FOUNDATIONS

CrossMark

Domination landscape in evolutionary algorithms and its applications

Guo-Sheng Hao[1] · Meng-Hiot Lim[2] · Yew-Soon Ong[3] · Han Huang[4] · Gai-Ge Wang[5]

Abstract

Evolutionary algorithms (EAs) are usually required to solve problems based on domination relationship among solutions. Often, the domination relationship is almost the sole source of knowledge that EAs can utilize, especially when the problem solving engine concerned is taken as a black box. In this paper, the domination landscape (DL), onto which an optimization problem (OP) can be mapped, is introduced. A DL may correspond to a cluster of OPs, implying that a class of OPs may have the same DL. To illustrate DL, we consider its representation as a directed graph, with its corresponding matrix and function. Of the various properties of DL, the domination-preserving property is used for the analysis of DL-equivalent OPs, and for the basis for classification of OPs. Taking DL as a tool for theoretical analysis, parameters determination for fitness scaling, the convergence property of EAs and the analysis of robustness in light of fitness noise are presented. The study of DL in this paper establishes the necessary theoretical foundation for future applications of DL equality and similarity based optimization.

Keywords Domination · Relationship · Graph · Problems classification · Domination preserving

---

✉ Meng-Hiot Lim
emhlim@ntu.edu.sg

Guo-Sheng Hao
hgskd@jsnu.edu.cn

Yew-Soon Ong
asYSOng@ntu.edu.sg

Han Huang
hhan@scut.edu.cn

Gai-Ge Wang
gaigewang@163.com

[1] School of Computer Science and Technology, Jiangsu Normal University, Xuzhou 221116, Jiangsu, China

[2] School of Electrical and Electronic Engineering, Nanyang Technological University, Jurong West 639798, Singapore

[3] School of Computer Engineering, Nanyang Technological University, Jurong West 639798, Singapore

[4] School of Software Engineering, South China University of Technology, Guangzhou 510641, Guangdong, China

[5] Department of Computer Science and Technology, Ocean University of China, Qingdao 266100, China

1 Introduction

Well-characterized optimization problems can usually be dealt with common mathematical methods, such as derivative based techniques. However, in reality, there are still many problems whereby only domination relationship is available (Borenstein and Poli 2007). For such problems, evolutionary algorithms (EAs) have been shown to perform relatively well. In recent years, researchers have focused on developing sophisticated EAs for solving hard optimization problems (OPs). Among EAs, Genetic Algorithm (GA) (Holland 1992) has emerged as one of the early success stories. Since then, various nature-inspired meta-heuristic algorithms such as evolutionary strategy (Beyer 2001), memetic algorithm (Lim et al. 2011; Ong et al. 2010), particle swarm optimization (Kennedy and Eberhart 2002), ant colony optimization (Sttzle 2004) algorithm, differential evolution (Storn and Price 1997), moth search algorithm (Wang 2016), interactive evolutionary computation (Gong et al. 2015), etc., have emerged as variable alternatives.

During evolution, the algorithm uses the domination relationship among solutions as a basis for selective and preferential exploitation. Domination relationship is commonly used in multi- or many-objective optimization to represent the relationship between individuals (Martł et al. 2016). Here,

Springer

we also use its original meaning, regardless of the number of objectives. One of the expressions of domination relationship is partial order, which has been the issues focused on in many papers. For example, Kang and his team (Liu et al. 2004) studied the multi-objective optimization problems from the viewpoint of partial ordering. Rudolph (2004) studied the domination relationship among solutions with fitness noise and described the change of domination relationship through partial order. However, these studies have the following shortcomings: (1) The expression style of domination relationship is limited to graphs and (2) domination relationship was studied only at the level of population and was not extended to the level of the whole search space. Hence, the application of domination relationship is limited in this sense.

Nevertheless, with regard to domination relationship, there are still many questions or issues to be addressed:

– Are there other modes of expressing domination relationship, for example, in matrix or function form?
– As a result of competition among solutions, can domination relationship be combined with information landscape (Borenstein and Poli 2005) to gain further insights into EAs?
– It is well known that EAs possess property of convergence robustness against fitness noise. What is the mechanism behind the robustness? Can this be explained based on domination relationship?
– Fitness scaling is important not only in adjusting the selection pressure but also in improving the performance of algorithms. Are there any rules that can be formulated to guide the determination of parameters for fitness scaling from the viewpoint of domination relationship?
– There are many difficult OPs for EAs, and some OPs have the same or similar domination relationship. An intuitive question to note is whether more difficult OPs can be solved by taking advantage of domination relationship of other OPs.

We will try to address these issues with a more formal concept, referred to as domination landscape (DL) perspective. To study DL, we restrict our attention to finite search spaces. For any continuous search space, when represented on a digital computer, although it is quite large, the continuum of all possible fitness values is essentially finite (Borenstein and Poli 2006).

This paper is organized as follows. Section 2 concepts on domination and evolution history information are introduced. In Sect. 3, we study DL, DL-equivalent OPs and DL-based OPs classification. Applications of DL are addressed in Sect. 4, which includes the theoretical analysis on the parameters determination of fitness scaling and EAs' convergence robustness against fitness noise. Finally, we conclude this paper in Sect. 5.

