A Linkage-Learning Niching in Estimation of Distribution Algorithm

Tsung-Yu Ho
Taiwan Evolutionary Intelligence Laboratory
Department of Electrical Engineering
National Taiwan University
No.1,Sec.4, Roosevelt Rd., Taipei, Taiwan
tsung@teilab.ee.ntu.edu.tw

Tian-Li Yu
Taiwan Evolutionary Intelligence Laboratory
Department of Electrical Engineering
National Taiwan University
No.1,Sec.4, Roosevelt Rd., Taipei, Taiwan
tianliyu@cc.ee.ntu.edu.tw

ABSTRACT

This work proposes a linkage-learning niching method that improves the capability of estimation of distribution algorithms (EDAs) on reducing spurious linkages which increase problems difficulty. Concatenated parity function (CPF), a class of allelic pairwise independent problems, causes exponential scalability for hierarchical Bayesian optimization algorithm (hBOA), which is one of powerful EDAs. Empirical results show that restricted tournament replacement (RTR) that hBOA employs results in spurious linkages and increases difficulty on solving CPF. Our research consists of these goals: (1) proposing a mutual information matrix to approximate the implicit linkage-information during EDAs' execution, (2) reducing spurious linkages by utilizing new metric of similarity, and (3) maintaining diversity of population. The results show that hBOA with our proposed niching method reduces the spurious linkages and solves CPF in the polynomial time.

Categories and Subject Descriptors

G.1.6 [**Optimization**]: Global Optimization

General Terms

Algorithm

Keywords

Genetic Algorithm, Estimation of Distribution Algorithm, Pairwise Linkage Independent, Linkage Learning , Niching

1. INTRODUCTION

In 2007, Coffin and Smith [2] assumed that the allelic pairwise independent problems may increase the problems' difficulty to estimation of distribution algorithm (EDAs) because most EDAs are based on the pairwise linkage detection. The previous works tried the concatenated parity function (CPF), where every pairwise variables in this problem are independent. The results showed that hierarchical Bayesian optimization algorithm (hBOA) scaled in the exponential time. Because hBOA is one of the most effective and powerful EDAs on bounded-difficulty problems, the allelic pairwise independent problems are believed to be difficult to most EDAs that employ pairwise linkage detection. However, another related work [1] tried compact genetic algorithm (CGA) and extended compact genetic algorithm (eCGA) to solve CPF, where both algorithms are simple EDAs. Surprisingly, those two simple EDAs which still employ pairwise linkage detection solved CPF in the polynomial time. It is contrary to the previous assumption that CPF may be difficult for most EDAs. The key difference between hBOA and CGA/eCGA is that hBOA employs the restricted tournament replacement (RTR), a well-known niching method.

Some researches thought that CPF is too simple for CGA and eCGA. Nevertheless, it is unclear that such powerful hBOA solves CPF in the exponential time, so this work does not discuss the issue of limitation. Our work only considers why RTR causes difficulty for hBOA and discusses the insufficiency of RTR for linkage-learning. Finally, we propose the new niching method that is more related to linkage-learning. The experiment shows that hBOA with proposed method solves CPF in the polynomial time.

2. PAIRWISE LINKAGE DETECTION

According to the linkage-learning mechanism of EDAs, the complexity is $O(l^k)$, where k is bounded size and l is problems' size. Hence, most EDAs begins with pairwise linkage detection which is $O(l^2)$. The phenomenon implies that allelic pairwise independent problems may increase problem difficulty. Here, we introduce the CPF with k-bits substructure as follows:

$$CPF(X) = \sum_{i=0}^{m-1} parity(x_{ik} \ldots x_{ik+(k-1)}). \quad (1)$$

The parity function is shown as follows:

$$parity(X) = \left\{ \begin{array}{ll} C_{even} & \text{if } u(X) \text{ is even} \\ C_{odd} & \text{otherwise} \end{array} \right., \quad (2)$$

where $u(X)$ is the unitary of string X, and C_{even} and C_{odd} are defined constants.

