國立臺灣大學電機資訊學院電機工程學系

碩士論文

Department of Electrical Engineering
College of Electrical Engineering and Computer Science
National Taiwan University

Master Thesis

# 由供應面向探討採用最佳混合之選擇重組基因演算法之族群大小

# Population Sizing for Selecto-Recombinative Genetic Algorithms Using Optimal Mixing in The View of Supply

廖翊雲
Yi-Yun Liao

指導教授：于天立博士
Advisor: Tian-Li Yu, Ph.D.

中華民國 108 年 8 月
August, 2019

doi:10.6342/NTU201902839

doi:10.6342/NTU201902839

# 國立臺灣大學碩士學位論文
# 口試委員會審定書

## 由供應面向探討採用最佳混合之選擇重組基因演算法之族群大小

## Population Sizing for Selecto-Recombinative Genetic Algorithms Using Optimal Mixing in The View of Supply

本論文係廖翊雲君（學號 R06921032）在國立臺灣大學電機工程學系完成之碩士學位論文，於民國 108 年 7 月 10 日承下列考試委員審查通過及口試及格，特此證明。

口試委員：

_____（簽名）
（指導教授）

_____ _____

_____ _____

_____ _____

_____ _____

系主任 _____（簽名）

doi:10.6342/NTU201902839



doi:10.6342/NTU201902839

# 誌謝

首先，感謝于天立教授的指導。從專題一路到碩士班的指導下，對於做研究究竟是什麼以及對於從更全局的觀點去看事物，雖說不敢自稱已經完全了解，但也有了更進一步的體悟。也很感謝教授肯給我機會擔任他的助教，從中反思到對於許多事情在分寸上的拿捏有不同的可能性。最後也很感謝教授各個方面容許不同的意見，雖說有不少次意見上的衝撞，但也因此讓我很多反思以及建立表達意見自信的機會。

再來，感謝泓崴肯讓我在專題時中途加入這個題目並且提供相當的協助幫助我快速進入狀況。當時的討論雖然最後留下成果的部分不多，但是對於後來研究方向的確定以及一些證明中用到的核心概念的啟發有決定性的啟發，真的相當感謝。

接下來，感謝碩士班期間在 TEIL 一起努力的同伴們，感謝你們為這段時間添增不少風采。其中特別感謝在碩一時浩倫幫助盡快融入了實驗室並進入狀況以及遠任維持了實驗室的氣氛。也感謝耿豪跟郁文在研究遇到困難的時候願意聽一下研究進度幫助理清思緒，雖說大多時候好像都是我在唱獨腳戲。

最後，感謝家人以及朋友在這段時間的鼓勵以及支持，能讓我專注在研究的時候生活步調不至於會被打亂，心理狀態也能維持在一定程度。最後的最後，再次感謝上述提到以及未被提到的人在這段時間的幫助，謝謝。



doi:10.6342/NTU201902839



doi:10.6342/NTU201902839

# 摘要

使用最佳混合之基因演算法在實務上有不錯的表現，但在理論方面的支持卻不足。本論文於供應面向下探討採用最佳混合的族群大小。更精確的來說，進行了對於供應包括期望值以及下界等較精確的分析。此外，考慮使用剩餘族群重組一隨機生成之染色體使之達到全域最佳，在此情形下用以提供由神諭機選出的合適片段之緊的族群大小之邊界也被導出。在神諭機引導下的全域族群大小上界也被推導。最後，對於有環狀拓樸結構適應函數之問題，緊的族群大小之邊界也被導出。基於環狀拓樸證明中的概念，一類特定的問題拓樸結構，層狀結構，被定義並且對於可將適應函數視爲層狀結構之問題，族群大小之上界也被導出。環面、超立方以及小世界結構作爲層狀結構之例子被舉出以展示層狀結構的可應用性。

關鍵字： 基因演算法, 最佳混合, 族群大小, 初始供應



doi:10.6342/NTU201902839



doi:10.6342/NTU201902839

# Abstract

Genetic algorithms using optimal mixing have shown promising results, but lack theoretical supports. This thesis investigates population sizing from the supply aspect under the optimal mixing scenario. Specifically, more precise analyses on supply, including the expectation and the lower bound, are made. Furthermore, considering recombining one randomly generated chromosome with the rest of the population to achieve the global optimum, the tight bounds on the size of the population providing proper fragments chosen by restricted oracles are derived. A global upper bound on the size of the population with the guide of an oracle is also derived. Finally, for problem dependent cases, tight bound on the size of the population on problems with fitness functions with ring topologies is derived. Based on the intuition in the proof of the ring topologies case, a category of problem topologies, layered structures, is defined, and upper bounds on the size of the population on problems with fitness functions that can be viewed as layered structures are derived. Examples of layered structures, such as torus and hypercube, are provided to show the applicability of the layered structures.

**Keywords:** Genetic Algorithms, Optimal Mixing, Population Sizing, Initial Supply



doi:10.6342/NTU201902839



doi:10.6342/NTU201902839

# Contents





doi:10.6342/NTU201902839

doi:10.6342/NTU201902839

# List of Figures





doi:10.6342/NTU201902839



doi:10.6342/NTU201902839

# List of Symbols

| | |
|---|---|
| $f$ | Fitness function. |
| $\ell$ | Problem size. |
| $\alpha$ | Exponent of problem size in failure rate. |
| $\eta$ | Exponent to problem size in failure rate. |
| $\boldsymbol{x}, \boldsymbol{y}$ | Chromosomes. |
| $C$ | Set containing all chromosomes. |
| $P$ | Population. |
| $\omega, \psi$ | Operators. |
| $R$ | Set of operators. |
| $\oplus_\beta$ | Addition modulo $\beta$ operator. |
| $\ominus_\beta$ | Subtraction modulo $\beta$ operator. |
| $\oplus$ | $\oplus_\ell$. |
| $\ominus$ | $\ominus_\ell$. |
| $\mathbb{G}_\beta$ | An addative group with $\beta$ elements and operator $\oplus_\beta$. |
| $S_\ell$ | Index set. |
| $M$ | Mask. |
| $\Delta$ | Set of masks. |
| $\lvert M \rvert$ | Size of a mask. |
| $\kappa$ | Maximal size of a mask. |
| $\boldsymbol{x}_M$ | Part of $\boldsymbol{x}$ indicated by a mask $M$. |
| $\nu_\ell$ | Expected number of chromosomes with length of $\ell$ for binary-encoded GAs satisfying having all the schemas of order 1 in the population. |


doi:10.6342/NTU201902839

| | |
|---|---|
| $c$ | Constant used in $c$-composite oracle. |
| $\lambda$ | Layers of a layered structure. |
| $d$ | Dimension of a structure. |
| $r$ | Radius of a structure. |
| $\iota$ | Number of bridges in a small-world topology. |
| $\mathbb{1}$ | Indicator function. |
| $\langle\rangle$ | Ordered set. |


doi:10.6342/NTU201902839

# Chapter 1

# Introduction

Optimal mixing (OM) is an operator widely adopted in modern genetic algorithms (GAs) (Holland, 1975) such as LT-GOMEA (Bosman and Thierens, 2012; Thierens and Bosman, 2011) and DSMGA-II (Hsu and Yu, 2015) and has shown promising results in many applications (Luong *et al.*, 2018; Orphanou *et al.*, 2018; Virgolin *et al.*, 2017). Unlike previous successful developments in GA - estimation of distribution algorithms (Mühlenbein and Paaß, 1996), GAs using OM only requires a relatively small population. This is one of the possible reasons why GAs using OM usually outperform the estimation of distribution algorithms. However, the real reason is yet unknown due to the slow development of theoretical support.

Population sizing is one of the focuses of theoretical developments of GAs. Since it greatly affects the performance of GAs. The amount of information obtained in the initialization process, mainly related to the population size, bounds the performance of a certain run. However, an unnecessarily large population only consumes extra function evaluations than needed to find the global optimum. For traditional GAs, by analyzing facetwise models, there are sound theoretical derivations on population sizing (Goldberg *et al.*, 1992, 2001; Harik *et al.*, 1997; Pelikan *et al.*, 2006; Tung, 2015; Yu *et al.*, 2007), but for GAs using OM, the only applicable results are the supply issue for simple GA addressed by Goldberg *et al.* (2001) and for GAs using OM with non-overlapping masks addressed by Tung (2015). For other models, experiments (Bosman and Thierens, 2012; Hsu and Yu, 2015) showed that the population size does not match. To conclude, new


doi:10.6342/NTU201902839

theories about population sizing for GAs using OM are in need. Based on the background, the following thesis objectives are proposed.

## Thesis Objectives

- Adopt the concept of oracles to provide a tool to analyze the supply model.
- Derive the bounds on the population size for selcto-recombinative GAs using OM for the supply model.
- Apply the adoption of the concept of oracles to specific problems to derive bounds on the population size and computation complexity.

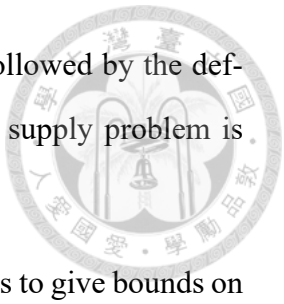
This work focuses on the supply model of selecto-recombinative GAs using OM as the first step. Also, empirically, the population required for GAs using OM is sublinear to the chromosome length (Bosman and Thierens, 2012; Hsu and Yu, 2015; Thierens and Bosman, 2011), which is closest to the supply bound among facetwise models mentioned above. To discuss supply issues only, this work derives bounds on the population size for GAs using OM with an oracle, an automaton that guides the recombination of chromosomes. Because an oracle is a perfect model builder and a perfect decision maker, our derivations focus on the supply issue. As a result, the derived bounds are lower bounds on population size to achieve a certain success rate for GAs using OM. Also, since this work focuses on selecto-recombinative GAs, mutation is not considered here. Nevertheless, the effect of mutation can be integrated into the results of this work via facetwise approaches.

## Roadmap

The rest of the thesis is organized as follows.

- Chapter 2 provides the necessary background of the thesis. Selecto-recombinative GAs and OM are introduced. In addition, related works are included in this chapter.


doi:10.6342/NTU201902839



- Chapter 3 begins the study by a simple problem formulation followed by the definition of oracles used throughout this thesis. Also, the basic supply problem is analyzed in this chapter.
- Chapter 4 extends the discussion and defines $c$-composite oracles to give bounds on the population size with restrictions on the behavior of the oracles.
- Chapter 5 applies the results in Chapters 3 and 4 into specific problems with ring topologies. Computation complexity when using an oracle in those problem types is also discussed.
- Chapter 6 generalizes the results in Chapter 5 into problems with layered structures. Computation complexity when using an oracle in those problem types is also discussed.
- Chapter 7 concludes the thesis with a summary of what has been done, future works, and contributions this thesis provides.


doi:10.6342/NTU201902839



doi:10.6342/NTU201902839

# Chapter 2

# Background

In this chapter, background knowledge of thesis is provided. First of all, the background of GAs and selecto-recombinative GAs are provided. Next, the most important operator in this thesis, OM, is introduced. Also, family of subsets and some associated notations are defined since it is important to GAs using OM. Finally, a brief introduction to facetwise models used in the population sizing of GAs is given, followed by previous works related to supply models in a more detailed manner.

## 2.1 Genetic Algorithms and Selecto-recombinative Genetic Algorithms

Proposed in the late 1960s (Holland, 1975), GAs are a type of stochastic population-based metaheuristic algorithms solving black-box optimization problems inspired by the concept of natural selection. Solving optimization problems is an important task in the field of both science and engineering, and the two concerns are the value of the objective function and the computational cost required. However, for problems proved to be NP-Hard, such as traveling salesman problem and k-minimum spanning tree problem, no deterministic algorithms that can give the exact global optimum in polynomial time, and metaheuristic algorithms are good choices to solve these problems. Instead of searching the whole feasible which is usually exponential to the problem size, metaheuristic algorithms try to


doi:10.6342/NTU201902839

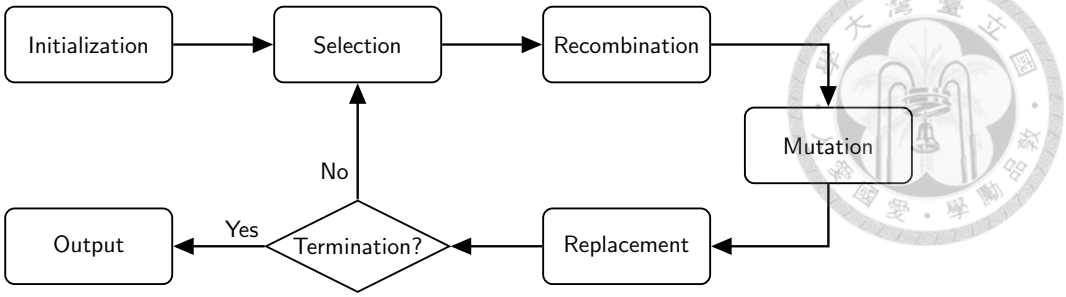



Figure 2.1: Illustration of a general flowchart of GAs.

gain more knowledge about the problem by combining prior knowledge and the characteristics found in the process and to propose a more promising feasible solution. Some metaheuristic algorithms are inspired by natural phenomena, and GAs are one on them. In the process of natural selection, creatures suit the environment thrive while those do not vanishes, and in GAs, the candidate solutions, encoded into bit strings with required cardinality and problem length and called **chromosomes**, with better values on the objective function, denoted as the **fitness function**, survive through the process while others are discarded.

Figure 2.1 is an illustration of a flowchart of GAs. In general, a GA can be decomposed into the following phases: initialization, selection, recombination, mutation, and replacement. Initialization is a phase where the set of chromosomes, the **population**, is set. Each chromosome is initialized independently and identically, and in most cases, without any prior knowledge of the fitness function, in any chromosome, for each position, called **locus**, the value, denoted as the **allele**, is determined by a distribution such that all feasible solutions are equally likely to be sampled. After initialization is done, the rest of the phases form a cycle, and each cycle is called a generation.

In each generation, firstly, a selection is performed literally to preserve the part of a population with relatively better fitness values and enhance the probability that a more promising chromosome is observed. After selection, recombination, representing the reproduction in nature, processes are conducted on the preserved chromosomes in hopes of creating a chromosome with its fitness value being closer to the one of the global optimum. To increase the diversity of the population, mutation is performed by randomly


doi:10.6342/NTU201902839

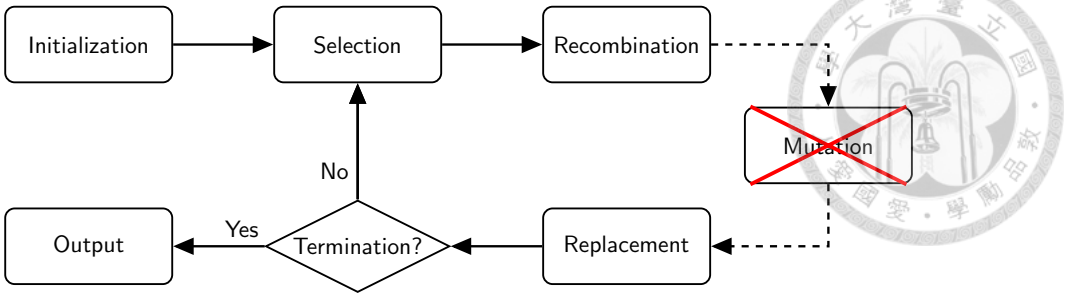



Figure 2.2: Illustration of a general flowchart of selecto-recombinative GAs.

changing alleles. Just like the case in nature, mutation happens with a relatively low likelihood. Finally, in nature, as the generation changes, those are old will be replaced with those new. A replacement phase is needed to determine the chromosomes that can go to the next iteration or so-called **generation**. The simplest replacement stage is replacing the chromosomes being fed into the recombination operators with those newly generated, while other skills such as niching can also be applied here.

Termination happens when either of the two conditions is satisfied. The first one is that the limit of computation, usually determined by the number of the fitness function is queried and denoted as the number of function evaluations (nFE), is met. The second one is that under no limitation on the computation power, all chromosomes in the population converge, meaning no better solution can be generated without mutation. Note that none of the conditions guarantees the output solution to be optimal.

The convergence of GAs can be viewed as the vanishing of the diversity of information. In this sense, mutation tries to maintain the diversity of information or at least tries to postpone the vanishment of diversity. However, since mutation is a stochastic process that seldom happens, it increases the difficulty in the analysis of convergence. Thus, selecto-recombinative GAs (Goldberg *et al.*, 1993) are proposed. Selecto-recombinative GAs are GAs without mutation. Figure 2.2 is an illustration of a general flowchart of selecto-recombinative GAs.


doi:10.6342/NTU201902839

## 2.2 Optimal Mixing

This section introduces the only recombination operator discussed in this thesis, OM. Proposed by Thierens and Bosman (2011), OM is a recombination operator. In the operation of OM, two chromosomes, namely the donor and the receiver, are involved. The donor copies some of its bits to the receiver's corresponding loci. The action, also known as a **donation**, is accepted by the receiver only if the value of the fitness function does not decrease. Figure 2.3 gives an example of one operation of OM when solving the one max problem.