2 Preliminary

This section introduces the concept of domination in optimization algorithms, domination graphs and evolution history information.

2.1 Domination in evolutionary algorithms

Let the search space be denoted as S and the real space be R. For single-objective optimization, the optimized function can be written as $f : S \rightarrow R$. The task is to find solutions $\mathbf{X}^* \in S$, such that $\forall \mathbf{X} \in S$, there is $\mathbf{X}^* \succeq \mathbf{X}$. Here, $\mathbf{X}^* \succeq \mathbf{X}$ means that $\mathbf{X}^*$ dominates $\mathbf{X}$ according to the OP f. If f is to be maximized, $\mathbf{X}^* \succeq \mathbf{X}$ means $f(\mathbf{X}^*) \geq f(\mathbf{X})$, and if f is to be minimized, $\mathbf{X}^* \succeq \mathbf{X}$ means $f(\mathbf{X}^*) \leq f(\mathbf{X})$ (Hachicha et al. 2011).

The domination relationship between solutions $\mathbf{X}_1$ and $\mathbf{X}_2$ can be expressed as $\mathbf{X}_1 \succ \mathbf{X}_2$, which is interpreted accordingly as $\mathbf{X}_1$ dominates $\mathbf{X}_2$.

Domination relationship has the property of transitivity such that if $\mathbf{X}_i \succ \mathbf{X}_j$, $\mathbf{X}_j \succ \mathbf{X}_k$, it implies that $\mathbf{X}_i \succ \mathbf{X}_k$.

2.2 Domination graphs

In a directed graph D, denote a vertex as v and an arc (or edge) as e. In Fig. 1, the arc from v_2 to v_3 can be written in partial order style as $< v_2, v_3 >$ which means v_3 dominates v_2.

Let the precursor set for vertex v be denoted as $\Gamma^-(v)$ (Geng and Qu 2014). For example, consider v_3 in Fig. 1 where two arcs, $< v_1, v_3 >$ and $< v_2, v_3 >$ are incident upon it. Therefore, $\Gamma^-(v_3) = \{v_1, v_2\}$, which means that v_3 directly dominates v_1 and v_2. Similarly, denote the successor set for a vertex v as $\Gamma^+(v)$ (Geng and Qu 2014). For v_3, two arcs, $< v_3, v_5 >$ and $< v_3, v_4 >$, emanate from v_3, therefore, $\Gamma^+(v_3) = \{v_4, v_5\}$, which means that v_3 is directly dominated by v_4 and v_5.

The generalized-precursor set $\Gamma^-\{v\}$ is the extension of $\Gamma^-(v)$ based on domination transitivity. For example, the generalized-precursor set $\Gamma^-\{v_5\} = \{v_1, v_2, v_3, v_4\}$. Similarly, the generalized-successor set $\Gamma^+\{v\}$ (Hao et al. 2010b) is the extension of $\Gamma^+(v)$ based on domination transitivity. In Fig. 1, $\Gamma^+\{v_6\} = \{v_7, v_8\}$. Therefore, $\forall v^- \in \Gamma^-\{v\}$ and
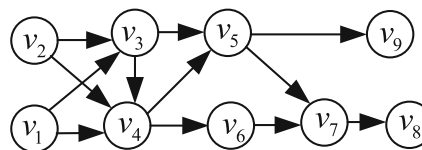

Fig. 1 A directed graph D

Table 1 Cardinal number of $\Gamma^-\{v_i\}$

v_i	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
$\lvert\Gamma^-\{v_i\}\rvert$	0	0	2	3	4	4	6	7	5

$\forall v^+ \in \Gamma^+\{v\}$, we can deduce:

$$v^+ \succ v \succ v^- \tag{1}$$

If all the arcs that can be obtained according to transitivity property are deleted, the new graph is also known as Hasse diagram (Geng and Qu 2014). On the other hand, if all the arcs, that can be obtained according to transitivity property, are added, the new graph (Hao et al. 2010b) is also known as complete diagram. Table 1 shows cardinal number $\lvert\Gamma^-\{v_i\}\rvert$ for the vertices of graph in Fig. 1.

2.3 Evolution history information

Let $I_h(t)$ denote the evolutionary information at generation t. We can then define evolutionary information as:

$$I_h(t) = \{(\mathbf{X}_i(t), f(\mathbf{X}_i(t))) \mid i = 1, 2, \ldots, |P|\} \tag{2}$$

where P is the population and $|P|$ is the population size. Hence, the accumulation of evolution history information can be expressed as (3) below:

$$I_H(t) = \bigcup_{i=0}^{t} I_h(t) \tag{3}$$

$I_H(t)$ provides an information repository for deducing domination relationship.

3 Domination landscape

In this section, we consider the domination landscape (DL) and its representation. Subsequently, the DL-equivalent OPs (DLEOPs) is proposed and after that, the classification of OPs will be described.

3.1 Domination landscape

DL refers to the collective representation of domination relationship among the solutions.

