Fletcher, "A new approach to variable metric algorithm," *Comput. J.*, vol. 7, no. 4, pp. 303–307, 1964.

[38] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *Comput. J.*, vol. 7, no. 4, pp. 303–307, 1964.

[39] R. Hooke and T. A. Jeeves, "Direct search solution of numerical and statistical problems," *J. ACM*, vol. 8, no. 2, pp. 212–229, Apr. 1961.

[40] J. A. Nelder and R. Meade, "A simplex method for function minimization," *Comput. J.*, vol. 7, pp. 308–313, 1965.

[41] T. Jones, "Evolutionary Algorithms, Fitness Landscapes and Search," Ph.D. dissertation, Univ. of New Mexico, Albuquerque, NM, 1995.

[42] P. Merz and B. Freisleben *et al.*, "Fitness landscapes and memetic algorithm design," in *New Ideas in Optimization*, D. Corne *et al.*, Eds. New York: McGraw-Hill, 1999, pp. 245–260.

[43] N. K. Bambha, S. S. Bhattacharyya, J. Teich, and E. Zitzler, "Systematic integration of parameterized local search into evolutionary algorithms," *IEEE Trans. Evol. Comput.*, vol. 8, pp. 137–154, 2004.

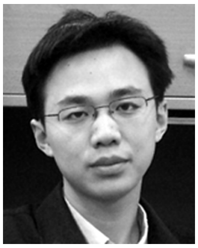
Yew-Soon Ong received the Bachelor's and Master's degrees in electrical and electronics engineering from Nanyang Technology University, Singapore, in 1998 and 1999, respectively, and the Ph.D. degree from he Computational Engineering and Design Center, University of Southampton, Southampton, U.K., in 2003.

He is currently an Assistant Professor with the School of Computer Engineering, Nanyang Technological University. His research interests lie in evolutionary computation spanning: design optimization, surrogate-assisted evolutionary algorithms, memetic algorithms, evolutionary computation in dynamic and uncertain environments, response surface methods for data modeling, and grid-based computing.

Meng-Hiot Lim received the B.Sc., M.Sc., and Ph.D. degrees from the University of South Carolina, Columbia.

He joined the faculty at the School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore, in 1989, where he is currently an Associate Professor. His research interests include computational intelligence, combinatorial optimization, web-based applications, evolvable hardware systems, reconfigurable circuits and architecture, computational finance, and graph theory. He has co-edited a volume on Recent Advances in Simulated Evolution and Learning published by World Scientific Publishing.

Dr. Lim is an Associate Editor of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS.



Ning Zhu received the B.Eng. degree from the School of Computers, WuHan University (WHU), WuHan, China, in 2002.

He is currently a postgraduate student in the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include genetic algorithms and memetic algorithms.

Kok-Wai Wong received the B.Eng. (Hons.) and Ph.D. degrees from the School of Electrical and Computer Engineering, Curtin University of Technology, Western Australia, in 1994 and 2000, respectively.

He joined the School of Information Technology, Murdoch University, Perth, WA, Australia, as a Lecturer in 2005. Prior to this appointment, he was an Assistant Professor at the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include intelligent data analysis, fuzzy rule based systems, computational intelligence, and game technology.