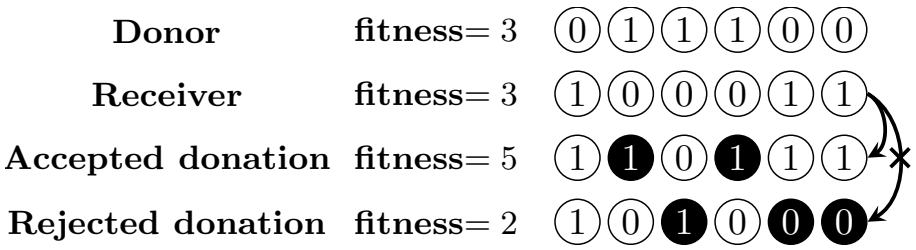



Figure 2.3: Example of OM. Suppose solving the one max problem. The black circles indicate the donated bits. Both examples of accepted and rejected donation are given.

OM is a hill climber with a customized neighborhood and has shown promising results in many applications (Luong *et al.*, 2018; Orphanou *et al.*, 2018; Virgolin *et al.*, 2017). The difference between OM and other recombination operators in GA is that in the operation of OM, the result is determined by the change in the value of fitness function. This enables OM to filter out chromosomes with less desired information in the view of the value of the fitness function before it is introduced to the population.

## 2.3 Family of Subsets

The operation of OM can be viewed as actions taken on bits with linkage defined by GAs, and the family of subsets (FOS) (Thierens, 2010), denoted $\mathcal{F}$, is originally designed to show the linkage of bits. To define FOS, the following notations introduced. A chromosome $\boldsymbol{x}$ is an ordered set of bits with each bit having cardinality $\chi$. The length of the


doi:10.6342/NTU201902839

chromosome is denoted $\ell$, and the set of natural numbers is denoted as $\mathbb{N}$. Based on these notations, some definition relating to FOS can be defined as follows.

**Definition 2.1.** *For any $\ell \in \mathbb{N}$, an* **index set**

$$S_\ell = \{0, 1, \ldots, \ell - 1\}.$$

*Any subset of $S_\ell$ is called a* **mask**.

**Definition 2.2.** *For an index set $S_\ell$, a* **FOS** $\mathcal{F}$ *satisfies the following properties.*

1. $\mathcal{F} = \left\langle M^1, M^2, \ldots, M^{|\mathcal{F}|} \right\rangle$ [1], *where* $\forall\, i \in \{1, 2, \ldots, |\mathcal{F}|\}$, $M^i \subseteq S_\ell$.
2. $\cup_{i \in \{1,2,\ldots,|\mathcal{F}|\}} M^i = S_\ell$.

**Definition 2.3.** *For an index set $S_\ell$, a FOS $\mathcal{F} = \left\langle M^1, M^2, \ldots, M^{|\mathcal{F}|} \right\rangle$ is a* **disjoint FOS** *if*

$$\forall\, i, j \in \{1, 2, \ldots, |\mathcal{F}|\}, i \neq j, M^i \cap M^j = \emptyset,$$

*where $\emptyset$ denotes an empty set.*

**Definition 2.4.** *For an index set $S_\ell$ and $k \geq 2$ such that $(k-1) \mid \ell$, a FOS is a* **homogeneous overlapping FOS**, *denoted as $\mathcal{F}_{\langle k \rangle}$, if*

$$\mathcal{F}_{\langle k \rangle} = \left\langle M^1, M^2, \ldots, M^{\frac{\ell}{k-1}} \right\rangle,$$

*where* $\forall\, i \in \left\{1, 2, \ldots, \frac{\ell}{k-1}\right\}$,

$$M^i = \{j | j \in S_\ell, (i-1)(k-1) \leq j \leq (i(k-1) \bmod \ell)\}.$$

Taking an example when $\ell = 3$. $S_\ell = \{0, 1, 2\}$. $\{0\}$ and $\{0, 1\}$ are masks. $\langle \{0, 1\}, \{1, 2\} \rangle$, $\langle \{1, 2\}, \{0, 1\} \rangle$, and $\langle \{1, 2\}, \{0, 1\}, \{1, 2\} \rangle$ are three FOSs, while $\langle \{1, 2\} \rangle$ is not a FOS since $0$ does not appear in the union of masks. $\langle \{0\}, \{1, 2\} \rangle$ is a disjoint FOS since $\{0\} \cap \{1, 2\} = \emptyset$. However, $\langle \{0, 1\}, \{1, 2\} \rangle$ is not a disjoint FOS since $\{0, 1\} \cap \{1, 2\} =$

[1] $\langle\rangle$ denotes an ordered set.


doi:10.6342/NTU201902839

$\{1\}$. $\mathcal{F}_{\langle 2\rangle} = \langle \{0,1\}, \{1,2\}, \{0,2\} \rangle$. Note that the order of masks matter in FOS, and there may be duplicated masks.

Elements in FOS are called masks because, during OM, variables are mixed according to the masks in FOS, and the operation can be viewed as a variable-wise mask operation. Also, note that the homogeneous overlapping FOS defined in this thesis is different from the one defined by Tung (2015). The homogeneous overlapping FOS defined by Tung consists of some pairs of masks, and both masks in every pair share one common index, while masks from different pairs are disjoint. In contrast, in the homogeneous overlapping FOS defined in this work, masks can no longer be separated into pairs.

## 2.4 Facetwise Models

Since GAs are a kind of stochastic population-based algorithm. The control of the population size, called population sizing, plays an important role in the performance since the population size upper bounds the amount of information stored in the population at once. However, in GAs, there are many stochastic steps taken. Thus, difficulties arise when directly analyzing the relationship between the population size and the whole process. To gain more understanding regarding the run of a GA, facetwise models are used. Facetwise model, first used by Goldberg, is a simplified model used to know how a part of a complex system works.

Concerning population sizing, three major facetwise models, which are supply, decision making, and model building, are often addressed. Supply model focuses on the population size needed to ensure having sufficient information to recombine the global optimum. An important issue in the supply model is collecting **schemas** (Holland, 1975). A schema is a collection of alleles at specific loci. In order words, collecting schemas is equivalent to collecting chromosomes with desired bit values under certain masks. More detailed results will be discussed in the next section.

Addressed by Goldberg *et al.* (1992) and Harik *et al.* (1997), decision making models focus on the population size needed to distinguish chromosomes with global optimum fragments from those without. The need of decision making models arises because even


doi:10.6342/NTU201902839

under the assumption that global optimum fragment is optimal in the sub-function, the sampling noise due to inter-competition between schemas may result in filtering out some of the fragment accidentally. Thus, a sufficient amount of chromosomes is needed to conquer the noise.

As GAs develope, more efficient information retrieving and recombining techniques are in need. A major way is to build a model that approximates the fitness function based on known information. Thus, the need for model building, which examines the population size needed to learn a high-quality model to generate high-quality chromosomes, arises. There are derivations for model building derived by Pelikan *et al.* (2006) and Yu *et al.* (2007). However, the main focuses of these works are identifying the correct linkage between bits, and more operator-oriented analyses still need to be done.

## 2.5 Related Works

In this section, a review of the previous results on supply is given. Among the previous works regarding the population sizing using the supply model, two works are the most related. Specifically, results derived by Goldberg *et al.* (2001) and Tung (2015) are introduced.

Goldberg *et al.* (2001) derived that in order to guarantee having all schemas of all fragments of a chromosome with cardinality of $\chi$ with $m$ building blocks with each fragment having an order of $k$ with a success rate of $1 - \frac{1}{m}$, assuming $m$ is large, let $n$ be the population size, $n$ must satisfy

$$1 - \frac{1}{m} = \exp\left(-\chi^k \exp\left(-\frac{n}{\chi^k}\right)\right),$$

and solving $n$ yields

$$n = \mathrm{O}\left(\chi^k \left(k \log \chi + \log m\right)\right).$$

Tung (2015) derived that in order to guarantee all the global optimal fragments of order


doi:10.6342/NTU201902839

$k$ exists in the population with a probability of $p$, the following equation holds.

$$p = \left(1 - \left(1 - \chi^{-k}\right)^{n}\right)^{m},$$

where $\chi$ is the cardinality of genes, $m$ is the number of building blocks in the chromosomes, and $n$ is the population size. Letting $p$ be constant and solving $n$ yields

$$n = \Theta\left(\chi^{k} \log m\right).$$

Also, there is a brief discussion on problem with separable overlapping structures, though no theorems are developed.

Both works mainly discuss problems with fully separable non-overlapping problem structures. The result derived by Tung (2015) implies that the bound derived by Goldberg *et al.* (2001) is tight if the success rate is constant. However, in the real world, fully separable non-overlapping problem structures is a restriction that is too strict. Thus, a study of the supply model in more general cases is in need.


doi:10.6342/NTU201902839

# Chapter 3

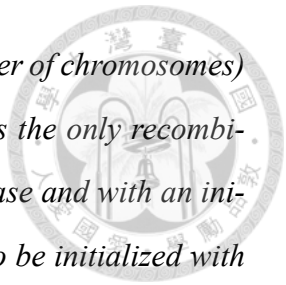
# Oracles and Basic Supply Problem

To begin the study, in this chapter, the problem formulation is specified. Next, since this thesis derives bounds in the view of supply, the concept of oracles is introduced to optimally handle the run of GAs. The last part of the section the most basic supply problem: having all the schemas (Holland, 1975) of order 1 in the population for selecto-recombinative GAs using OM is studied.

## 3.1 Problem formulation

As mentioned in Chapter 1, the development in the population sizing of GAs using OM is still in an early stage. This thesis focuses on the population sizing of GAs using OM in the view of supply. However, this is still a problem too big to be solved at once. Thus, this thesis focuses on the **binary-encoded** chromosomes as the prototype. Also, **selecto-recombinative GAs** are the main focus in order to rule out the process of mutation, which is not considered in the supply model. For the same reason, the selection phase is neglected. In addition, without further prior knowledge about the problem formulation, the initialization process should not inject certain inter or intra tendencies among bits. Thus, each bit of the chromosome is assumed to be initialized with **identically and independently distributed Bernoulli distribution** with $\frac{1}{2}$ being the probability determined to be 1. As the last restrictions, OM is the only recombination operator allowed to be used, and simple replacement is chosen. To sum up, the problem to be answered is as follows.


doi:10.6342/NTU201902839



*Let $\ell$ be the chromosome length, what is the population size (number of chromosomes) required for binary-encoded selecto-recombinative GAs using OM as the only recombination operator followed by simple replacement with no selection phase and with an initialization process in which each bit of the chromosome is assumed to be initialized with identically and independently distributed Bernoulli distribution with $\frac{1}{2}$ being the probability determined to be $1$ to have a success rate of $1-\Theta\left(\ell^{-\alpha}\right)$ in the view of supply?*

The problem asked above does not show how to make the success rate solely relies on supply, and the workaround will be discussed in the next section, but before that, to simplify to notation, in the following content, if not specified $\ell$ is the chromosome length and binary-encoded selecto-recombinative GAs using OM as the only recombination operator followed by simple replacement with no selection phase and with initialization process in which each bit of the chromosome is assumed to be initialized with identically and independently distributed Bernoulli distribution with $\frac{1}{2}$ being the probability determined to be $1$ is abbreviated as *binary-encoded selecto-recombinative GAs using OM*.

To make clear and simplify the notations used when discussing arguments relating to chromosoes. For $\beta \in \mathbb{Z}^+$, define operators $\oplus$ and $\ominus$ as $\forall\ s,t \in \mathbb{Z},\ s \oplus_\beta t = (s+t) \bmod \beta$ and $\forall\ s,t \in \mathbb{Z},\ s \ominus_\beta t = (s-t) \bmod \beta$. Then a chromosome $\boldsymbol{x} = \langle x_{[i]} \rangle$, where $[i] \in \mathbb{G}_\ell$ in which $\mathbb{G}_\beta$ is an additive group of $\beta$ elements with the set $\{[0],[1],\ldots,[\beta-1]\}$ and operator $\oplus_\beta$. Also, for any mask $M$, $\boldsymbol{x}_M = \left\langle x_{[i]} \middle| i \in M \right\rangle$, and if not specified, $\oplus = \oplus_\ell$ and $\ominus = \ominus_\ell$.

## 3.2 Oracle

In the last section, the problem to be solved is specified. Since this thesis concentrates on the supply issue, to have the success rate of GAs solely relies on supply, the concept of oracles is adopted to optimally handle the recombination tasks under certain restriction as desired. An oracle is defined as follows.

**Definition 3.1.** ***oracle****: For GAs on an optimization problem, an oracle is an automaton with input in the form of $3$-tuple $(f, P, \{\omega\})$. $f$ is the fitness function. $P$ is a population, a*


doi:10.6342/NTU201902839

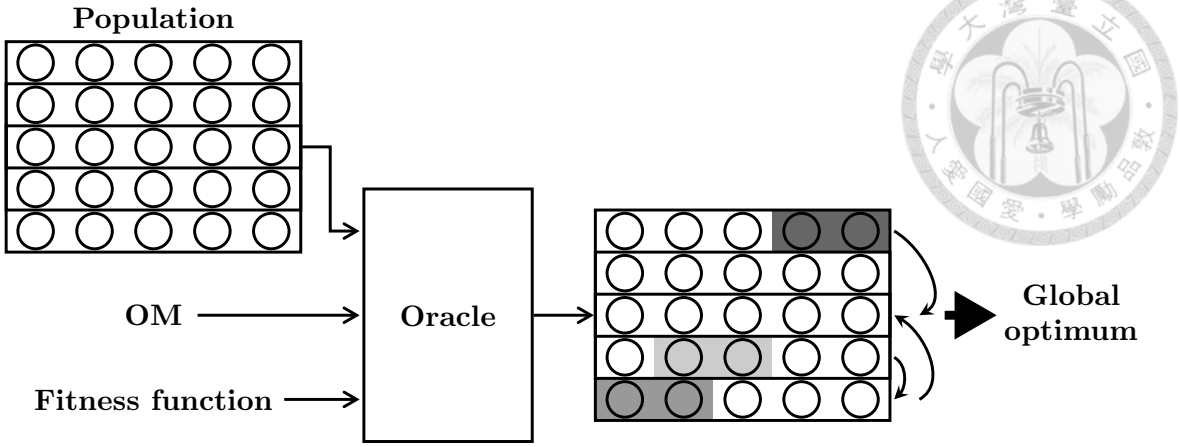



Figure 3.1: Illustration of an oracle considered in this thesis. Each row in the population represents a chromosome, and each circle in a chromosome represents a bit. Each arrow on the right of the population represents a donation, with the chromosome next to the base being the donor. The colored area on the donor is the mask used in the donation. Donations are sorted with the one on the left happened first.

*set of chromosomes. In other words, let $C$ be the set containing all chromosomes, $\forall\ \boldsymbol{x} \in P, \boldsymbol{x} \in C$. $\omega$ is an operator, defined as a function $\in R$, where $R : \left(C^{\zeta}, \langle\Delta\rangle\right) \to C^{\zeta'}$ for $\zeta,\ \zeta' \in \mathbb{N}$. $\Delta$ is a set of masks (Hsu and Yu, 2015), which is a set of loci. The operator must satisfy the following constraint: for each allele in the output set, there must exist one chromosome in the input set with the same allele at the same locus. The oracle optimizes $f$ by recombining chromosomes among $P$ using $\{\omega\}$ as recombining operators. If any chromosome in $P$ can achieve the optimal solution, the oracle returns a sequence of $(\langle\boldsymbol{y}\rangle, \psi, \langle M\rangle)$, where $\boldsymbol{y} \in C$, $\psi \in R$, and $(\langle\boldsymbol{y}\rangle, M)$ is the corresponding input of $\psi$. If the optimal solution is not achievable, the oracle returns a* FAIL.

In other words, in this thesis, the oracle takes an input of fitness function, a population, and description of OM and the oracle returns a sequence of donors, description of OM, and masks. Figure 3.1 gives an illustration of an oracle discussed in this thesis. Notice that any stochastic operators such as randomized mutation and operators that generate alleles not existing in the input set such as complement a binary-encoded chromosome are not valid. Thus, the amount of information the oracle can have is upper bounded by the input of the


doi:10.6342/NTU201902839

oracle. Since if given sufficient supply to solve the problem, the oracle never outputs FAIL, the probability of the supply is sufficient to solve a problem and the success rate of GA using OM with an oracle are of the same value. In this way, the probability of the supply is sufficient to solve a problem and the success rate of GAs are directly related. However, since the two issues are different, in the following derivations success probability instead of success rate is used when the solving scheme uses an oracle. Also, since only focusing on cases with the aid of an oracle, the derived bounds are lower bounds on the population size for binary-encoded GAs using OM and are tight when using an oracle in Chapters 3 and 4 whereas only upper bounds when using an oracle are derived in Chapters 5 and 6.