Formally, we define a mapping based on the domination relationship among solutions in search space S to a set as follows:

$$S \times S \rightarrow \{0, 1, -1\} \tag{4}$$

The mapping of solutions is such that: $\forall \mathbf{X}_i, \mathbf{X}_j \in S$, the dominance value $d(\mathbf{X}_i, \mathbf{X}_j)$ is expressed as follows:

$$d(\mathbf{X}_i, \mathbf{X}_j) = \begin{cases} 1, & \mathbf{X}_i \succeq \mathbf{X}_j \\ 0, & \mathbf{X}_i = \mathbf{X}_j \\ -1, & \mathbf{X}_j \succeq \mathbf{X}_i \end{cases} \tag{5}$$

In order to explain DL, we use three types of representation for DL, which includes the directed graph, the corresponding adjacency matrix and function.

- In the directed graph representation for DL, the vertices represent solutions, and the arcs represent the domination relationship, i.e. if there is $\mathbf{X}_i \succeq \mathbf{X}_j$, then there is an arc from the corresponding vertex for $\mathbf{X}_j$ to the vertex for $\mathbf{X}_i$.
- The matrix representation for DL is the adjacency matrix of the above directed graph, and the indices of the row or column correspond to the vertices in the directed graph. The elements of the matrix are from $\{0, 1\}$ and represent the relationship among vertices. If there is an arc from $\mathbf{X}_j$ to $\mathbf{X}_i$, then the value of the element on row i and column j is 1, which represents the domination relation $\mathbf{X}_i \succeq \mathbf{X}_j$.
- In the function representation for DL, the function value of each solution is the number of solutions that it dominates, and the function is defined as:

$$f_{\mathrm{DL}}(\mathbf{X}_i) = \frac{|\Gamma^-\{\mathbf{X}_i\}|}{|S|} = \sum_{\mathbf{X}_j \in \Gamma^-\{\mathbf{X}_i\}} \frac{d(\mathbf{X}_i, \mathbf{X}_j)}{|S|} \tag{6}$$

where $|S|$ is the number of vertices in the DL graph.

From (5) and (6), it is clear that DL function has a desirable property as being single-order function, which transform the original OPs into a simple form, and this property makes DL a powerful tool for many applications. For example, consider an OP max $F_5(x) = x^5, x \in [0, 1]$. For simplicity, we express DL of F_5 with its discrete format and the search space as $\{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$. F_5 and its DL function, DL graph and DL matrix are shown in Fig. 2a–c, respectively. Figure 2b is one of the DL graphs for F_5. Although not all the domination relationships are represented, other domination relationships can be deduced according to the domination transitivity property.

In DL matrix, the upper/lower triangular matrix area marks the distinct elements of the DL. Diagonal elements are all 0 because an element is equivalent to itself according to the domination relationship. The corresponding relationship among DL matrix, DL graph and the curve of DL function is depicted in Fig. 3, where a maximizing OP and three variables $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$ are exemplified.
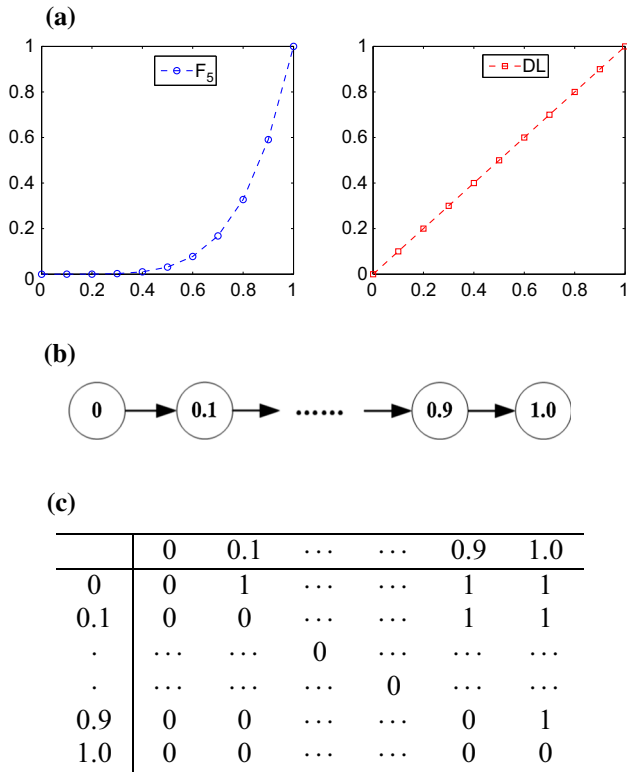
(a)

(b)

(c)

	0	0.1	$\cdots$	$\cdots$	0.9	1.0
0	0	1	$\cdots$	$\cdots$	1	1
0.1	0	0	$\cdots$	$\cdots$	1	1
.	$\cdots$	$\cdots$	0	$\cdots$	$\cdots$	$\cdots$
.	$\cdots$	$\cdots$	$\cdots$	0	$\cdots$	$\cdots$
0.9	0	0	$\cdots$	$\cdots$	0	1
1.0	0	0	$\cdots$	$\cdots$	0	0