It is interesting that CGA and eCGA solve CPF in the polynomial time, even if they do not detect the correct linkage structure. Because both algorithms do not employ RTR,
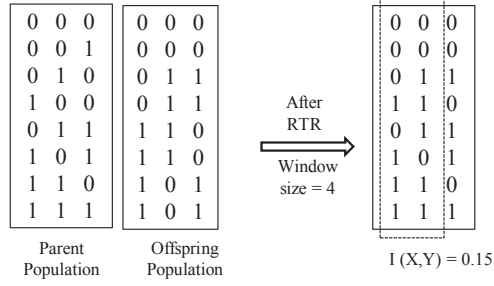
Figure 1: RTR causes extra interaction on the CPF, where any pair of CPF should be independent. The $I(X;Y)$ is mutual information between X and Y location, where $I = 0$ means the independent pair. The bias of extra interaction may form spurious linkage.

the exponential scalability of hBOA may be induced by RTR. To explain this confusion about RTR, we try the simple GA (SGA). In our empirical result, SGA still solves CPF in the polynomial time, but SGA including RTR increases the number of function evaluations. The reason is that RTR increases the extra linkages on those independent pairs and forms the spurious linkages. Although RTR has ability on keeping diversity and maintaining better schemata, the increasing spurious linkages may be the problems, especially on allelic pairwise independent problems. Figure 1 shows how RTR increases extra interaction on independent pair.

3. LINKAGE-LEARNING NICHING

This section introduces our proposed niching method for linkage-learning, especially for those EDAs without linkage information. The main idea of our method enhances the relation between linkage-learning and the measurement of similarity. Some EDAs, like eCGA, have explicit linkage, but we need to establish an approximating linkage-structure for those EDAs without linkage information, like hBOA. The following algorithm shows the approximating method:

Algorithm 1 Mutual Information Matrix Estimation

Input: Sample N individuals with l problem size.
**Output:** Mutual information matrix.
1: **for** $i = 1 \rightarrow l - 1$ **do**
2:     **for** $j = i + 1 \rightarrow l$ **do**
3:         Evaluate MI $I(i; j)$ on N individuals.
4:         $I(j; i) = I(i; j)$
5: **for** $i = 1 \rightarrow l$ **do**
6:     $I(i; i) = 0$

We evaluate the mutual information by following equation:

$$I(X;Y) = \sum_x \sum_y p(x,y) \log \frac{p(x,y)}{p(x)p(y)}, \qquad (3)$$

where X and Y are two random variables, x and y are the outcomes of X and Y respectively, and $p(a)$ is the probability of a. The value of $I(X;Y)$ is between 1 and 0, and the higher value means the higher relation between two variables X and Y.
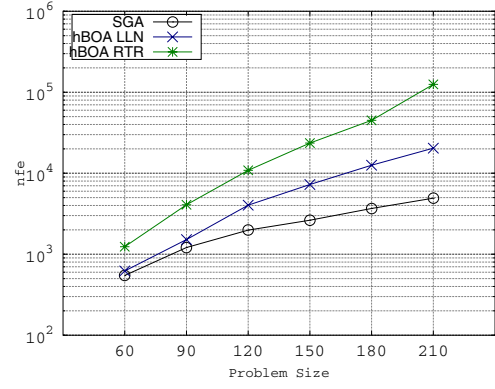
Figure 2: The number of function evaluation on SGA, hBOA with linkage-learning niching, and hBOA with RTR.

The $O(l^2)$ computation time increases the number of function evaluations. However, we only calculus the matrix once in each generation, so the computation time does not affect the total too much. In the next step, we utilize the evaluated matrix to calculate the similarity of two individuals. We define the linkage difference as follows:

$$L_D = \sum_{i=1}^{l} (Diff_i \times \sum_{j=1}^{l} (1 - I(i,j))), \qquad (4)$$

where i and j are the loci of the problem and $Diff$ is the difference of two individuals at loci i.

$$Diff_i = \begin{cases} 1 & \text{if the bits are different} \\ 0 & \text{otherwise} \end{cases} . \qquad (5)$$

L_D accumulates the $1 - I(i,j)$ and presents the distance of two individuals by measuring their linkage similarity. In the final step, we utilize the method to enhance RTR with replacing Hamming distance by our new method. The results show in Figure 2.

4. CONCLUSIONS

Multivariate EDAs have successfully solved many bounded-difficulty problems. However, the relations between linkage-learning and niching are rarely studied. This work proposed a method to approximate the implicit linkage information because some EDAs do not have explicit linkage information. Subsequently, we define a new metric to estimate the distance of two individuals. Empirical results show that our proposed niching with hBOA reduces spurious linkage and solves CPF in the polynomial time. Hence, this idea may be a start to analyze the relation between linkage-learning and niching in the future work.

5. REFERENCES

[1] S.-C. Chen and T.-L. Yu. Difficulty of linkage learning in estimation of distribution algorithms. In *GECCO '09*, pages 397–404, Montréal, Québec, Canada, July 2009.

[2] D. Coffin and R. Smith. The limitations of distribution sampling for linkage learning. In *CEC 2007*, pages 364–369, Sept. 2007.