## 3.3 Basic Supply Problem

Consider the most basic supply problem: having all the schemas of order 1 in the population. This can be viewed as the supply problem for selecto-recombinative GAs using an arbitrary recombination operator. Thus, this can be viewed as a requirement of solving a problem by an omnipotent oracle, defined as follows.

**Definition 3.2.** ***omnipotent oracle****: An omnipotent oracle is an oracle that uses an arbitrary set of operators defined in Definition 3.1.*

The omnipotent oracle can be viewed as a selecto-recombinative GA with an arbitrary recombination operator. Thus, the requirement on the population size for selecto-recombinative GAs using an omnipotent oracle is a lower bound for every GA using an oracle.

Also, this problem is in effect a special case of problems already discussed by Goldberg *et al.* (2001), with cardinality of 2 and order of building block of 1; however, their result is an upper bound and is not applicable here. The results derived by Tung (2015) are not applicable here either since the need is to collect all schemas, not only the ones that the global optimum consists of. Though this problem is a variation of the coupon collector's problem (de Moivre, 1711), no applicable result specialized for GAs exists.


doi:10.6342/NTU201902839

### 3.3.1 Expectation

Before deriving the lower bound on the population size required to have the success probability higher than certain value directly, the expectation, which is an easier task and provides as an upper bound on the lower bound desired, is investigated. Specifically, this subsection focuses on the expected number of chromosomes with length of $\ell$ for binary-encoded GAs satisfying having all the schemas of order $1$ in the population, denoted by $\nu_\ell$. Based on the definition of $\nu_\ell$, the following Lemmas hold.

**Lemma 3.3.** *$\nu_\ell$ is O $(\log \ell)$.*

*Proof.* Consider solving $\nu_\ell$ using recursive relation. Let $\nu'_\ell$ be the expected number of chromosomes needed after the first chromosome is initialized and thus, $\nu'_{chi,\ell} = \nu_\ell - 1$. Since each bit is initialized identically and independently, the difference of uncollected schemas before and after another chromosome is initialized is a random variable with Binomial distribution with $\frac{1}{2}$ being the probability of success and the number of uncollected schemas before another initialization the being the number of sample size. Thus, the recursive relation of $\nu'_\ell$ can be written as follows.

$$\nu'_\ell = \begin{cases} 1 + \sum_{i=0}^{\ell-1} \binom{\ell}{i} 2^{-\ell} \nu'_{\ell-i} & l \geq 1 \\ 0 & \ell = 0 \end{cases}.$$

Now define another recursive relation $\nu^+_\ell$ as follows to bound $\nu'_\ell$.

$$\nu^+_\ell = \begin{cases} 1 + \frac{1}{2}\nu^+_{\lfloor \frac{\ell}{2} \rfloor} + \frac{1}{2}\nu^+_\ell & \ell \geq 1 \\ 0 & \ell = 0 \end{cases}.$$

Since $\nu_\ell$, $\nu'_\ell$, and $\nu^+_\ell$ are monotonically increasing, and the median of a Binomial distribution is between the ceiling of success probability multiplied by number of sample size and the floor of the same value, by mathematical induction, $\forall\ i \leq \ell, \nu_i \leq \nu'_i \leq \nu^+_i$. Since $\nu^+_\ell \leq 2\log_2 \ell + 2$, $\nu_\ell$ is O $(\log \ell)$. □


doi:10.6342/NTU201902839

**Lemma 3.4.** *$\nu_\ell$ is $\Omega(\log \ell)$.*

*Proof.* Consider solving $\nu_\ell$ directly. First, $\Pr\{$a schemas of order $1$ for one locus are collected with a population with size smaller than or equal to $i\} = 1 - 2^{-(i-1)}$. Since all loci are independent, $\Pr\{$all schemas of order $1$ are collected with a population with size smaller than or equal to $i\} = \left(1 - 2^{-(i-1)}\right)^\ell$. Therefore, let $\lambda$ be an integer greater than or equal to $\lfloor \log_2 \ell \rfloor + 1$, $\nu_\ell =$

$$
\begin{aligned}
&\sum_{i=2}^{\infty} i \left( \left(1 - 2^{-(i-1)}\right)^\ell - \left(1 - 2^{-(i-2)}\right)^\ell \right) \\
\geq &\sum_{i=2}^{\lambda} i \left( \left(1 - 2^{-(i-1)}\right)^\ell - \left(1 - 2^{-(i-2)}\right)^\ell \right) + (\lambda+1)(1 - (1 - 2^{-(\lambda-1)})^\ell) \qquad (3.1) \\
= &- \sum_{i=2}^{\lambda-1} \left(1 - 2^{-(i-1)}\right)^\ell + \lambda + \left(1 - 2^{-(\lambda-1)}\right)^\ell \\
\geq &- \sum_{i=2}^{\lambda-1} \left(1 - 2^{-(i-1)}\right)^\ell + \lambda \\
\geq &- \sum_{i=2}^{\lfloor \log_2 \ell \rfloor + 1} \left(1 - 2^{-(i-1)}\right)^\ell - \sum_{i=\lfloor \log_2 \ell \rfloor + 2}^{\lambda-1} \left(1 - 2^{-(i-1)}\right)^\ell + \lambda \\
\geq &- \frac{1}{e} \lfloor \log_2 \ell \rfloor - (\lambda - \lfloor \log_2 \ell \rfloor + 2) + \lambda \\
\geq &(1 - \frac{1}{e}) \lfloor \log_2 \ell \rfloor - 2.
\end{aligned}
$$

Therefore, $\nu_\ell$ is $\Omega(\log \ell)$. □

Combining Lemmas 3.3 and 3.4, the tight bound of $\nu_\ell$ is given in the following theorem:

**Theorem 3.5.** *$\nu_\ell$ is $\Theta(\log \ell)$. That is to say, for a binary-encoded GA with chromosome with length of $\ell$, the expected number of chromosomes satisfying having all the schemas of order $1$ in the population is $\Theta(\log \ell)$.*

*Proof.* The result can be directly inferred from Lemmas 3.3 and 3.4. □

Some experiment is conducted to verify the derivation (Figure 3.2) (Liao *et al.*, 2019). The experiment is conducted on $\ell = 2^i$ for $i = 1$ to $15$, and for each $\ell$, each bit of a chromosome is assigned with Bernoulli distribution with $\frac{1}{2}$ being the probability of success.


doi:10.6342/NTU201902839

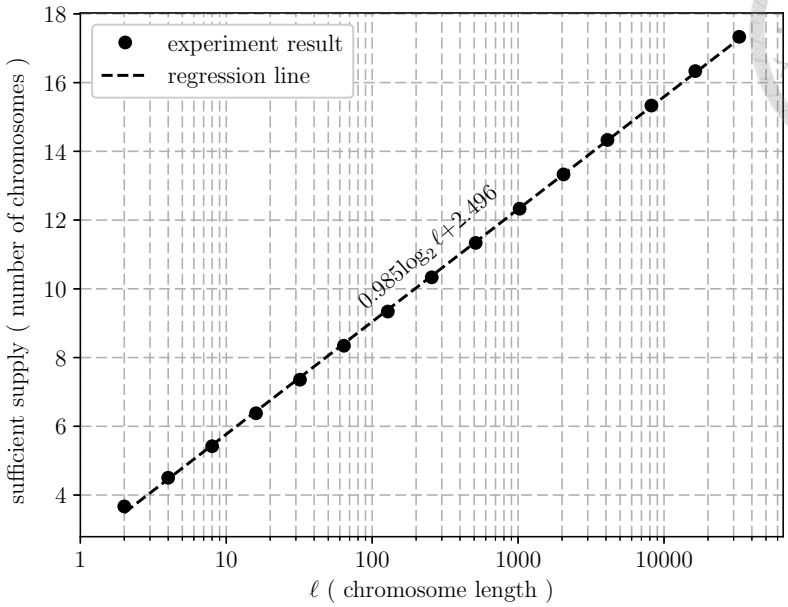



Figure 3.2: Experiment on the expectation of the basic supply problem. Dots are the mean of one million trials of number of chromosomes needed to have all the schemas of order $1$ in the population with variance added versus $\ell$. The maximal variance is $0.0019$. The dashed line shows the curve fitting result.

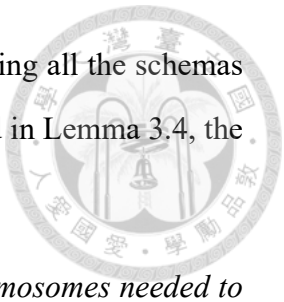
The minimal number of chromosomes to have all the schemas of order $1$ in the population is counted. Experiment result is shown in Figure 3.2. The dots are the mean over one million tests on the number of chromosome needed to have all the schemas of order $1$ with variance added with different $\ell$ in $\log$ scale. The dashed line shows the curve fitting result, which shows that Theorem 3.5 agrees with the result.

### 3.3.2 Lower bound

In the previous subsection, the expected number of chromosomes with length of $\ell$ for binary-encoded GAs to satisfy having all the schemas of order $1$ in the population is derived. However, to guarantee a success probability of $1 - \frac{1}{\ell}$, the minimal number of chromosomes, rather than the expected number of chromosomes, that satisfies the condition is needed. Because of that, in this subsection, the lower bound on the number of


doi:10.6342/NTU201902839

chromosomes with length of $\ell$ for binary-encoded GAs to satisfy having all the schemas of order 1 in the population is derived. Based on Equation 3.1 derived in Lemma 3.4, the following theorem hold:



**Theorem 3.6.** *For positive constant $\alpha$, the minimual number of chromosomes needed to guarantee a success probability of $1-\Theta\left(\ell^{-\alpha}\right)$ for binary-encoded GAs using an omnipotent oracle with chromosomes with length of $\ell$ is $\Theta\left(\alpha\log\ell\right)$.*

*Proof.* Suppose $\exists\ \ell_0>0$ such that $\forall\ell\geq\ell_0$, the success probability is bounded between $1-\frac{\eta^+}{\ell^\alpha}$ and $1-\frac{\eta^-}{\ell^\alpha}$, where $\eta^-\geq\eta^+$ are non-negative constants. From Equation 3.1, $\forall\ \beta\geq 2$, $\Pr\{\text{fail to collect all schemas of order 1 with a population with size of }\beta\}=1-\left(1-2^{-(\beta-1)}\right)^\ell$. Thus, $\forall\ \ell\geq 2,\alpha>0$, $\Pr\{\text{fail to collect all schemas of order 1 with a population size of }\lceil(\alpha+1)\log_2\ell-\log_2\eta^-\rceil\}\leq 0-\left(1-\frac{\eta^-}{\ell^{1+\alpha}}\right)^\ell$. Estimate the upper bound on $1-\left(1-\frac{\eta^-}{\ell^{1+\alpha}}\right)^\ell$. The upper bound would be

$$1-\left(1-\frac{\eta^-}{\ell^{1+\alpha}}\right)^\ell\leq 1-\left(1-\frac{\eta^-}{\ell^{1+\alpha}}\ell\right)\leq\frac{\eta^-}{\ell^\alpha}.$$

Thus, with $\lceil\alpha\log_2\ell-\log_2\eta^-\rceil$ chromosomes, success probability would greater than or equal to $1-\frac{\eta^-}{\ell^\alpha}$. Next, consider the probability with a population size of $\lfloor(\alpha+1)\log_2\ell-\log_2\eta^+-1\rfloor$, which is greater than or equal to $1-\left(1-\frac{2\eta^+}{\ell^{1+\alpha}}\right)^\ell$.

For $\ell\geq\max\left(2,\left(\frac{1}{2\eta^+}\right)^{\frac{-1}{1+\alpha}}\right)$, to lower bound the value, let $L\left(x\right)=\left(1-x\right)^\ell$. Since $\forall\ x\leq 1$, $\frac{d^2L(x)}{dx^2}\geq 0$, $L\left(x\right)\leq 1+x\frac{dL(x)}{dx}$. Because $1-\left(1-\frac{\eta^+}{\ell^{1+\alpha}}\right)^\ell=1-L\left(\frac{2\eta^+}{\ell^{1+\alpha}}\right)$ and $\frac{2\eta^+}{\ell^{1+\alpha}}\leq 1$,

$$\begin{aligned}
&1-\left(1-\frac{2\eta^+}{\ell^{1+\alpha}}\right)^\ell\\
\geq&1-\left(1-\ell\left(1-\frac{2\eta^+}{\ell^{1+\alpha}}\right)^{\ell-1}\frac{2\eta^+}{\ell^{1+\alpha}}\right)\\
=&\frac{2\eta^+}{\ell^\alpha}\left(1-\frac{2\eta^+}{\ell^{1+\alpha}}\right)^{\ell-1}\\
\geq&\frac{2\eta^+}{\ell^\alpha}\left(1-(\ell-1)\frac{2\eta^+}{\ell^{1+\alpha}}\right)\\
=&\frac{2\eta^+}{\ell^\alpha}-(\ell-1)\left(2\eta^+\right)^2\ell^{-2\alpha}
\end{aligned}$$


doi:10.6342/NTU201902839

$$
\begin{aligned}
&\geq \frac{2\eta^+}{\ell^\alpha} - \left(2\eta^+\right)^2 \ell^{-2\alpha} \\
&= \frac{2\eta^+}{\ell^\alpha}\left(1 - \frac{2\eta^+}{\ell^\alpha}\right) \\
&\geq \frac{\eta^+}{\ell^\alpha}.
\end{aligned}
$$

Because $\forall\ \ell \geq \max\left(\ell_0, 2, \left(\frac{1}{2\eta^+}\right)^{\frac{-1}{1+\alpha}}\right), \alpha > 0$, the success probability can be bounded between $1 - \frac{\eta^+}{\ell^\alpha}$ and $1 - \frac{\eta^-}{\ell^\alpha}$ with the the number of chromosomes required are bounded between $\lceil (\alpha+1)\log_2 \ell - \log_2 \eta^- \rceil$ and $\lfloor (\alpha+1)\log_2 \ell - \log_2 \eta^+ - 1 \rfloor$, the minimual number of chromosomes needed to guarantee a success probability of $1 - \Theta\left(\ell^{-\alpha}\right)$ for binary-encoded GAs using an omnipotent oracle with chromosomes with length of $\ell$ is $\Theta\left(\alpha \log \ell\right)$. □

To verify the derivation, success probability with population size of $\lceil 2\log_2 \ell \rceil + 2$ and $\lfloor 2\log_2 \ell \rfloor - 1$ are tested on $\ell = 2^i$ for $i = 1$ to $15$, and the results are shown in Figure 3.3 (Liao *et al.*, 2019). Each bit is determined using Bernoulli distribution with $\frac{1}{2}$ being the probability of success. It can be shown that Theorem 3.6 agrees with the result, and Lemma 3.3 is a close approximation.

Finally, based on Theorem 3.6, the lower bound of the supply is derived:

**Theorem 3.7.** *For positive constant $\alpha$, the population size for binary-encoded selecto-recombinative GAs using OM to guarantee a success rate of $1 - \Theta\left(\ell^{-\alpha}\right)$ is $\Omega\left(\alpha \log \ell\right)$.*

*Proof.* The result can be inferred from Theorem 3.6 and the fact that having all the schemas of order $1$ is a minimal requirement to solve the problem. □


doi:10.6342/NTU201902839

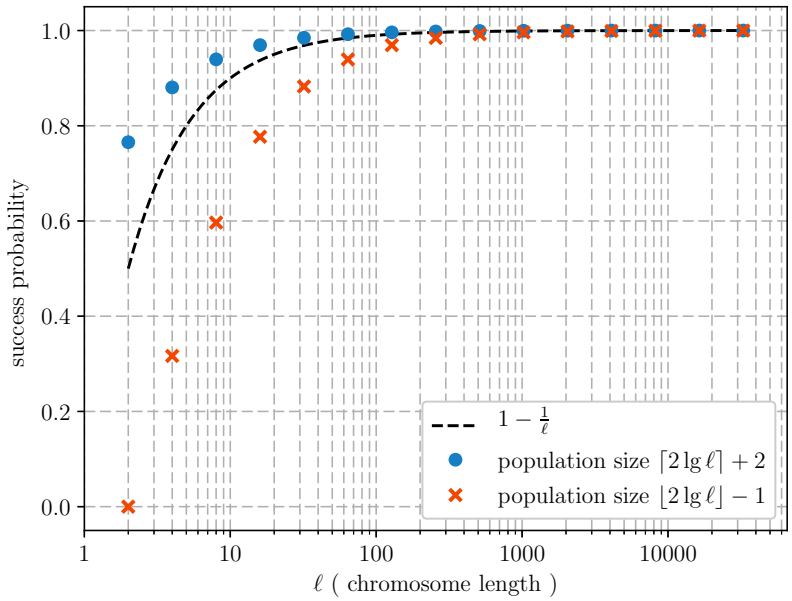



Figure 3.3: Experiment on the probability of successfully solving the basic supply problem. Dots are the results of one million trials on the success probability of whether $\lceil 2\log_2 \ell \rceil + 2$ chromosomes can have all the schemas of order $1$ in the population with $95\%$ confidence interval added versus $\log \ell$. Crosses are the results of one million trials on the success probability of whether $\lfloor 2\log_2 \ell \rfloor - 1$ chromosomes can have all the schemas of order $1$ in the population with $95\%$ confidence interval added. Maximum one-sided range for population size $\lceil 2\log_2 \ell \rceil + 2$ and $\lfloor 2\log_2 \ell \rfloor - 1$ are $0.00083$ and $0.00096$ respectively. The dashed line shows the success probability of $1 - \frac{1}{\ell}$ for each $\ell$.


doi:10.6342/NTU201902839

# Chapter 4

# $c$-composite Oracles and Supply

In Chapter 3, bounds on the most basic supply issue are derived. The issue is equivalent to the success probability when solving with an omnipotent oracle. However, an omnipotent oracle is more powerful than binary-encoded selecto-recombinative GAs using OM with an oracle because of the variety of operators an omnipotent oracle holds. Thus, in this chapter, oracles only using OM as the operator are focused. Even so, the behavior of the oracle is still hard to estimate because it can be very problem dependent. Thus, we put restrictions on the oracle and derive the bounds on the population size for using a $c$-composite oracle, defined as follows.