Fig. 2 $\arg\max F_5 = x^5$ and its DL representation **a** curve of F_5 at left side and DL function curve for F_5 at right side; **b** a DL graph for F_5; **c** DL matrix for F_5
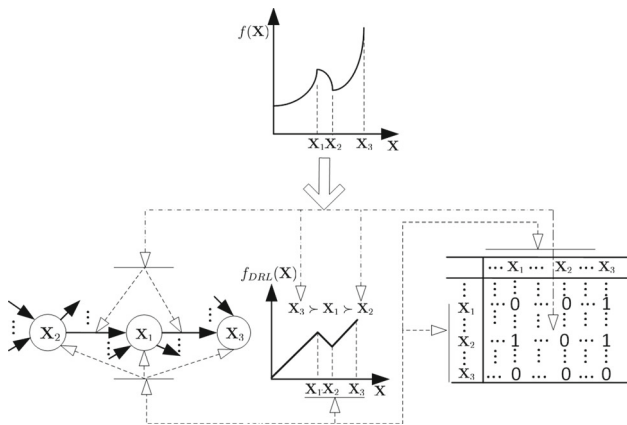

Fig. 3 An OP and its corresponding DL graph, DL function, DL matrix

The advantage of DL with graph expression is that the concrete value of $\mathbf{X}$ is unrelated to the graph. In other words, only the topology of the graph needs to be considered. As such, the absolute values of the solutions can be ignored in the DL. Hence, we can say that DL embodies one of the natural features of OPs.

DL has some useful properties, such as transitivity property. Here, we introduce another property, Domination Preserving (DP), in the form of a theorem.
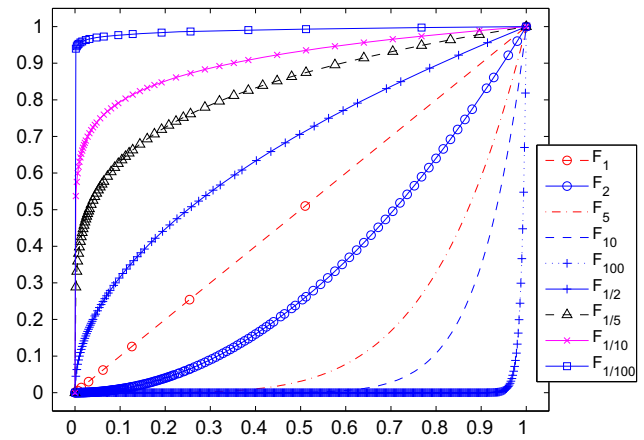

Fig. 4 Curves of $F_k(x) = x^k$, $k = 1, 2, 5, 10, 100, 1/2, 1/5, 1/10$ and $1/100$

Theorem 1 *As a mapping, DL has the property of DP.*

Proof Considering an OP $f(\mathbf{X})$ and its corresponding DL function $f_{\mathrm{DL}}(\mathbf{X})$, for $\forall \mathbf{X}_i, \mathbf{X}_j \in S$, mark $\Delta_f = f(\mathbf{X}_i) - f(\mathbf{X}_j) \neq 0$, $\Delta_{\mathrm{DL}} = f_{\mathrm{DL}}(\mathbf{X}_i) - f_{\mathrm{DL}}(\mathbf{X}_j) \neq 0$. If we can show that the non-zero condition of the product of Δ_f and Δ_{DL} as in (7) is satisfied, then the domination is preserved:

$$\Delta_f \cdot \Delta_{\mathrm{DL}} > 0 \tag{7}$$

We consider the conditions of $\Delta_f > 0$ and $\Delta_f < 0$. For the first case $\Delta_f > 0$, it follows that:

$$\Delta_f > 0 \Rightarrow f(\mathbf{X}_i) > f(\mathbf{X}_j) \Rightarrow \mathbf{X}_i \succ \mathbf{X}_j$$

According to the definition of $\Gamma^-\{\cdot\}$, $\forall \mathbf{X}_j \in \Gamma^-\{\mathbf{X}_i\}$, there is $\Gamma^-\{\mathbf{X}_j\} \subset \Gamma^-\{\mathbf{X}_i\}$. This implies that:

$$\frac{|\Gamma^-\{\mathbf{X}_j\}|}{|S|} < \frac{|\Gamma^-\{\mathbf{X}_i\}|}{|S|} \Rightarrow \Delta_{\mathrm{DL}} > 0.$$

As such, since $\Delta_f > 0$ implies that $\Delta_{\mathrm{DL}} > 0$, the DP property is proven, as per condition (7).

Similarly, it can be proven for the case $\Delta_f < 0$. $\square$

Consider a series of OPs as follows:

$$\max F_k(x) = x^k, \quad x \in [0, 1], \quad k \in R^+ \tag{8}$$

The plots of some of these OPs are shown in Fig. 4. These OPs have the same DL as that of $F_1(x) = x$.