**Definition 4.1.** ***composite****: A composite is a composition of one or many non-overlapping masks.*

**Definition 4.2.** ***$c$-composite oracle****: For a constant nonnegative integer $c$, a $c$-composite oracle is an oracle designed for optimization problems such that if the oracle does not return* FAIL*, for each output sequence of the oracle, all the masks in the sequence where no two masks are identical can form at most $c$ composites in the most compact way.*

Figure 4.1 gives an example of a $c$-composite oracle optimizing a 6-bits problem. The masks can form 3 composites in the most compact way. Since the number $c$ restricts the maximal number of composites can be formed in the process, a 3-composite oracle can optimize this problem with the given population, whereas a 2-composite oracle can not.


doi:10.6342/NTU201902839

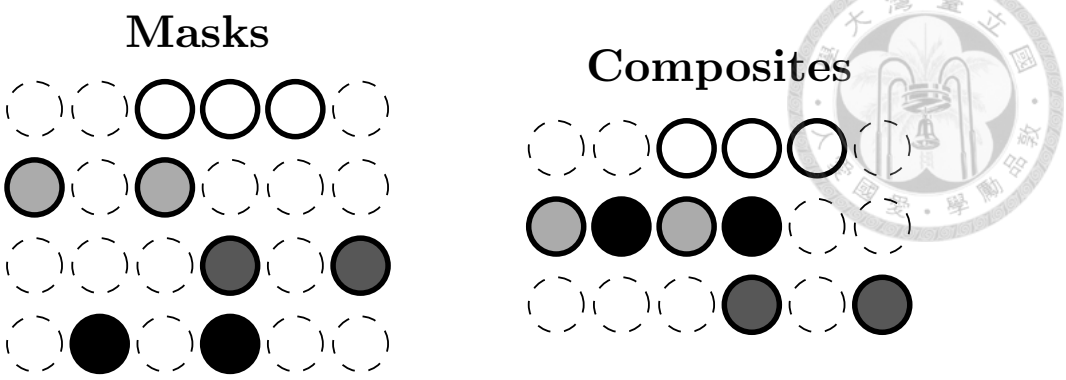



Figure 4.1: Example for a $c$-composite oracle. Suppose optimizing a 6-bits problem with a given population and the masks of the process yeilding minimum number of composites in the most compact way are given. The solid circles represent the masks, and the dotted line circles are the unchanged bits. Color of the circle indicates that it belongs to the same mask with the ones with the same color. The masks can form 3 composites in the most compact way. Thus, a 3-composite oracle can optimize this problem with the given population, whereas a 2-composite oracle can not.

## 4.1 Investigation on 1-composite oracles

In this section, a special case of $c$-composite oracle, 1-composite oracle, is focused on. Firstly, based on the derivation done by Tung (2015), a tight bound on the population size for binary-encoded selecto-recombinative GAs using OM with a 1-composite oracle with the longest mask with length of $\kappa$. Then, based on a qualified scheme, a general upper bound for binary-encoded selecto-recombinative GAs using OM with a 1-composite oracle is derived along with some investigation on the tightness of the bound.

### 4.1.1 Result with prior knowledge on the masks

Inferred from Definition 4.2, when using a 1-composite oracle, each locus can be changed at most once. This inference implies a 1-composite oracle on OM forces the donors to donate bits the same as the ones of the optimal solution only or there must be some locus being changed more than once. Thus, the requirement of recombining a certain chromo-


doi:10.6342/NTU201902839

some to globabl optimum with the aid of a 1-composite oracle with a success probability of $1-\Theta\left(\ell^{-\alpha}\right)$ is the same as the requirement of having the same rate of having one certain schema in every segment with respect to the partition of the process in which the chromosome are recombined.

Derived by Tung (2015), for a problem composed of $m$ subproblems with size $k$, to guarantee a constant success rate, the population size should be $\mathrm{O}\left(2^k \ln m\right)$. However, a success probability of $1-\Theta\left(\ell^{-\alpha}\right)$ is not a constant one. Thus, new bounds need to be derived.

**Theorem 4.3.** *Population size to have a success probability of* $1-\Theta\left(\ell^{-\alpha}\right)$ *for binary-encoded selecto-recombinative GAs using OM with a* 1*-composite oracle with the largest mask with size* $\kappa$ *is* $\Theta\left(2^{\kappa}\left(1+\alpha\right)\log\ell\right)$*, where* $\ell$ *is the chromosome length and* $\alpha$ *is a positive constant.*

*Proof.* Suppose for some $\ell_0 \geq 0$, $\forall\ \ell \geq \ell_0$, the required success probability is bounded between $1-\frac{\eta^+}{\ell^\alpha}$ and $1-\frac{\eta^-}{\ell^\alpha}$, where $\eta^- \geq \eta^+$ are non-negative constants. First, define $n^-, n^-, n^+$ as follows.

$n^+$ : minimum population size needed to guarantee having the designated schema out of $2^\kappa$ patterns for $\ell$ distinct building blocks with a success probability of $1-\frac{\eta^+}{\ell^\alpha}$.

$n$ : minimum population size needed to guarantee a success probability bounded between $1-\frac{\eta^+}{\ell^\alpha}$ and $1-\frac{\eta^-}{\ell^\alpha}$ for solving an arbitrary chromosome for a binary-encoded selecto-recombinative GA using OM with an 1-composite oracle with chromosome with length of $\ell$ and with the largest mask with size $\kappa$.

$n^-$ : minimum population size needed to guarantee having the designated schema out of $2^\kappa$ patterns with a success probability of $1-\frac{\eta^-}{\ell^\alpha}$.

Based on the above descriptions, the following equations hold:

$$\left(1-\left(1-2^{-\kappa}\right)^{n^+}\right)^{\ell} = 1-\frac{\eta^+}{\ell^\alpha},$$


doi:10.6342/NTU201902839

$$1-\frac{\eta^-}{\ell^\alpha} \leq \prod_{i=1}^{\tau}\left(1-\left(1-2^{-\kappa_i}\right)^n\right) \leq 1-\frac{\eta^+}{\ell^\alpha},$$

$$\text{and} \qquad \left(1-\left(1-2^{-\kappa}\right)^{n^-}\right) = 1-\frac{\eta^-}{\ell^\alpha},$$

where $\tau$ is the number of masks needed to be collected, and $\kappa_i$s represent the masks needed to be collected where $\forall\ i \in \{1,2,\ldots,\tau\}, \kappa_i > 0$, $\sum_{i=1}^{\tau}\kappa_i \leq \ell$, and $\max_{i\in\{1,2,\ldots,\tau\}}\kappa_i = \kappa$. Since $\forall\ \zeta \in \mathbb{N}$,

$$\left(1-\left(1-2^{-\kappa}\right)^{\zeta}\right) \geq \prod_{i=1}^{\tau}\left(1-\left(1-2^{-\kappa_i}\right)^{\zeta}\right),$$

and

$$\begin{aligned}\prod_{i=1}^{\tau}\left(1-\left(1-2^{-\kappa_i}\right)^{\zeta}\right) &\geq \prod_{i=1}^{\tau}\left(1-\left(1-2^{-\kappa}\right)^{\zeta}\right)\\ &\geq \left(1-\left(1-2^{-\kappa}\right)^{\zeta}\right)^{\ell},\end{aligned}$$

and $\prod_{i=1}^{\tau}\left(1-(1-2^{-\kappa_i})^{\zeta}\right)$ is increasing with $\zeta$, $n^+ \geq n \geq n^-$. For computation simplicity, we focus on finding the asymptotic order of $n^+$ and $n^-$. Based on the results derived by Tung (2015),

$$n^+ = -\ln\left(1-\left(1-\frac{\eta^+}{\ell^\alpha}\right)^{\frac{1}{\ell}}\right)\Theta\left(2^\kappa\right), \tag{4.1}$$

while

$$n^- = -\ln\left(1-\left(1-\frac{\eta^-}{\ell^\alpha}\right)^{1}\right)\Theta\left(2^\kappa\right). \tag{4.2}$$

Equation 4.2 can be simplified as $-\ln\left(\frac{\eta^-}{\ell^\alpha}\right)\Theta\left(2^\kappa\right)$ and is hence $\Theta\left(2^\kappa\alpha\log\ell\right)$. Thus, we only need to focus on Equation 4.1. Applying the same technique used by Tung (2015), consider convex function $g(x) = (1-\gamma)^x$, where $\gamma > 0, x \in (0,1]$. Based on the two inequalities:

$$(1-x)\,g(0) + xg(1) \geq g(x) \geq g(0) + x\frac{dg(0)}{dx}.$$


doi:10.6342/NTU201902839

By substituting $x$ and $\gamma$ with $\frac{1}{\ell}$ and $\frac{\eta^+}{\ell^\alpha}$, respectively, and with simple arithmetic calculations, $\forall\ \ell > (\eta^+)^{\frac{1}{\alpha}}$,

$$\frac{\eta^+}{\ell\left(\ell^\alpha - \eta^+\right)} = \frac{1}{\ell}\left(\frac{1}{1-\frac{\eta^+}{\ell^\alpha}} - 1\right) \geq \frac{1}{\ell}\ln\frac{1}{1-\frac{\eta^+}{\ell^\alpha}} \geq 1 - \left(1-\frac{\eta^+}{\ell^\alpha}\right)^{\frac{1}{\ell}} \geq \frac{\eta^+}{\ell^{1+\alpha}}$$

$$\Rightarrow -\ln\left(\frac{\eta^+}{\ell^{1+\alpha}}\right) \geq -\ln\left(1-\left(1-\frac{\eta}{\ell}\right)^{\frac{1}{\ell}}\right) \geq -\ln\frac{\eta^+}{\ell\left(\ell^\alpha - \eta^+\right)}.$$

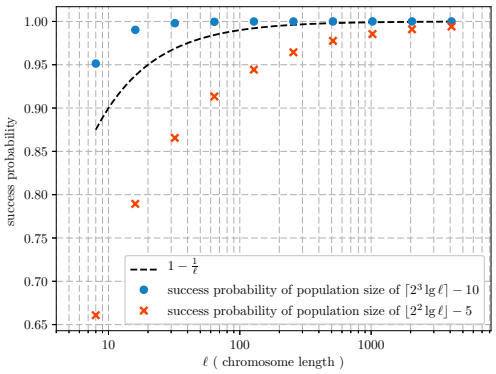
Thus, Equation 4.1 is $\Theta\left(2^\kappa \alpha \log \ell\right)$. Since both the upper and lower bounds are $\Theta\left(2^\kappa \alpha \log \ell\right)$, population size required for a binary-encoded GA using OM with a 1-composite oracle with the largest mask with size $\kappa$ to have a success probability of $1 - \Theta\left(\ell^{-\alpha}\right)$, where $\ell$ is the chromosome length and $\alpha$ is a positive constant, is $\Theta\left(2^\kappa \alpha \log \ell\right)$. □

To verify the derivation, the success probability of solving a randomized chromosome with length of $\ell$ with the longest segment with length of $\kappa$ for a binary-encoded GA using OM with a 1-composite oracle is simulated (Liao *et al.*, 2019). The experiments are conducted on $\ell = 2^i$ for $i = 3$ to 12 for $\kappa = 2$ and 5 respectively. Also, each bit of a chromosome is assigned with Bernoulli distribution with $\frac{1}{2}$ being the probability of success. Partition is also determined at random. Without loss of generality, assume the optimal solution is the chromosome with all 1s. Success probability with supply $\lceil 2^{\kappa+1} \ln \ell \rceil - 10$ or $\lceil 2^{\kappa+1} \ln \ell \rceil - 100$ and $\lfloor 2^\kappa \ln \ell \rfloor - 5$ or $\lfloor 2^\kappa \ln \ell \rfloor - 50$ are estimated using the average over one million independent trials with $95\%$ confidence interval using normal approximation. Experiment result is shown in Figure 4.2.

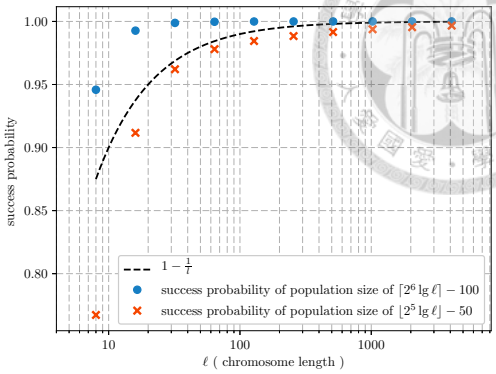
### 4.1.2 General upper bound

Theorem 4.3 gives a tight bound on the population size when the size of the largest mask is known, but how can one know the size of the largest mask when being guided by a 1-composite oracle? Consider the situation that in the population there exist two chromosomes that have all the schemas of order 1. Since the optimal solution is the chromosome with the best fitness, with the aid of a 1-composite oracle the global optimum can be


doi:10.6342/NTU201902839




(a) Results for $\kappa = 2$




(b) Results for $\kappa = 5$

Figure 4.2: Experiments on the success probability of 1-composite oracles. 4.2(a) and 4.2(b) show the experiment results for the largest mask with size $\kappa = 2$ and 5 respectively. In both Figures 4.2(a) and 4.2(b), dots are the averages over one million independent trials of whether $\lceil 2^{\kappa+1} \lg \ell \rceil - 10$ or $\lceil 2^{\kappa+1} \lg \ell \rceil - 100$ chromosomes are enough to solve a chromosome with length of $\ell$ with the largest mask with size $\kappa$ using OM with a 1-composite oracle for $\kappa = 2$ and 5 respectively with $95\%$ confidence interval added versus $\ell$, whereas crosses are the ones of $\lfloor 2^{\kappa} \lg \ell \rfloor - 5$ or $\lfloor 2^{\kappa} \lg \ell \rfloor - 50$ chromosomes for $\kappa = 2$ and 5 respectively with $95\%$ confidence interval added versus $\ell$. Maximum one-sided intervals are $0.0009$ and $0.0008$ respectively. The dashed line shows the probability of $1 - \frac{1}{\ell}$ for each $\ell$.

reached by one donation. Thus, if after the initialization there exist two chromosomes that have all the schemas of order 1, a 1-composite can solve the optimization problem by proposing a scheme using one donation. With no further assumptions on the fitness function, the upper bound of the population size derived in this scheme generally fits. However, Theorem partition states that for 1-composite oracles, the maximum number of bits involved in one OM donation will exponentially affect the population size required. Thus, to know more about the underlying distribution of the minimum of the maximal number of bits used in one donation under the scheme, the distribution of the maximal


doi:10.6342/NTU201902839

number of schemas of order $1$ that a chromosome in the population and the global optimum shares given the population size needs to be examined.

**Theorem 4.4.** *For binary-encoded selecto-recombinative GAs, the minimal population size required for the maximal number of schemas of order* $1$ *that a chromosome in the population and the global optimum shares being greater than or equal to* $\lfloor (1-\rho)\,\ell \rfloor$*, where* $\ell$ *is the chromosome length and* $0 < \rho < \frac{1}{2}$ *is a constant, with a probability of* $1 - \frac{\eta}{\ell^\alpha}$*, where* $\alpha$ *and* $\eta$ *are positive constants, is* $O\left(\left(\ell+1\right)^2 e^{\ell(\rho \log_2 2\rho + (1-\rho)\log_2 2(1-\rho))}\left(\log \frac{1}{\eta} + \alpha \log \ell\right)\right)$.

*Proof.* Let $n$ be the population size and $\tau_{\ell,n}$ be the the maximal number of schemas of order $1$ that a chromosome of length $\ell$ in the population of size $n$ and the global optimum share, then

$$
\begin{aligned}
& \Pr\left\{\tau_{\ell,n} \geq \lfloor (1-\rho)\,\ell \rfloor\right\} \geq 1 - \frac{\eta}{\ell^\alpha} \\
\iff & \Pr\left\{\tau_{\ell,n} < \lfloor (1-\rho)\,\ell \rfloor\right\} \leq \frac{\eta}{\ell^\alpha} \\
\iff & \Pr\left\{\tau_{\ell,1} < \lfloor (1-\rho)\,\ell \rfloor\right\}^n \leq \frac{\eta}{\ell^\alpha} \\
\iff & \left(1 - \Pr\left\{\tau_{\ell,1} \geq \lfloor (1-\rho)\,\ell \rfloor\right\}\right)^n \leq \frac{\eta}{\ell^\alpha} \\
\impliedby & \left(1 - \Pr\left\{\tau_{\ell,1} = \lfloor (1-\rho)\,\ell \rfloor\right\}\right)^n \leq \frac{\eta}{\ell^\alpha} \\
\impliedby & \left(1 - \frac{1}{\left(\ell+1\right)^2} e^{-\ell(\rho \log_2 2\rho + (1-\rho)\log_2 2(\rho))}\right)^n \leq \frac{\eta}{\ell^\alpha},
\end{aligned}
$$

where $\Pr\left\{\tau_{\ell,1} = \lfloor (1-\rho)\,\ell \rfloor\right\} \leq \frac{1}{(\ell+1)^2} e^{-\ell(\rho \log_2 2\rho + (1-\rho)\log_2 2(\rho))}$ is proved by Cover and Thomas (2006). Since $\forall\ x > 1$, $\left(1 - \frac{1}{x}\right)^x < \frac{1}{e}$, $n = \left(\ell+1\right)^2 e^{\ell(\rho \log_2 2\rho + (1-\rho)\log_2 2(1-\rho))}\left(\ln \frac{1}{\eta} + \alpha \ln \ell\right)$ is sufficient to fulfill the last inequality and $\Pr\left\{\tau_{\ell,1} < \lfloor (1-\rho)\,\ell \rfloor\right\}^n$ decreases as $n$ increases, for binary-encoded selecto-recombinative GAs, the minimal population size required for the maximal number of schemas of order $1$ that a chromosome in the population and the global optimum shares being greater than or equal to $\lfloor (1-\rho)\,\ell \rfloor$, where $\ell$ is the chromosome length and $0 < \rho < \frac{1}{2}$ is a constant, with a probability of $1 - \frac{\eta}{\ell^\alpha}$, where $\alpha$ and $\eta$ are positive constants, is $\mathrm{O}\left(\left(\ell+1\right)^2 e^{\ell(\rho \log_2 2\rho + (1-\rho)\log_2 2(\rho))}\left(\log \frac{1}{\eta} + \alpha \log \ell\right)\right)$. □

Based on Theorems 4.3 and 4.4, a upper bound for selecto-recombinative GAs using


doi:10.6342/NTU201902839

OM with a 1-composite oracle can be derived.

**Theorem 4.5.** *The minimal population size required for binary-encoded selecto-recombinative GAs using OM with an oracle to have a success probability of* $1-\Theta\left(\ell^{-\alpha}\right)$*, where* $\ell$ *is the chromsome length and* $\alpha$ *is a positive constant is* $O\left(2^{\frac{11\ell}{40}}\alpha\log\ell\right)$.

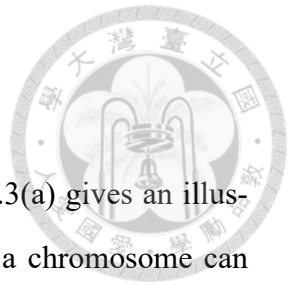
*Proof.* By Theorems 4.3 and 4.4, $\mathrm{O}\left(\left(\ell+1\right)^{2}e^{\ell(\rho\log_2 2\rho+(1-\rho)\log_2 2(1-\rho))}\alpha\ln\ell\right)$ $+\Theta\left(2^{\lceil\rho\ell\rceil}\alpha\log\ell\right)$ chromosomes, where $0<\rho<\frac{1}{2}$ is a constant, are sufficient of have a chromosome in the population with number of schemas of order 1 that shares with the global optimum is greater than or equal to $\lfloor(1-\rho)\,\ell\rfloor$ and another chromosome that can recombine with the previous chromosome to reach the global optimum with probability of $1-\Theta\left(\ell^{-\alpha}\right)$. Since $e^{\ell\left(\frac{11}{40}\log_2\frac{22}{40}+\left(1-\frac{11}{40}\right)\log_2 2\left(1-\frac{11}{40}\right)\right)}<2^{\frac{11}{40}\ell}$ and $\mathrm{O}\left(\left(\ell+1\right)^{2}e^{\ell(\rho\log_2 2\rho+(1-\rho)\log_2 2(1-\rho))}\alpha\ln\ell\right)$ increases as $\rho$ increases, whereas $\Theta\left(2^{\lceil\rho\ell\rceil}\alpha\log\ell\right)$ decreases as $\rho$ increases, $\rho$ such that balances between the asymptotic orders of $\mathrm{O}\left(\left(\ell+1\right)^{2}e^{\ell(\rho\log_2 2\rho+(1-\rho)\log_2 2(1-\rho))}\alpha\ln\ell\right)$ and $\Theta\left(2^{\lceil\rho\ell\rceil}\alpha\log\ell\right)$ would be smaller than $\frac{11}{40}$, meaning the minimal population size required for binary-encoded selecto-recombinative GAs using OM with a 1-composite oracle to have a success probability of $1-\Theta\left(\ell^{-\alpha}\right)$, where $\ell$ is the chromsome length and $\alpha$ is a positive constant is $\mathrm{O}\left(2^{\frac{11\ell}{40}}\alpha\log\ell\right)$. Since 1-compoiste oracle is a restricted oracle, the upper bound of population size derived for 1-composite oracle is also an upper bound for an oracle under the same condition and the theorem statement holds.

□

Theorem 4.5 gives an upper bound of supply. Since the bound can be applied to any problem in the worst case scenario and is still exponential to the chromosome length, the tightness of the bound can be challenged. An example type of problem is presented here to show that Theorem 4.5 is fairly tight. Consider a maximization problem with its fitness function being a special case of deceptive-trap function (Deb and Goldberg, 1993), which can be written as follows.

$$f^{\mathrm{TRAP}}\left(\boldsymbol{x}\right)=\begin{cases}1 & \text{if } \sum_{i=0}^{\ell-1}\mathbb{1}\left\{x_{[i]}=1\right\}=\ell\\ \frac{0.9\left(\ell-1-\sum_{i=0}^{\ell-1}\mathbb{1}\left\{x_{[i]}=1\right\}\right)}{\ell-1} & \text{else}\end{cases},$$


doi:10.6342/NTU201902839



where $\boldsymbol{x}$ is a chromosome, and $\mathbb{1}$ is the indicator function. Figure 4.3(a) gives an illustration for the fitness function. The function is designed such that a chromosome can only lower the hamming distance to the optimal solution by having all the non-optimal genes replaced in one donation. This property ensures that given a population, either a 1-composite oracle can optimize in one donation or selecto-recombinative GAs using OM can not reach the optimal solution using this population.

On the other hand, there are cases such as maximizing onemax problem or maximizing needle in a haystack problem that a 1-composite oracle with a population of size logarithm of the chromosome length is sufficient to guarantee a success probability required. Let $\boldsymbol{x}$ be a chromosome, the fitness functions of onemax problem or needle in a haystack problem are:

$$f^{\text{ONEMAX}}(\boldsymbol{x}) = \sum_{i=0}^{\ell-1} \mathbb{1}\left\{x_{[i]} = 1\right\},$$
$$f^{\text{NEEDLE}}(\boldsymbol{x}) = \mathbb{1}\left\{\boldsymbol{x} = \boldsymbol{x}^*\right\},$$

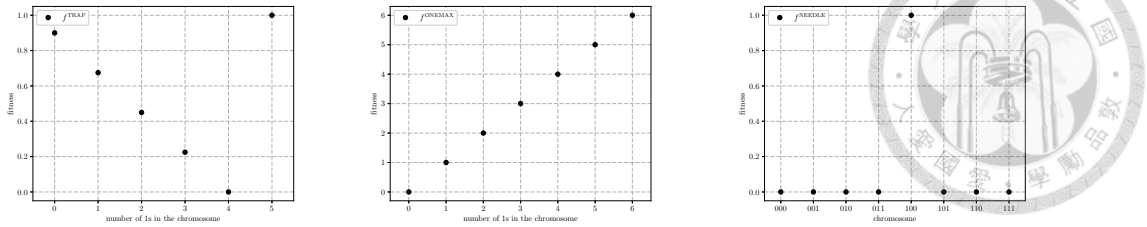
where $\mathbb{1}$ is the indicator function, and $\boldsymbol{x}^*$ is the global optimum. Onemax problem can be viewed as maximizing the number of 1s in the chromosome whereas needle in a haystack problem aims to find the optimum with no information given from the problem. Figures 4.3(b) and 4.3(c) gives an illustration for both problems.

Based on the properties of the problems, the following theorems hold:

**Theorem 4.6.** *The population size required for binary-encoded selecto-recombinative GAs using OM with an oracle to solve a one max problem with a success probability of* $1 - \Theta\left(\ell^{-\alpha}\right)$*, where* $\ell$ *is the chromosome length and* $\alpha$ *is a positive constant, is* $\Theta\left(\alpha \log \ell\right)$.

*Proof.* Since the lower bound is derived in Theorem 3.7, only the upper bound need to be derived. Consider the following strategy: randomly pick a chromosome in the population and change each bit to 1 one bit at a time using OM until the chromosome reaches the global optimum. Since increasing the number of 1s will not worsen the fitness of


doi:10.6342/NTU201902839




(a) Deceptive-trap function (b) Onemax problem (c) Needle in a haystack problem

Figure 4.3: Illustrarion of the landscapes that 1-composite oracles can solve. Figure 4.3(a) is a 5-bit deceptive-trap function, Figure 4.3(b) is a 6-bit onemax function, and Figure 4.3(c) is a 3-bit needle in a haystack problem with the optimal solution $\boldsymbol{x}^* = 100$.

the chromosome, each move will be accepted. Also, this process can be done using a 1-composite oracle with maximal size of the mask being one. Therefore, by Theorem 4.3, $\Theta(\alpha \log \ell)$ chromosomes are sufficient for binary-encoded selecto-recombinative GAs using OM with a 1-composite oracle to have a success probability of $1 - \Theta(\ell^{-\alpha})$. Since a 1-composite oracle is a more restricted oracle, the oppulation size for binary-encoded selecto-recombinative GAs using OM with an oracle to solve a one max problem with a success probability of $1 - \Theta(\ell^{-\alpha})$ is $\mathrm{O}(\alpha \log \ell)$. Thus, the population size required for binary-encoded selecto-recombinative GAs using OM with an oracle to solve a one max problem with a success probability of $1 - \Theta(\ell^{-\alpha})$, where $\ell$ is the chromosome length and $\alpha$ is a positive constant, is $\Theta(\alpha \log \ell)$. □

**Theorem 4.7.** *The population size required for binary-encoded selecto-recombinative GAs using OM with an oracle to solve a needle in a haystack problem sith a success probability of* $1 - \Theta(\ell^{-\alpha})$*, where* $\ell$ *is the chromosome length and* $\alpha$ *is a positive constant.*

*Proof.* Since the lower bound is derived in Theorem 3.7, only the upper bound need to be derived. Consider the following strategy: randomly pick a chromosome in the population and change each bit to be the same the one of the global optimum one bit at a time using OM until the chromosome reaches the global optimum. Since increasing the number of 1s will not worsen the fitness of the chromosome, each move will be accepted. Also, this process can be done using a 1-composite oracle with maximal size of the mask


doi:10.6342/NTU201902839

being one. Therefore, by Theorem 4.3, $\Theta(\alpha \log \ell)$ chromosomes are sufficient for binary-encoded selecto-recombinative GAs using OM with a 1-composite oracle to have a success probability of $1 - \Theta(\ell^{-\alpha})$. Since a 1-composite oracle is a more restricted oracle, the oppulation size for binary-encoded selecto-recombinative GAs using OM with an oracle to solve a needle in a haystack problem with a success probability of $1 - \Theta(\ell^{-\alpha})$ is $O(\alpha \log \ell)$. Thus, the population size required for binary-encoded selecto-recombinative GAs using OM with an oracle to solve a needle in a haystack problem with a success probability of $1 - \Theta(\ell^{-\alpha})$, where $\ell$ is the chromosome length and $\alpha$ is a positive constant, is $\Theta(\alpha \log \ell)$. □

## 4.2 General case for c-composite oracles

In the previous section, bounds on the population when using a 1-composite oracle, which is a specific case, are derived. In this section, the general cases are discussed. For a 1-composite oracle, the oracle is forced to have donors only donate bits the same as the ones of optimal solution. However, in Theorem 4.3, the goal is equivalent to gathering designated non-overlapping schemas parallelly with maximum length $\kappa$. This equivalence is useful when discussing the general case.

**Theorem 4.8.** *For positive constant $\alpha$, the population size for solving a chromosome with length of $\ell$ for binary-encoded GAs using OM with a $c$-composite oracle with the largest mask with size $\kappa$ with a success probability of $1 - \Omega(\ell^{-\alpha})$ is $\Omega(2^{\kappa}\alpha \log \ell)$.*

*Proof.* Suppose $\forall\ \ell > \ell_0$, where $\ell_1$ is a constant, the required success probability is bounded below by $1 - \frac{\eta_1}{\ell^{\alpha}}$, where $\eta_1$ is a positive constant. By Theorem 4.3, the population size required for gathering designated schemas parallelly with maximum length of $\kappa$ and total length of all the schemas is less than or equal to $\ell$ with a success probability bounded between $1 - \frac{\eta_2^+}{\ell^{\alpha}}$ and $1 - \frac{\eta_2^-}{\ell^{\alpha}}$, where $\eta_2^- > \eta_2^+$ are positive constants, is $\Theta(2^{\kappa}\alpha \log \ell)$. Because of that, we can have a lower bound of order $\Omega(2^{\kappa}\alpha \log \ell)$, since the total requirement for a $c$-composite oracle is at most having designated schemas parallelly with maximum length $\kappa$ and total length of all the schemas is less than or equal to


doi:10.6342/NTU201902839

$\ell$ on $c$ different chromosomes instead of 1 chromosome. □

**Theorem 4.9.** *For positive constants $\alpha_1$ and $\alpha_2$, the population size for solving a chromosome with length of $\ell$ for binary-encoded GAs using OM with a $c$-composite oracle, where $c$ is $O(\ell^{\alpha_2})$, with the largest mask with size $\kappa$ with a success probability of $1 - O(\ell^{-\alpha_1})$ is $O(c2^{\kappa}(\alpha_1 + \alpha_2)\log\ell)$.*

*Proof.* Suppose $\forall\ \ell > \ell_0$, where $\ell_1$ is a constant, the required success probability is bounded above by $1 - \frac{\eta_1}{\ell^{\alpha}}$, where $\eta_1$ is a positive constant and $c$ is upper bounded by $\eta_2\ell^{\alpha_2}$, where $\eta_2$ is a positive constant. By Theorem 4.3, the population size required for gathering designated schemas parallelly with maximum length of $\kappa$ and total length of all the schemas is less than or equal to $\ell$ with a success probability bounded above by $1 - \frac{\eta_3}{\ell^{\alpha_1+\alpha_2}}$, where $\eta_3$ is a positive constant, is $\mathrm{O}(2^{\kappa}(\alpha_1 + \alpha_2)\log\ell)$. Taking $\eta_3 = \frac{\eta_1}{\eta_2}$, by union bound, $\mathrm{O}(c2^{\kappa}(\alpha_1 + \alpha_2)\log\ell)$ chromosomes are enough to guarantee solving a chromosome with length of $\ell$ using a binary-encoded GA with OM with a $c$-composite oracle with longest segment with length of $\kappa$ with the success probability bounded above by $1 - \frac{\eta_1}{\ell^{\alpha}}$ . Thus, for positive constants $\alpha_1$ and $\alpha_2$, the population size for solving a chromosome with length of $\ell$ for binary-encoded GAs using OM with a $c$-composite oracle, where $c$ is $\mathrm{O}(\ell^{\alpha_2})$, with the largest mask with size $\kappa$ with a success probability of $1 - \mathrm{O}(\ell^{-\alpha_1})$ is $\mathrm{O}(c2^{\kappa}(\alpha_1 + \alpha_2)\log\ell)$. □

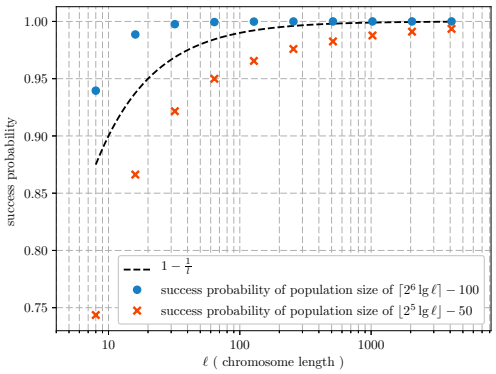
The asymptotic bounds derived in Theorems 4.8 and 4.9 do not match generally. However, when $c$ is constant, tight bound exists.