Although the OPs in (8) are different from an optimization viewpoint, they have the same DL. This is because any pair where $x_1 > x_2$, for two different $F_{k1}(x)$ and $F_{k2}(x)$, $k_1 \neq k_2$, the condition $x_1 \succ x_2$ holds. From this viewpoint, $\Delta_{F_{k1}} \cdot \Delta_{F_{k2}} = (F_{k1}(x_1) - F_{k1}(x_2)) \cdot (F_{k2}(x_1) - F_{k2}(x_2)) \geq 0$.

Based on the DP property of DL, the DL-equivalence is studied in the following subsection.

3.2 DL-equivalent optimization problems

With DL, we can derive the notion of DL-equivalent OPs (DLEOPs).

Definition 1 For two optimization problems, OP$_1$ and OP$_2$, if they have the same DL, we say that OP$_1$ is DL-equivalent to OP$_2$ and we denote it as DL(OP$_1$) = DL(OP$_2$).

We also use =$_{DL}$ to express the DL-equivalence relationship, then we express DL(OP$_1$) = DL(OP$_2$) with =$_{DL}$ (OP$_1$, OP$_2$) or OP$_1$ =$_{DL}$ OP$_2$.

It can be further deduced that the DL-equivalence relationship has the properties of symmetry, reflexivity and transitivity. They are:

- =$_{DL}$ (OP$_1$, OP$_2$) ⇔=$_{DL}$ (OP$_2$, OP$_1$) for symmetry;
- =$_{DL}$ (OP$_1$, OP$_1$) for reflexivity;
- =$_{DL}$ (OP$_1$, OP$_2$)∧ =$_{DL}$ (OP$_2$, OP$_3$) ⇒=$_{DL}$ (OP$_1$, OP$_3$) for transitivity.

Hence, as a mapping, DL satisfies the condition of relation of equivalence (Geng and Qu 2014).

For two OPs, how can one judge whether two OPs are DLEOPs? The following theorem provides the justifications.

Theorem 2 *Consider two OPs with the same search space S, and the corresponding functions are f_1 and f_2. If $\forall \mathbf{X}_1, \mathbf{X}_2 \in S$, $\Delta_1 = f_1(\mathbf{X}_1) - f_1(\mathbf{X}_2) \neq 0$, $\Delta_2 = f_2(\mathbf{X}_1) - f_2(\mathbf{X}_2) \neq 0$, and if $\Delta_1 \cdot \Delta_2 > 0$, then the two OPs are DLEOPs.*

Proof For $\Delta_1 \cdot \Delta_2 > 0$, there are two cases as: (1) $\Delta_1 > 0$ and $\Delta_2 > 0$ or (2) $\Delta_1 < 0$ and $\Delta_2 < 0$.

For the first case, $\Delta_1 > 0 \Rightarrow f_1(\mathbf{X}_1) > f_1(\mathbf{X}_2) \Rightarrow \mathbf{X}_1 \succ \mathbf{X}_2$ and $\Delta_2 > 0 \Rightarrow f_2(\mathbf{X}_1) > f_2(\mathbf{X}_2) \Rightarrow \mathbf{X}_1 \succ \mathbf{X}_2$. Therefore, both OPs have the same domination relationship. Because these two solutions are randomly selected, therefore, the result is suitable for all the solutions in the search space. Hence, the two OPs have the same DL.

For the second case, the proof is similar.

The proof is consistent for both maximizing and minimizing OPs. □

Based on the definition of DLEOPs, the classification of problems is studied in the following subsection and the parameters determination for fitness scaling is studied in the next section.

3.3 Optimization problems classification

Based on DL, OPs can be divided into different classes. On the left part of Fig. 5, we illustrate that these functions have the same DL with $F_k(x)$ as in (8). The center of the left part of Fig. 5 is the DL function, and other functions surrounding
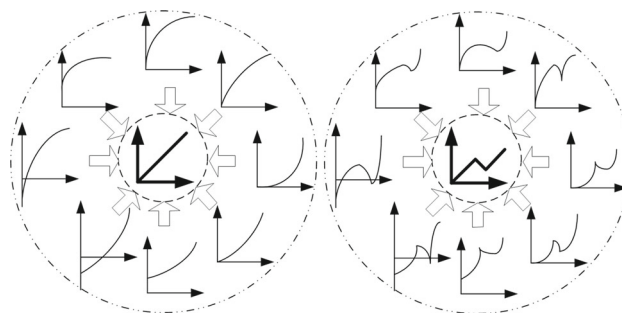
Fig. 5 Classification of optimization problems according to DL: the left DL is composed of one line and the right DL is composed of three lines

it are those problems with the same DL as $F_k(x)$. From this viewpoint, all these OPs belong to the same class. On the right part of Fig. 5, we illustrate a DL with three line segments in the center, around which, many other complex problems have the same DL.