**Theorem 4.10.** *For positive constants $\alpha$ and $c$, the population size for solving a chromosome with length of $\ell$ for binary-encoded GAs using OM with a $c$-composite oracle with the largest mask with size $\kappa$ with a success probability of $1 - \Theta(\ell^{-\alpha})$ is $\Theta(2^{\kappa}\alpha\log\ell)$.*

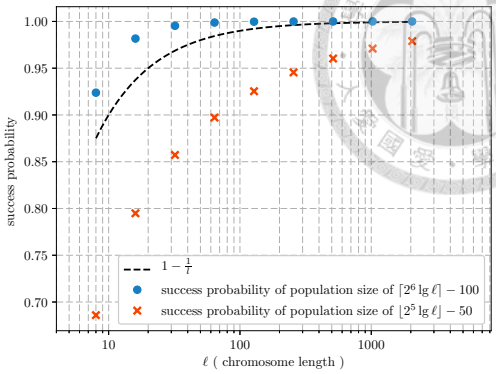
*Proof.* The result can be directly inferred from Theorems 4.8 and 4.9. □

To verify the derivations, the success probability of solving a random chromosome with length of $\ell$ with largest mask with size $\kappa = 5$ for a binary-encoded GA using OM with a $c$-composite oracle is simulated (Liao *et al.*, 2019). The experiments are conducted


doi:10.6342/NTU201902839




(a) Results for $c = 2$




(b) Results for $c = 5$

Figure 4.4: Experiments on the success probability of $c$-composite oracles. In both Figures 4.4(a) and 4.4(b), dots are the averages over one million independent trials of whether $\lceil 2^{\kappa+1} \lg \ell \rceil - 100$ chromosomes are enough to solve chromosome with length of $\ell$ with the largest mask size $\kappa = 5$ for a GA OM with a $c$-composite oracle with $95\%$ confidence interval added versus $\ell$ for $c = 2$ and $5$ respectively, whereas the crosses are the ones of $\lfloor 2^{\kappa} \lg \ell \rfloor - 50$ chromosomes with $95\%$ confidence interval added versus $\ell$. Maximum one-sided intervals are $0.0008$ and $0.0009$ respectively. Black dots show the success probability of $1 - \frac{1}{\ell}$ for each $\ell$.

on $\ell = 2^i$ for $i = 3$ to $11$ for $c = 2, 5$ respectively. Each bit of a chromosome is determined using Bernoulli distribution with $\frac{1}{2}$ being the probability of success. Partition is also determined at random. Without loss of generality, we assume the optimal solution is the chromosome with all 1s. Success probabilities with population size $\lceil 2^{\kappa+1} \lg \ell \rceil - 100$ and $\lfloor 2^{\kappa} \lg \ell \rfloor - 50$ are estimated using the average over one million independent trials. Experiment result is shown in Figure 4.4 and Theorem 4.10 agrees with it.



doi:10.6342/NTU201902839



doi:10.6342/NTU201902839

# Chapter 5

# Results with Problems with Ring Topologies

In Chapters 3 and 4, except Theorems 4.6 and 4.7, the derived bounds are all problem structure independent. Thus, in this chapter, bounds on the populaiton size in some special problem structures are derived.

## 5.1 Ring topology and reduction

This section focuses on problems with ring topology. To be more specific, only a small group of NK landscape problems defined by Pelikan *et al.* (2009) with step size equals to $k$ is focused. Let $\boldsymbol{x}$ be a chromosome and $\mathcal{F}_{\langle k+1\rangle}$ be a homogeneous overlapping FOS, and $\boldsymbol{x}_M$ be the part of $\boldsymbol{x}$ indicated by a mask $M$, the fitness function $f^{\text{ring}}$ is defined as follows.

$$f^{\text{ring}}(\boldsymbol{x}) = \sum_{M \in \mathcal{F}_{\langle k+1\rangle}} f_M^{\text{sub-ring}}(\boldsymbol{x}_M).$$

Figure 5.1 gives a illistration of the ring topology. Ring topology is chosen to be examined becacuse it is homogeneous. However, to make each bit really being homogeneous, no difference between whether one bit will be used in two sub-functions or in just one sub-function should exist. Thus, define the reduced problem with chromosomes with length


doi:10.6342/NTU201902839

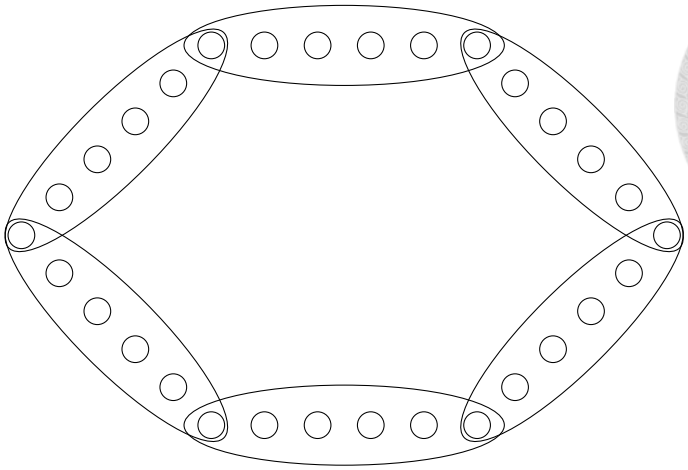


Figure 5.1: An illustration of the ring topology focused in this thesis with $\ell = 30$ and $k = 5$. Each ellipse represents a sub-function. Each circle represents a bit.

of $\frac{\ell}{k}$ and with fitness function $f^{\text{reduced-ring}}$:

$$f^{\text{reduced-ring}}(\boldsymbol{y}) = \sum_{i=0}^{\frac{\ell}{k}-1} f_i^{\text{sub-reduce}}\left(y_{[i]}, y_{[i]\oplus_{\frac{\ell}{k}}[1]}\right),$$

where $\boldsymbol{y} = \langle y_{[i]} \rangle$, in which $i \in \mathbb{G}_{\frac{\ell}{k}}$, $y_{[i]} = x_{[k\cdot i]\oplus[1]}$, and for maximation problems, $f_i^{\text{sub-reduce}}\left(y_{[i]}, y_{[i]\oplus_{\frac{\ell}{k}}[1]}\right)$ is

$$\max_{x_{[i\cdot k]\oplus[2]},\ldots,x_{[i\cdot k]\oplus[k]}} f_i^{\text{sub-ring}}\left(y_{[i]}, x_{[i\cdot k]\oplus[2]}, \ldots, x_{[i\cdot k]\oplus[k]}, y_{[i]\oplus_{\frac{\ell}{k}}[1]}\right),$$

while for minimation problems, $f_i^{\text{sub-reduce}}\left(y_{[i]}, y_{[i]\oplus_{\frac{\ell}{k}}[1]}\right)$ is

$$\min_{x_{[i\cdot k]\oplus[2]},\ldots,x_{[i\cdot k]\oplus[k]}} f_i^{\text{sub-ring}}\left(y_{[i]}, x_{[i\cdot k]\oplus[2]}, \ldots, x_{[i\cdot k]\oplus[k]}, y_{[i]\oplus_{\frac{\ell}{k}}[1]}\right).$$

Figure 5.2 is an illustration of the reduction. As shown in the figure, no bits belonging to only one sub-function is in the reduced problem. Also, each sub-function consists of two bits. Based on the reduction, the following theorem can be derived:

**Theorem 5.1.** *For a problem with chromosomes with length of $\ell$, fitness function $f^{\text{ring}}$ and the corresponding reduced problem having fitness function $f^{\text{reduced-ring}}$, suppose for binary-encoded selecto-recombinative GAs using OM, solving $f^{\text{reduced-ring}}$ with a $c$-composite or-*


doi:10.6342/NTU201902839

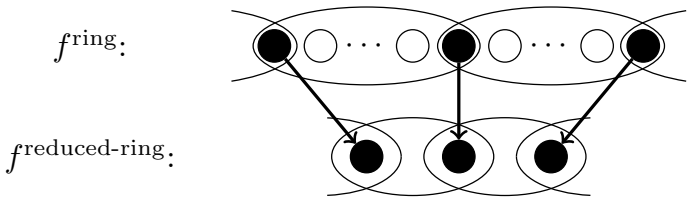



Figure 5.2: Illustration of problem reduction. Each ellipse represents a sub-function. Each solid circle represents a shared bit between the original problem and the reduced problem, whereas a hollow circle represents a reduced bit.

*acle with a success probability of* $1-\Theta\left(\ell^{-\alpha}\right)$*, where* $\alpha$ *is a positive constant, needs population size of order O* $(n)$*. Then, population size for solving a chromosome with length of* $\ell$ *for binary-encoded GAs using OM with a* $c$*-composite oracle with a success probability of* $1-\Theta\left(\ell^{-\alpha}\right)$*is O* $\left(2^k\left(\alpha\log\ell+k\right)+n\right)$.

*Proof.* ‘ For $k=0$, the reduced problem is the same as the original problem. Thus, population size of order O $(n)$ is enough for the original problem to have a success probability of $1-\Theta\left(\ell^{-\alpha}\right)$.

For $k\geq 1$, consider all the bits used in $f^{\text{ring}}$ but being reduced in $f^{\text{reduced-ring}}$. Since they can form $\frac{\ell}{k}$ non-overlapping segments of length $k-1$ and changing the middle $k-1$ bits in a sub-function into its local optimum does not destroy any bits used in two $f_i^{\text{sub-ring}}$s, if provided with all the schemas for the non-overlapping segments, with the aid of the oracle, the solving process can be extended from the one of solving $f^{\text{reduced-ring}}$ using O $(n)$ chromosomes.

Since O $(n)$ chromosomes can guarantee a success probability of $1-\Theta\left(\ell^{-\alpha}\right)$, to guarantee the whole solving process having a success probability at least $1-\Theta\left(\ell^{-\alpha}\right)$, the probability of not providing all the non-overlapping schemas for all the $\frac{\ell}{k}$ segments should be O $\left(\ell^{-\alpha}\right)$. By the result derived by Goldberg *et al.* (2001), this can be done with O $\left(2^k\left(\alpha\log\ell+k\right)\right)$ chromosomes. Thus, population size for solving a chromosome with length of $\ell$ using a binary-encoded selecto-recombinative GA using OM with a success probability of $1-\Theta\left(\ell^{-\alpha}\right)$ is O $\left(2^k\left(\alpha\log\ell+k\right)+n\right)$. □


doi:10.6342/NTU201902839

## 5.2 Tight bounds on population size

Theorem 5.1 forms a relationship between the population size before and after the reduction. As long as $k$ is a constant, the difference between the bound for the problem before reduction and the one after the reduction would be at most at the order of $\log \ell$, which is acceptable in practice. Therefore, the following context focused on problems with the following form:

$$f^{\text{homo-ring}}(\boldsymbol{x}) = \sum_{M \in \mathcal{F}_{\langle 2 \rangle}} f_M^{\text{sub-homo-ring}}(\boldsymbol{x}_M), \tag{5.1}$$

where $\mathcal{F}_{\langle 2 \rangle}$ is a homogeneous overlapping FOS and $\boldsymbol{x}_M$ is the part of $\boldsymbol{x}$ indicated by a mask $M$. To know more about problems with fitness functions of the same form as in Equation 5.1, the relationship between the optimal solution and the non-optimal ones is investigated.

**Lemma 5.2.** *A problem with its fitness function of the same form as in Equation 5.1 has optimal substructure property in the following form: for a consecutive segment, the segment of the golbal optimum is the optimal segment, given the bits before and after are the sames as the ones of the global optimum.*

*Proof.* Without loss of generality, assume the problem is a maximization problem. Let $\boldsymbol{x}^* = \langle x^*_{[i]} \rangle$, where $[i] \in \mathbb{G}_\ell$, be the optimal solution and the fitness function be $f$. Suppose $\boldsymbol{x}$ is a chromosome such that $x_{[i_1]} = x^*_{[i_1]}$ and $x_{[i_2]} = x^*_{[i_2]}$, where $i_1 \leq i_2$.

**Case 1.** If $i_1 = i_2$, consider the set of chromosomes $C_{\boldsymbol{x},i_1} = \left\{ \boldsymbol{y} \middle| y_{[i_1]} = x_{[i_1]} \right\}$. Since $\boldsymbol{x}^* \in C_{\boldsymbol{x},i_1}$, if $\langle x^*_{[i_1 \oplus 1]}, x^*_{[i_1 \oplus 2]}, \ldots, x^*_{[i_1 \oplus (\ell-1)]} \rangle$ is not the optimal substructure, then there exists $\boldsymbol{x}' \in C_{\boldsymbol{x},i_1}$ such that $f\left(\boldsymbol{x}'\right)$ is better than $f(\boldsymbol{x}^*)$, which is a contradiction.

**Case 2.** If $i_1 + 1 = i_2$, using similar technique in Case 1., consider the set of chromosomes $C_{\boldsymbol{x},i_1,i_2} = \left\{ \boldsymbol{y} \middle| y_{[i_1]} = x_{[i_1]}, y_{[i_2]} = x_{[i_2]} \right\}$. Since $\boldsymbol{x}^* \in C_{\boldsymbol{x},i_1,i_2}$, if $\langle x^*_{[i_1 \oplus 2]}, x^*_{[i_1 \oplus 3]}, \ldots, x^*_{[i_1 \oplus (\ell-1)]} \rangle$ is not the optimal substructure, then there exists $\boldsymbol{x}' \in C_{\boldsymbol{x},i_1,i_2}$ such that $f\left(\boldsymbol{x}'\right)$ is better than $f(\boldsymbol{x}^*)$, which is a contradiction.


doi:10.6342/NTU201902839

**Case 3.** If $i_1 + 1 < i_2$, define index sets $M_{i_1,i_2} = \langle i | i_i < i < i_2 \rangle$ and $\overline{M_{i_1,i_2}} = \langle i_2 \oplus 1, i_2 \oplus 2, \ldots, i_2 \oplus (i_1 - i_2 + \ell - 1) \rangle$. Consider the set of chromosomes $C_{\boldsymbol{x},M_{i_1,i_2}} = \left\{ \boldsymbol{y} \middle| \forall\, i \in M_{i_1,i_2} \cup \{i_1, i_2\}, y_{[i]} = x_{[i]} \right\}$. If the optimal substructure of the index set $M_{i_1,i_2}$ is not $\langle x^*_{[i_1 \oplus 1]}, x^*_{[i_1 \oplus 2]}, \cdots, x^*_{[i_1 \oplus (i_2 - i_1 - 1)]} \rangle$, suppose it is $\langle x'_{[i_1 \oplus 1]}, x'_{[i_1 \oplus 2]}, \cdots, x'_{[i_1 \oplus (i_2 - i_1 - 1)]} \rangle$, then if the problem is a maximization problem,

$$
\begin{aligned}
& f_{i_1,i_1 \oplus 1}\left(x^*_{[i_1]}, x'_{[i_1 \oplus 1]}\right) + f_{i_2 \ominus 1,i_2}\left(x'_{[i_2 \ominus 1]}, x^*_{[i_2]}\right) + \sum_{i=i_1+1}^{i_2-2} f_{i,i \oplus 1}\left(x'_{[i]}, x'_{[i \oplus 1]}\right) + \sum_{i \in \overline{M_{i_1,i_2}}} f_{i,i \oplus 1}\left(x^*_{[i]}, x^*_{[i \oplus 1]}\right) \\
< & f_{i_1,i_1 \oplus 1}\left(x^*_{[i_1]}, x^*_{[i_1 \oplus 1]}\right) + f_{i_2 \ominus 1,i_2}\left(x^*_{[i_2 \ominus 1]}, x^*_{[i_2]}\right) + \sum_{i=i_1+1}^{i_2-2} f_{i,i \oplus 1}\left(x^*_{[i]}, x^*_{[i \oplus 1]}\right) + \sum_{i \in \overline{M_{i_1,i_2}}} f_{i,i \oplus 1}\left(x^*_{[i]}, x^*_{[i \oplus 1]}\right) \\
= & f\left(\boldsymbol{x}^*\right),
\end{aligned}
$$

which is a contradiction. Similar result happens when the problem is a minimization problem. Thus, given $x_{[i_1]} = x^*_{[i_1]}$ and $x_{[i_2]} = x^*_{[i_2]}$, the optimal substructure of the index set $M_{i_1,i_2}$ is $\langle x^*_{[i_1 \oplus 1]}, x^*_{[i_1 \oplus 2]}, \ldots, x^*_{[i_1 \oplus (i_2 - i_1 - 1)]} \rangle$. For the same reason, given $x_{[i_1]} = x^*_{[i_1]}$ and $x_{[i_2]} = x^*_{[i_2]}$, the optimal substructure of the index set $\overline{M_{i_1,i_2}}$ is $\langle x^*_{[i_2 \oplus 1]}, x^*_{[i_2 \oplus 2]}, \ldots, x^*_{[i_2 \oplus (i_1 - i_2 + \ell - 1)]} \rangle$. Therefore, a problem with its fitness function of the same form as in Equation 5.1 has optimal substructure property. □

Based on optimal substructure property, the following theorem holds.