DL can serve as the basis for the classification of some complex problems. Accordingly, we say that those problems sharing the same DL belong to the same class. One example is $F_k(x)$ in (8). Another two examples are shown in Fig. 5. In fact, many other OPs are in the same class as $F_1(x)$, such as $F_k(x) + c$, $x \in [0, 1]$, $k \in R^+$ and $c \in R$. Furthermore, another example that a cluster OPs belongs to the same classification is shown in (9):

$$\max F_{k,c_1,c_2}(x) = x \sin^k(10\pi x + c_1) + c_2, \quad x \in [-1, 2] \quad (9)$$

which has three real constant parameters, $k > 0$, $c_1, c_2 \in R$, and k is a odd number. We depicted some of their figures in Fig. 6, where $k = 1, 9, 39, 509$ and 809. (For the cases of $k = 39, 509$ and 809, the curves are so steep that it is difficult to figure their value out when $x > 0$ until we enlarge part of them as shown in (d)). For this class of OPs, although they have different optima which depends on c_1, they have the same DL.

Formally, let W_i and W_j be the DL matrices of OP$_i$ and OP$_j$ respectively. If they satisfy the equation as per (10):
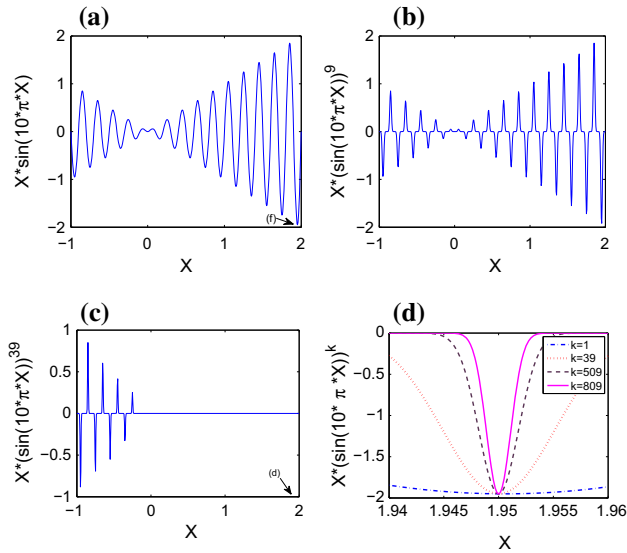
$$W_i = W_j \quad (10)$$

OP$_i$ and OP$_j$ belong to the same class. Intuitively, the optimum can be transferred among the OPs that belong to the same class. The issues pertaining to this will be addressed in another paper.

4 Application of domination landscape

We apply DL to the analysis of parameters determination for fitness scaling, and fitness noise robustness of EAs.

Fig. 6 Another example of OPs in (9) with $c_1 = c_2 = 0$ belong to the same classification from the viewpoint of DL. **a** $k = 1$, **b** $k = 9$, **c** $k = 39$, **d** $k = 1, 9, 39, 509, 809$

4.1 Parameters determination for fitness scaling

Fitness scaling (Jong 1975) plays an important role not only in adjusting the selection pressure, but also in improving the performance of algorithms. There are many popular fitness scaling functions, including the linear function, logarithmic function, power function, exponential function, etc.

The parameters of fitness scaling are not randomly determined. Although there have been some theoretical studies on how these parameters should be determined (Hao et al. 2010a), however, their theoretical basis was not given. Here, we show that the notion of DLEOPs can be a feasible and effective approach for such purpose.

DL-equivalence relation can be an effective representation of fitness scaling. If OPs before and after fitness scaling are not DLEOPs, then the domination relationship will be changed, which means that the rule of survival of the fittest will be violated and the population may proceed to deteriorate as it evolves. Therefore, it is the necessary condition for fitness scaling function to be a DL preserving function. Based on this condition, the parameters of fitness scaling can be determined.

Here, we consider the linear fitness scaling (LFS), which is:

$$F(\mathbf{X}) = \alpha f(\mathbf{X}) + \beta \tag{11}$$

where $f(\mathbf{X})$ is the original optimization function, and $F(\mathbf{X})$ is the function after LFS, and α and β are the constant parameters of LFS.

Theorem 3 *For the OP* max $f(\mathbf{X})$, *the parameter of LFS should satisfy the condition* $\alpha > 0$. *If roulette wheel selection operation is used, the parameters of LFS should satisfy the conditions* $\alpha > 0$ *and* $\beta \geq -\alpha f(\mathbf{X}_0)$, *where* $\mathbf{X}_0 = \arg \min_{\mathbf{X}_i \in P} f(\mathbf{X}_i)$ *where P is the current population.*

Proof Firstly, LFS is a DP map. Then $\forall \mathbf{X}_1, \mathbf{X}_2 \in S$:

$$\begin{aligned}
\Delta_F &= F(\mathbf{X}_1) - F(\mathbf{X}_2) \\
&= \alpha f(\mathbf{X}_1) + \beta - (\alpha f(\mathbf{X}_2) + \beta) \\
&= \alpha(f(\mathbf{X}_1) - f(\mathbf{X}_2)) \\
&= \alpha \Delta_f
\end{aligned}$$

Therefore, there is:

$$\Delta_F \cdot \Delta_f = \alpha(\Delta_f)^2$$

If we want $\Delta_F \cdot \Delta_f > 0$, then we should have $\alpha > 0$ except for the case when $\Delta_f = 0$.

Secondly, taking roulette wheel selection operation into consideration, there is a constraint for $F(\mathbf{X}) \geq 0$. For the solution $\mathbf{X}_0 = \arg \min_{\mathbf{X}_i \in P} f(\mathbf{X}_i)$, it follows that $F(\mathbf{X}_0) = \alpha f(\mathbf{X}_0) + \beta \geq 0$. Therefore, we get $\beta \geq -\alpha f(\mathbf{X}_0)$. □

For other fitness scaling functions, the determination of parameters is similar to the case of LFS.

4.2 EAs' convergence robustness against fitness noise

It has been reported that EAs are generally noise robust (Buche et al. 2002; Arnold and Beyer 2006). However, the reason for this robustness is rarely studied. Here, based on DL, we study EA's robustness subject to fitness noise.

In theory, robustness means that EAs will converge even under the condition of fitness noise. Fitness noise may change the difficulty level of problems, making them harder or easier (Beyer 2000). Here, we do not focus on the hardness issue, but rather study the convergence robustness subject to fitness noise. From the theory of EAs convergence (Rudolph 1998), only if the domination relationship remains unchanged, the convergence will not change. Therefore, the noise robustness can be studied based on DP, which means that we can study the fitness noise robustness based on DLEOPs.

If the noise does not alter the domination relationship among solutions, then EA will be convergence robust (Hao et al. 2006). In other words, if the original OP and the OP with fitness noise are DLEOPs, then EA would be convergence robust.

Denoting the original fitness as $f(\mathbf{X})$, and the fitness with noise as follows (Arnold and Beyer 2006):

$$F(\mathbf{X}) = f(\mathbf{X}) + N(\mathbf{X}) \tag{12}$$

where $N(\mathbf{X})$ is the mapping of fitness noise. $F(\mathbf{X})$ can be seen as the result of from $f(\mathbf{X})$ and can be denoted as:

$$f(\mathbf{X}) \xrightarrow{\text{noise}} F(\mathbf{X}) \tag{13}$$

From the above discussion, we can see that if there is $DL(F(\mathbf{X})) = DL(f(\mathbf{X}))$, then EAs are convergence robust. In other words, EAs are robust for the DLEOPs kind of fitness noise.

We study the robustness from two different viewpoints: (1) the whole search space; (2) the population.

Theorem 4 *EAs are convergence robust against the fitness noise which belongs to DP mapping on the whole search space or on the population.*

If we can prove that the domination of the optimum over other solutions is not changed by fitness noise, then EAs are convergence robust.

Proof Firstly, we consider the fitness noise on whole search space S. For $\forall \mathbf{X}_1, \mathbf{X}_2 \in S$, there is:

$$
\begin{aligned}
\Delta_F &= F(\mathbf{X}_1) - F(\mathbf{X}_2) \\
&= (f(\mathbf{X}_1) + N(\mathbf{X}_1)) - (f(\mathbf{X}_2) + N(\mathbf{X}_2)) \\
&= f(\mathbf{X}_1) - f(\mathbf{X}_2) + N(\mathbf{X}_1) - N(\mathbf{X}_2) \\
&= \Delta_f + \Delta_N
\end{aligned}
$$

Therefore, for the original fitness and fitness with noise, we can conclude that:

$$\Delta_F \cdot \Delta_f = (\Delta_f)^2 + \Delta_f \cdot \Delta_N$$

To ensure $\Delta_F \cdot \Delta_f > 0$, there should be $\Delta_f \cdot \Delta_N > 0$, which satisfies the condition as per Eq. (7). Therefore, EAs are convergence robust against the fitness noise that satisfies DP property.

Secondly, we consider the fitness noise that does not satisfy DP on the whole search space but satisfy DP only on the population.

Denote the tth population as $P(t)$ and the optimum in $P(t)$ as $\mathbf{X}^*(t)$.

Similar to the proof derived from a search space perspective, we have $\Delta_F \cdot \Delta_f > 0$ in $P(t)$.

According to the elitist preservation strategy, we can assume $\mathbf{X}^*(t) \in P(t + 1)$. This means that under the condition that fitness with noise belongs to DP, the population optimum will not only remain dominant over all other solutions in $P(t)$, but also will be retained into the next generation.

Since $S = \lim_{t \to \infty} \bigcup_{i=1}^{t} P(i)$, and the optimum in the population at any point in time exerts domination over the other solutions, the global optimum keeps its domination over all other solutions in the whole search space. □

For this theorem, we summarize it as follows. Firstly, from the viewpoint of search space, if the real OP and noised OP are DLEOPs, EAs are noise robust. However, this is a very strong condition, because it requires the DP property to hold for the whole search space. It is therefore not appropriate for consideration.

Secondly, a somewhat weak condition is given on the population. Since evolution is carried out on the population in each generation, it is not a strong condition for original OP and the noised OP to be DLEOPs in each population than that on the whole search space. This relaxes the condition for fitness noise mapping, which enables EAs to maintain convergence robustness by allowing the fitness noise mapping to be changeable during the evolution process.