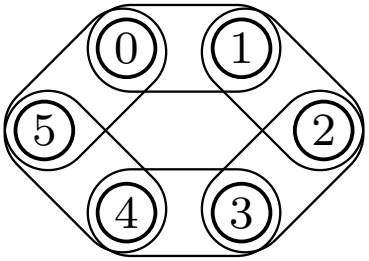
**Theorem 5.3.** *Population size required for binary-encoded GAs using OM with an oracle to solve a problem with its fitness function of the same form as in Equation 5.1 with a success probability of* $1 - \Theta\left(\ell^{-\alpha}\right)$*, where* $\ell$ *is the chromosome length and* $\alpha$ *is a positive constant, is* $\Theta\left(\alpha \log \ell\right)$.

*Proof.* Th process that Figure 5.4 conducts gives an intuition that leads to the proof. Consider optimizing a 6-bit problem with the population shown in Figure 5.3(b) and global optimum being the chromosomes with all 1s. Since the probelm has optimal substructure property, if there exists schemas of order 2 that in the bits within are the bits before and after a consecutive bits of a chromosome which have alleles the same as the ones of the global optimum in the problem structure, then the consecutive segment can be donated to


doi:10.6342/NTU201902839

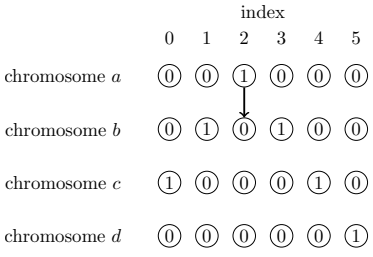
form a larger consecutive segment. For example, in Figure 5.3(c), the global optimum is the chromomsome with all 1s, and the bits with index $1$ and $3$ of chromosome $b$ encloses the bit with index $2$ of chromsome $a$. Thus, chromosome $a$ can donote the bit with index $2$ to chromosome $b$, and the result shows in Figure 5.3(d), where chromosome $b$ has a schema of order $3$ that the global optimum has. If the condition can be fulfilled continuously, the global optimum can be achieved by enlarging the number of the consecutive bits of a chromosome which have alleles the same as the ones of the global optimum, such as the operations shown from Figures 5.3(c) to Figure 5.3(h).


doi:10.6342/NTU201902839

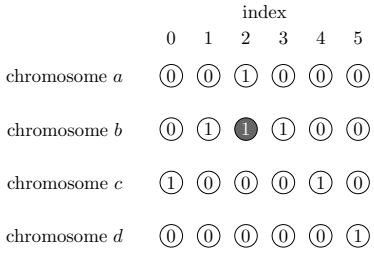



(a) Problem structure

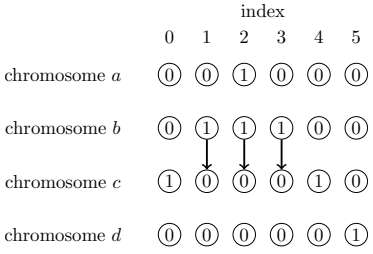

(b) Initialized population



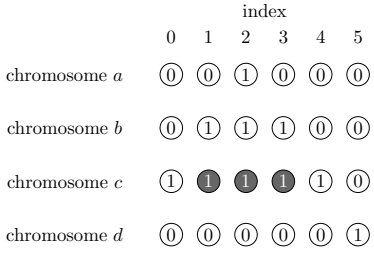

(c)



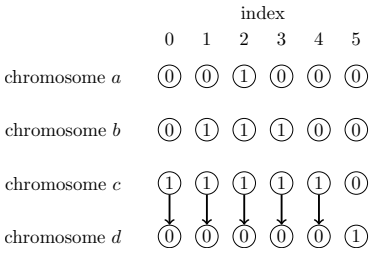

(d)




(e)




(f)




(g)


(h)



doi:10.6342/NTU201902839

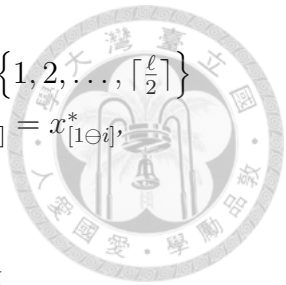
Figure 5.4: Illustration of a process that gives an intuition that leads to the proof of Theorem 5.3. Figure 5.3(a) shows the problem structure of a 6-bit problem with fitness function of the same form as the one in Equation 5.1, where a circle means a bit with the number inside the circle representing its index , each rounded rectangle represents a sub-function, and the global optimum is the chromosome with all 1s. Figure 5.3(b) shows the initialized population, where each row represents a chromosome, and the number on each column indicates the index. Figures 5.3(c) to 5.3(h) shows a process of making chromosome $d$ become the global optimum. The arrows in Figures 5.3(c), 5.3(e), and 5.3(g) means the bits donated from donor to the receiver, and Figures 5.3(d), 5.3(f), and 5.3(h) are the populations after the donations in Figures 5.3(c), 5.3(e), and 5.3(g) happens, respectively, with circles filled with grey identifying bits being donated.

Based on the intuition, Algorithm 1 is designed. Even though the overall process can not be guide by a 1-composite oracle, the overall requirement that Algorithm 1 can be conducted is having disjoint designated schemas of order at most 2 and total length of all the schemas is less than or equal to $\ell$, and by Theorem 4.3, a population with $\Theta(\alpha \log \ell)$ chromosomes can guarantee a sequence of donations described in the last paragraph that leads to global optimum from a single bit having the same allele as the one of the global optimum at the corresponding locus with probability $1 - \Theta(\ell^{-\alpha})$, where $\alpha$ is a positive constant. Therefore, population size required for binary-encoded GAs using OM with an oracle to solve a problem with its fitness function of the same form as in Equation 5.1 with a success probability of $1 - \mathrm{O}(\ell^{-\alpha})$, where $\ell$ is the chromosome length and $\alpha$ is a positive constant, is $\mathrm{O}(\alpha \log \ell)$. By combining the result with Theorem 3.7, the proof is compelete.

□

Since in the proof of Theoerm 5.3, not only the population size but also how to reach


doi:10.6342/NTU201902839



**Input** *Chromosomes* $\boldsymbol{x}^1$ *to* $\boldsymbol{x}^{\lceil \frac{\ell}{2} \rceil}$*, where,* $\boldsymbol{x}^i = \langle x^i_{[j]} \rangle$*, in which* $i \in \left\{1, 2, \ldots, \lceil \frac{\ell}{2} \rceil\right\}$ *and* $[j] \in \mathbb{G}_\ell$*, and* $\forall i \in \left\{1, 2, \ldots, \lceil \frac{\ell}{2} \rceil\right\}$*,* $x^i_{[i \ominus 1]} = x^*_{[i \ominus 1]}$*, and* $x^i_{[1 \ominus i]} = x^*_{[1 \ominus i]}$*, where* $\boldsymbol{x}^*$ *is the global optimum.*

**begin**

  **for** $i \leftarrow 1$ **to** $\lceil \frac{\ell}{2} \rceil - 1$ **by** $1$ **do**

    Choose $\boldsymbol{x}^i$ as donor and $\boldsymbol{x}^{i+1}$ as receiver, and donate mask $\{0, 1, \ldots, i-1\} \cup \{1 \ominus 1, 1 \ominus 2, \ldots, 1 \ominus i\}$.

  **Output** $\boldsymbol{x}^\lambda$

**Algorithm 1:** Process for reaching the global optimum.

the global optimum with sufficient number of chromosomes are specified. Thus, the nFE can be bounded.

**Theorem 5.4.** *With a probability of* $1 - O\left(\ell^{-\alpha}\right)$*, where* $\ell$ *is the chromosome length and* $\alpha$ *is a positive constant, for binary-encoded GAs using OM with an oracle to solve a problem with its fitness function of the same form as in Equation 5.1 with population size being* $O\left(\alpha \log \ell\right)$*, the nFE is* $O\left(\ell\right)$*.*

*Proof.* In the proof of Theorem 5.3, it is showed that $\Theta\left(\alpha \log \ell\right)$ chromosomes are sufficient for binary-encoded GAs using OM with an oracle to solve a problem with its fitness function of the same form as in Equation 5.1 with a success probability of $1 - \Theta\left(\ell^{-\alpha}\right)$ using a sequence of donations after which the number of a certain consecutive bits of a chromosome which have alleles the same as the ones of the global optimum is enlarged. Since the number of bits is strictly increasing with the number of donations, the number is upper bounded by $\ell$, the nFE is O (number of donation made), and only one specific case is discussed, with a probability of $1 - \mathrm{O}\left(\ell^{-\alpha}\right)$, where $\ell$ is the chromosome length and $\alpha$ is a positive constant, for binary-encoded GAs using OM with an oracle to solve a problem with its fitness function of the same form as in Equation 5.1 with population size being $\mathrm{O}\left(\alpha \log \ell\right)$, the nFE is $\mathrm{O}\left(\ell\right)$.

□



doi:10.6342/NTU201902839



doi:10.6342/NTU201902839

# Chapter 6

# Results with Problems beyond Ring Topologies

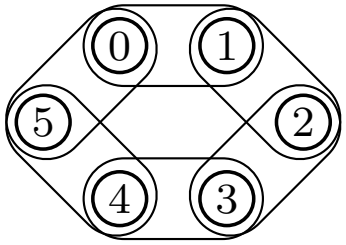
In Chapter 5, bounds on the population size for binary-encoded GAs using OM with an oracle to solve problems with fitness functions of the same form as in Equation 5.1 are derived. However, the problems discussed are very restricted, and not all problems in practice have that kind of fitness functions. Thus, in this chapter, bounds on the population size for binary-encoded GAs using OM with an oracle to solve problems with less restricted fitness functions, though may not be as tight, are derived.

## 6.1 Layered structure

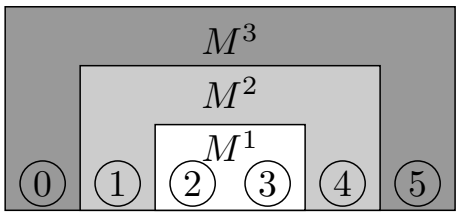
In the proof of Theorem 5.3, only the optimal substructure property proved in Lemma 5.2 is needed and one can reach the optimum by lengthening an optimal segment. In order to discuss a more generalized optimal substructure property, the layered structure of a fitness function and problem is defined.

**Definition 6.1.** ***layered sturcture****: A fitness function $f$ of a problem can be viewed as a layered structure of $\lambda$ layers if there exists a disjoint FOS $\langle M^1, M^2, \cdots, M^\lambda \rangle$ with $M^i$ being a non-empty mask $\forall\ i \in \{1, 2, \cdots, \lambda\}$ satisfying $\forall\ i \in \{1, 2, \cdots, \lambda\}$, exist functions $g_i$ and $h_i$ such that $f(\boldsymbol{x}) = g_i\left(\boldsymbol{x}_{\cup_{j=1}^{i} M^j}\right) + h_i\left(\boldsymbol{x}_{\cup_{j=i}^{\lambda} M^j}\right)$ where $\boldsymbol{x} = \langle x_{[s]} \rangle$ is a chromosome, and $\boldsymbol{x}_M$ is the part of $\boldsymbol{x}$ indicated by a mask $M$.*

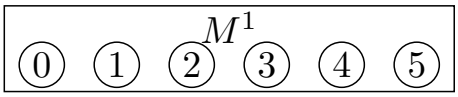

doi:10.6342/NTU201902839




(a) Example problem




(b) Layered structure of the problem




(c) Layered structure of the problem

Figure 6.1: Example of the layered structure. Figure 6.1(a) shows an illustration of the fitness function of a $6$-bit problem. The fitness function is the summation of $6$ sub-functions. Each circle represents a bit with its index inside, and each ellipse represents a sub-function. Figures 6.1(b) and 6.1(c) show two layered structures that the fitness function in Figure 6.1(a) can be viewed as. The indexes of the bits in the same colored region belong to the same mask.

Figure 6.1 gives an example of a $6$-bit problem having a layered structure of $3$ layers. Note that a fitness function can be viewed as multiple layered structures. For example, the fitness function in Figure 6.1(a) can also be viewed as layered structures shown in Figures 6.1(b) and 6.1(c). This example also shows that a fitness function of the same form as in Equation 5.1 can be viewed as a layered structure of $\lceil \frac{\ell}{2} \rceil$ layers with the FOS of the layered structure being the sequence of masks used to reach the global optimum in the way used in the proof of Theorem 5.3. This means that the layered structure preserves some property desired, and the optimal substructure property for the layered structure is derived in the following theorem to derive bounds on the population size for binary-encoded GAs using OM with an oracle to solve problems with fitness functions with more generalized structure.


doi:10.6342/NTU201902839

**Theorem 6.2.** *A problem with its fitness function can be viewed as a layered structure of* $\lambda$ *layers with FOS* $\mathcal{F} = \langle M^1, M^2, \cdots, M^\lambda \rangle$*and functions* $g_i$ *and* $h_i$ *such that* $f(\boldsymbol{x}) = g_i\left(\boldsymbol{x}_{\cup_{j=1}^{i} M^j}\right) + h_i\left(\boldsymbol{x}_{\cup_{j=i}^{\lambda} M^j}\right)$ *where* $\boldsymbol{x} = \langle x_{[s]} \rangle$ *is a chromosome, and* $\boldsymbol{x}_M$ *is the part of* $\boldsymbol{x}$ *indicated by a mask* $M$ *has optimal substructure property in the following form:* $\forall i \in \{1, 2, \cdots, \lambda\}$*, for any chromosome* $\boldsymbol{y}$ *if* $\boldsymbol{y}_{M^i} = \boldsymbol{x}^*_{M^i}$*, where* $\boldsymbol{x}^* = \langle x^*_{[s]} \rangle$ *is the optimal solution, then if the porblem is a maximization problem then* $\boldsymbol{x}^*_{\cup_{j=1}^{i} M^j} \in \operatorname{argmax}_{\boldsymbol{y}_{\cup_{j=1}^{i} M^j}} g_i\left(\boldsymbol{y}_{\cup_{j=1}^{i} M^j}\right)$ *and* $\boldsymbol{x}^*_{\cup_{j=i}^{\lambda} M^j} \in \operatorname{argmax}_{\boldsymbol{y}_{\cup_{j=i}^{\lambda} M^j}} h_i\left(\boldsymbol{y}_{\cup_{j=i}^{\lambda} M^j}\right)$*, and if the porblem is a minimization problem then* $\boldsymbol{x}^*_{\cup_{j=1}^{i} M^j} \in \operatorname{argmin}_{\boldsymbol{y}_{\cup_{j=1}^{i} M^j}} g_i\left(\boldsymbol{y}_{\cup_{j=1}^{i} M^j}\right)$ *and* $\boldsymbol{x}^*_{\cup_{j=i}^{\lambda} M^j} \in \operatorname{argmin}_{\boldsymbol{y}_{\cup_{j=i}^{\lambda} M^j}} h_i\left(\boldsymbol{y}_{\cup_{j=i}^{\lambda} M^j}\right)$.

*Proof.* If the problem is a maximization problem and suppose $\boldsymbol{x}^*_{\cup_{j=1}^{i} M^j} \notin \operatorname{argmax}_{\boldsymbol{y}_{\cup_{j=1}^{i} M^j}} g_i\left(\boldsymbol{y}_{\cup_{j=1}^{i} M^j}\right)$. If the chromosome $\boldsymbol{y}^* = \langle y^*_{[s]} \rangle$ satisifes $\boldsymbol{y}^*_{\cup_{j=i}^{\lambda} M^j} = \boldsymbol{x}^*_{\cup_{j=i}^{\lambda} M^j}$ and $\boldsymbol{y}^*_{\cup_{j=1}^{i} M^j} \in \operatorname{argmax}_{\boldsymbol{y}_{\cup_{j=1}^{i} M^j}} g_i\left(\boldsymbol{y}_{\cup_{j=1}^{i} M^j}\right)$, then

$$
\begin{aligned}
& f(\boldsymbol{y}^*) - f(\boldsymbol{x}^*) \\
=& \left(g_i\left(\boldsymbol{y}^*_{\cup_{j=1}^{i} M^j}\right) + h_i\left(y^*_{\cup_{j=i}^{\lambda} M^j}\right)\right) - \left(g_i\left(\boldsymbol{x}^*_{\cup_{j=1}^{i} M^j}\right) + h_i\left(x^*_{\cup_{j=i}^{\lambda} M^j}\right)\right) \\
=& g_i\left(\boldsymbol{y}^*_{\cup_{j=1}^{i} M^j}\right) - g_i\left(\boldsymbol{x}^*_{\cup_{j=1}^{i} M^j}\right) + h_i\left(y^*_{\cup_{j=i}^{\lambda} M^j}\right) - h_i\left(x^*_{\cup_{j=i}^{\lambda} M^j}\right) \\
=& g_i\left(\boldsymbol{y}^*_{\cup_{j=1}^{i} M^j}\right) - g_i\left(\boldsymbol{x}^*_{\cup_{j=1}^{i} M^j}\right) \\
>& 0.
\end{aligned}
$$

This violates the assumption that $\boldsymbol{x}^*$ is the optimal solution. Thus, $\boldsymbol{x}^*_{\cup_{j=1}^{i} M^j} \in \operatorname{argmax}_{\boldsymbol{y}_{\cup_{j=1}^{i} M^j}} g_i\left(\boldsymbol{y}_{\cup_{j=1}^{i} M^j}\right)$. The rest of the parts of the theorem can be proved in the similar way.