As an illustrative example of fitness noise, consider the case of interactive genetic algorithms. Fitness is assigned by one or more users according to preference or subjective judgment. Generally, it is assumed that the user knows nothing about the search space in advance before the evolutionary process. Therefore, the fitness he/she assigns in the first and later generation is likely to be very different from the "real fitness" (Wang et al. 2006) and tends to be noisy or uncertain. However, the noise does not break the domination relationship among the individuals in the same population, i.e., the DP property is preserved according to the user-assigned fitness.

5 Conclusions

In this paper, DL is studied, including its representations graph, matrix and function. Using the notion of DL, the classification of OPs based on DL-equivalence is presented. The application of DL in parameters determination for fitness scaling and EA's convergence robustness against fitness noise are presented. DL's properties such as transitivity and DP are utilized in DLEOPs justifications.

As mentioned before, there are many OPs which have the same DL, among which some are hard for EAs. By exploiting DP mapping, these hard OPs can be easy for EAs. Therefore, the judgement of DLEOPs and the methods for knowledge transfer from easy to hard problems is our direction for future work.

Besides DL equality, DL similarity is also a future topic. For those DL similar problems, knowledge derived from conquered problems can be transferred to hard problems to enhance the search process.

Compliance with ethical standards

References

Arnold DV, Beyer HG (2006) A general noise model and its effects on evolution strategy performance. IEEE Trans Evol Comput 10(4):380–391

Beyer HG (2000) Evolutionary algorithms in noisy environments: theoretical issues and guidelines for practice. Comput Methods Appl Mech Eng 186(2):239–267

Beyer HG (2001) The theory of evolution strategies. Springer, Berlin

Borenstein Y, Poli R (2005) Information landscapes. In: Conference on genetic and evolutionary computation. pp 1515–1522

Borenstein Y, Poli R (2006) Structure and metaheuristics. In: Genetic and evolutionary computation conference, GECCO 2006, proceedings, Seattle, Washington, USA, July, pp 1087–1094

Borenstein Y, Poli R (2007) Decomposition of fitness functions in random heuristic search. In: International conference on foundations of genetic algorithms. pp 123–137

Buche D, Stoll P, Dornberger R, Koumoutsakos P (2002) Multiobjective evolutionary algorithm for the optimization of noisy combustion processes. IEEE Trans Syst Man Cybern Part C (Appl Rev) 32(4):460–473

Geng SY, Qu WL (2014) Discrete mathematics, 5th edn. Higher Education Press, Beijing **(in Chinese)**

Gong D, Chen J, Sun X, Sun J (2015) Evaluating individuals in interactive genetic algorithms using variational granularity. Soft Comput 19(3):621–635

Hachicha N, Jarboui B, Siarry P (2011) A fuzzy logic control using a differential evolution algorithm aimed at modelling the financial market dynamics. Inf Sci 181(1):79–91

Hao GS, Huang YQ, Gong DW, Guo GS, Zhang Y (2006) Fitness noise in interactive evolutionary computation and the convergence robustness. In: Intelligent systems design and applications, ISDA'06. Sixth international conference on, vol 1. IEEE, pp 429–434

Hao GS, Yin YC, Wei KX, Gong G (2010) Parameters selection of fitness scaling in genetic algorithm and its application. In: Chinese control and decision conference. pp 2475–2480

Hao GS, Zhao XJ, Wei KX, Ren SJ (2010) Description of evolutionary computation based on graph theory. J Comput Inf Syst 6(2):653–660

Holland JH (1992) Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. MIT Press, Cambridge

Jong KAD (1975) Analysis of the behavior of a class of genetic adaptive systems. Ph.D. Thesis University of Michigan

Kennedy J, Eberhart R (2002) Particle swarm optimization. In: IEEE international conference on neural networks, 1995. Proceedings, vol 4. pp 1942–1948

Lim MH, Gustafson S, Krasnogor N, Ong YS (2011) Editorial to the first issue. Memet Comput 1(1–2):1–1

Liu MZ, Fen ZX, Kang LS (2004) Efficient multi-objective evolutionary algorithm based on partial order ranking. Mini-micro Syst 24(12):2102–2106

Martł L, Garcia J, Berlanga A, Molina JM (2016) Moneda: scalable multi-objective optimization with a neural network-based estimation of distribution algorithm. J Global Optim 66(4):729–768

Ong YS, Meng HL, Chen X (2010) Research frontier: memetic computation-past, present and future. IEEE Comput Intell Mag 5(2):24–31

Rudolph G (1998) Finite markov chain results in evolutionary computation: a tour d'horizon. Fundam Inform 35:67–89

Rudolph G (2004) A partial order approach to noisy fitness functions. In: Evolutionary computation, 2001. Proceedings of the 2001 congress on, vol 1. pp 318–325

Storn R, Price K (1997) Differential evolution: a simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 11(4):341–359

Sttzle T (2004) Ant colony optimization. IEEE Comput Intell Mag 1(4):28–39

Wang GG (2016) Moth search algorithm: a bio-inspired metaheuristic algorithm for global optimization problems. Memet Comput, pp 1–14

Wang SF, Wang XF, Takagi H (2006) User fatigue reduction by an absolute rating data-trained predictor in IEC. In: IEEE congress on evolutionary computation, pp 2195–2200

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.