□

Based on the results derived above, the upper bound on the population size for binary-encoded GAs using OM with an oracle to solve problems with fitness functions that can be viewed as a layered structure is derived.

**Theorem 6.3.** *The population size required for binary-encoded GAs using OM with an oracle to solve a problem with its fitness function that can be viewed as a layered structure*


doi:10.6342/NTU201902839

*of $\lambda$ layers with FOS $\mathcal{F} = \langle M^1, M^2, \cdots, M^\lambda \rangle$ with a success probability of $1 - \Theta\left(\ell^{-\alpha}\right)$, where $\ell$ is the chromosome length and $\alpha$ is a positive constant, is O $\left(\alpha 2^{\max_i |M^i|} \log \ell\right)$, where $|M|$ denotes the number of elements in $M$.*

*Proof.* For $i \in \{1, 2, \cdots, \lambda\}$, if there exists chromosomes $\boldsymbol{x}^i = \langle x^i_{[s]} \rangle$ such that $\boldsymbol{x}^i_{M^i} = \boldsymbol{x}^*_{M^i}$, where $\boldsymbol{x}^* = \langle x^*_{[s]} \rangle$ is the global optimmum, $\boldsymbol{x}_M$ is the part of $\boldsymbol{x}$ indicated by a mask $M$, and all $\boldsymbol{x}^i$s do not need to be disjoint, then the global optimum can be reached by Algorithm 2. The condition of the previous algorithm is gathering all the non-overlapping schemas of order $|M^1|, |M^2|, \cdots, \left|M^\lambda\right|$ respectively. By Theorem 4.3, O $\left((\alpha + 1)\, 2^{\max_i |M^i|} \log \ell\right)$ chromosomes are enough to guarantee the condition be fulfilled with a probability of $1 - \Theta\left(\ell^{-\alpha-1}\right)$. Since $\alpha$ is a constant, by applying union bound, the population size required for binary-encoded GAs using OM with an oracle to solve a problem with its fitness function that can be viewed as a layered structure of $\lambda$ layers with FOS $\mathcal{F} = \langle M^1, M^2, \cdots, M^\lambda \rangle$ with a success probability of $1 - \Theta\left(\ell^{-\alpha}\right)$ is O $\left(\alpha 2^{\max_i |M^i|} \log \ell\right)$.

**begin**
  **for** $i \leftarrow 1$ **to** $\lambda - 1$ **by** $1$ **do**
    Choose $\boldsymbol{x}^i$ as donor and $\boldsymbol{x}^{i+1}$ as receiver, and donate mask $\cup_{j=1}^{i} M^j$.
  **Output** $\boldsymbol{x}^\lambda$

**Algorithm 2:** Process for reaching the global optimum.

□

**Theorem 6.4.** *With a probability of $1 - O\left(\ell^{-\alpha}\right)$, where $\ell$ is the chromosome length and $\alpha$ is a positive constant, for binary-encoded GAs using OM with an oracle to solve a problem with its fitness function that can be viewed as a layered structure of $\lambda$ layers with FOS $\mathcal{F} = \langle M^1, M^2, \cdots, M^\lambda \rangle$ with population size being O $\left(\alpha 2^{\max_i |M^i|} \log \ell\right)$, the nFE is O $(\lambda)$.*

*Proof.* In the proof of Theorem 6.3, it is showed that O $\left(\alpha 2^{\max_i |M^i|} \log \ell\right)$ chromosomes are sufficient for binary-encoded GAs using OM with an oracle to solve a problem with its fitness functionthat can be viewed as a layered structure of $\lambda$ layers with FOS $\mathcal{F} = \langle M^1, M^2, \cdots, M^\lambda \rangle$ with a success probability of $1 - \Theta\left(\ell^{-\alpha}\right)$ using Algorithm 2. Since


doi:10.6342/NTU201902839

the number of bits is strictly increasing with the number of donations, the number is upper bounded by $\ell$, and the nFE is O (number of donation made) if using an oracle. Since a larger population only increases the success probability, with a probability of $1 - \mathrm{O}\left(\ell^{-\alpha}\right)$, where $\ell$ is the chromosome length and $\alpha$ is a positive constant, for binary-encoded GAs using OM with an oracle to solve a problem with its fitness function that can be viewed as a layered structure of $\lambda$ layers with FOS $\mathcal{F} = \langle M^1, M^2, \cdots, M^\lambda \rangle$ with population size being $\mathrm{O}\left(\alpha 2^{\max_i |M^i|} \log \ell\right)$, the nFE is $\mathrm{O}(\lambda)$.

□

Theorem 6.3 states an upper bound on the population size required for binary-encoded GAs using OM with an oracle to have a success rate of $1 - \Theta\left(\ell^{-\alpha}\right)$. In the following sections, to show its applicability to real cases, some fitness functions with specific topologies are discussed. To begin with, fitness functions with torus topologies are discussed.

## 6.2 Results on torus topologies

In this section, fitness functions with torus topologies are focused. To begin with, in this section, $\ell = r^d$, where $r, d \in \mathbb{Z}^+$. Also, re-define a chromosome $\boldsymbol{x} = \left\langle x_{[i_1],[i_2],\ldots,[i_d]} \right\rangle$, where $[i_1], [i_2], \ldots, [i_d] \in \underbrace{\mathbb{G}_r \times \mathbb{G}_r \times \cdots \times \mathbb{G}_r}_{d}$, a fitness function with a torus topology of dimension $d$, so called a $d$-torus, of radius $r$, $f^{\text{torus},d,r}$, is defined as follows:

$$f^{\text{torus},d,r}(\boldsymbol{x}) = \sum_{i_1=0}^{r-1} \sum_{i_2=0}^{r-1} \cdots \sum_{i_d=0}^{r-1} \sum_{j=1}^{d} f^{\text{sub-torus}}_{i_1,i_2\ldots,i_d,j} \left( x_{[i_1],[i_2],\ldots,[i_d]}, x_{[i_1],\ldots[i_{j-1}],[i_j]\oplus_r[1],[i_{j+1}],\ldots,[i_d]} \right). \tag{6.1}$$

Figure 6.2 shows an example of a torus topology of 2-torus of radius 12, where each intersection of line segments is a bit, and each line segment between two bits represents a sub-function. Torus is the first topology to be discussed because, by definition, a fitness function of the same form as in Equation 5.1 is a fitness function with a torus topology of dimension 1 of radius $\ell$, meaning that a torus topology is a generalized case for a ring topology. Since Theorem 6.3 states that the upper bound of the population size is exponen-


doi:10.6342/NTU201902839

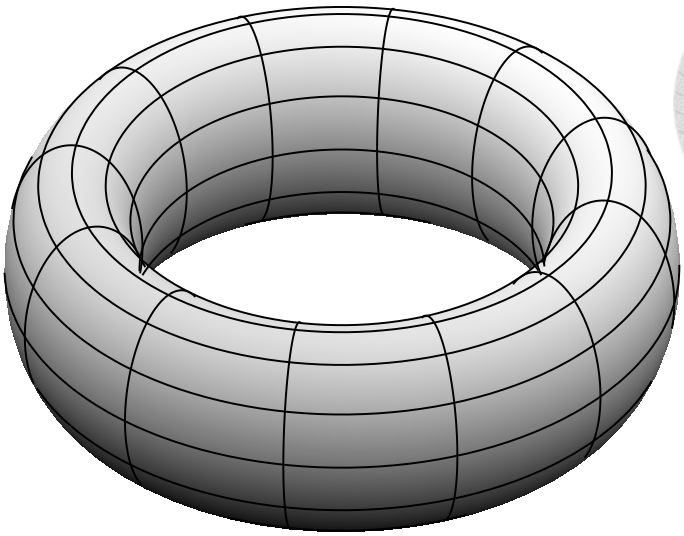


Figure 6.2: Example of a torus topology. The figure is a 2-torus of radius 12, where each intersection of line segments is a bit, and each line segment between two bits represents a sub-function.

tial to the size of the mask with the largest number of elements in the FOS of the layered structure, the next to be answered is: What kind of layered structure a fitness function of the same form as in Equation 6.1 should be viewed as in order to have a small maximal size in the FOS? In the following, a layered structure is discussed based on the FOS used in a ring topology. In Figure 6.1(b), how a 1-torus of radius 6 can be separated into a layered structure is shown. For a fitness function with a topology of a $d$-torus of radius $r$, if the same method can be applied along one axis, then a fitness function with a topology of a $d$-torus of radius $r$ can be viewed as a layered structure of $\lceil \frac{r}{2} \rceil$ layers with a maximal size in the FOS being $2r^{d-1}$. Figure 6.3 shows an example of how it can be applied on a 2-torus. Based on this intuition, the following theorems are derived.

**Theorem 6.5.** *The population size required for binary-encoded GAs using OM with an oracle to solve a problem with fitness function of the same form as in Equation 6.1 with a success probability of* $1 - \Theta\left(\ell^{-\alpha}\right)$*, where* $\ell$ *is the chromosome length and* $\alpha$ *is a positive constant, is* $O\left(\alpha 2^{2\ell^{\frac{d-1}{d}}} \log \ell\right)$.

*Proof.* Consider the following FOS $\mathcal{F} = \langle M^1, M^2, \cdots, M^{\lceil \frac{r}{2} \rceil} \rangle$, where $M^i = \left\{ [j_1], [j_2], \ldots, [j_d] \middle| \left\lfloor \left| j_1 - \frac{r-1}{2} \right| + 1 \right\rfloor = i \right\}$. A fitness function of the same form as in Equa-


doi:10.6342/NTU201902839

tion 6.1 can be viewed as a layered structure of $\lceil \frac{r}{2} \rceil$ layers with FOS $\mathcal{F}$ and $\forall\, i < \lceil \frac{r}{2} \rceil$,

$$g_i\left(\boldsymbol{x}_{\cup_{j=1}^{i} M^j}\right) = \sum_{\substack{\left\lfloor \frac{r-1}{2}-(j-1)\right\rfloor -1 \le s_1 \le \left\lceil \frac{r-1}{2}+(j-1)\right\rceil, \\ 0 \le s_2,\ldots,s_d \le r-1,\ 1 \le t \le d}} \cdots \sum f^{\text{sub-torus}}_{s_1,s_2\ldots,s_d,t}\left(x_{[s_1],\ldots,[s_d]}, x_{[s_1],\ldots[s_{t-1}],[s_t]\oplus_r[1],[i_{t+1}],\ldots,[s_d]}\right)$$

and

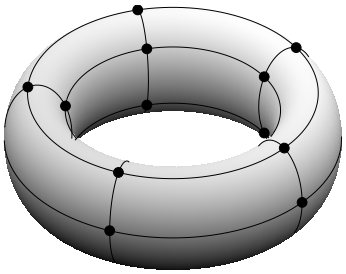
$$h_i\left(\boldsymbol{x}_{\cup_{j=i}^{\lceil \frac{r}{2} \rceil} M^j}\right) = \sum_{\substack{0 \le s_1 \le \left\lfloor \frac{r-1}{2}-(j-1)\right\rfloor -2 \\ \vee \left\lceil \frac{r-1}{2}+(j-1)\right\rceil +1 \le s_1 \le r-1, \\ 0 \le s_2,\ldots,s_d \le r-1,\ 1 \le t \le d}} \cdots \sum f^{\text{sub-torus}}_{s_1,s_2\ldots,s_d,t}\left(x_{[s_1],\ldots,[s_d]}, x_{[s_1],\ldots[s_{t-1}],[s_t]\oplus_r[1],[i_{t+1}],\ldots,[s_d]}\right),$$

where $\boldsymbol{x}$ is a chromosome, and $\boldsymbol{x}_M$ is the part of $\boldsymbol{x}$ indicated by a mask $M$. $g_{\lceil \frac{r}{2} \rceil}(\boldsymbol{x}) = f^{\text{torus},d,r}(\boldsymbol{x})$, and $h_{\lceil \frac{r}{2} \rceil} = 0$ for all inputs. The maximum size of $M^i$ in $\mathcal{F}$ is $2r^{d-1} = 2\left(r^d\right)^{\frac{d-1}{d}} = 2\ell^{\frac{d-1}{d}}$. By applying Theorem 6.3, the population size required for binary-encoded GAs using OM with an oracle to solve a problem with fitness function of the same form as in Equation 6.1 with a success probability of $1 - \Theta\left(\ell^{-\alpha}\right)$, where $\ell$ is the chromosome length and $\alpha$ is a positive constant, is $\mathrm{O}\left(\alpha 2^{2\ell^{\frac{d-1}{d}}} \log \ell\right)$. □

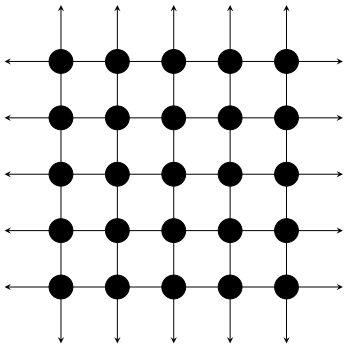
**Theorem 6.6.** *With a probability of $1 - O\left(\ell^{-\alpha}\right)$, where $\ell$ is the chromosome length and $\alpha$ is a positive constant, for binary-encoded GAs using OM with an oracle to solve a problem with fitness function of the same form as in Equation 6.1 with population size being $O\left(\alpha 2^{2\ell^{\frac{d-1}{d}}} \log \ell\right)$, the nFE is $O\left(\lceil \frac{r}{2} \rceil\right)$.*

*Proof.* The result can be directly inferred from Theorems 6.4 and 6.5. □

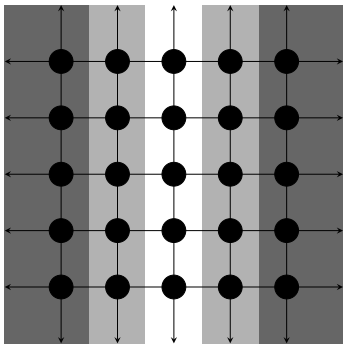

doi:10.6342/NTU201902839



(a) A 2-torus of radius 5.



(b) Planar representation of Figure 6.3(a).



(c) Layered structure of Figure 6.3(b).

Figure 6.3: Example of forming a layered structure from a torus topology. Figure 6.3(a) is an illustration of a fitness function with a topology of a 2-torus of radius 5, where each black circle represents a bit and each line segment between two circles represents a sub-function. Figure 6.3(b) is the planar representation of Figure 6.3(a), where the two arrows on the same line mean that the two line segments are connected and belong to the same sub-function. Figure 6.3(c) shows a layered structure that the fitness function in Figure 6.3(a) can be viewed as. The indexes of the bits in the same colored region belong to the same mask.



doi:10.6342/NTU201902839

## 6.3 Results on hypercube topologies

Torus topologies discussed in the last section is the generalization of ring topologies discussed in Chapter 5. However, the homogeneity that torus topologies hold makes the visualization or imagination more difficult, let alone discussing layered structures on them. Thus, in this section, fitness functions with hypercubic topologies that have something in common with torus topologies but break down the homogeneity are focused. To begin with, in this section, $\ell = (r+1)^d$, where $r, d \in \mathbb{Z}^+$. Also, re-define a chromosome $\boldsymbol{x} = \left\langle x_{[i_1],[i_2],\ldots,[i_d]} \right\rangle$, where $[i_1], [i_2], \ldots, [i_d] \in \underbrace{\mathbb{G}_{r+1} \times \mathbb{G}_{r+1} \times \cdots \times \mathbb{G}_{r+1}}_{d}$, a fitness function with a hypercubic topology of dimension $d$, so-called a $d$-cube, of radius $r$, $f^{\text{cube},d,r}$, is defined as follows:

$$f^{\text{cube},d,r}(\boldsymbol{x}) = \sum_{\substack{0 \le i_1, i_2, \ldots, i_d \le r, \\ j \in \{1,2,\ldots,d\},\ i_j \neq r}} \cdots \sum f^{\text{sub-cube}}_{i_1,i_2\ldots,i_d,j} \left( x_{[i_1],[i_2],\ldots,[i_d]}, x_{[i_1],\ldots[i_{j-1}],[i_j] \oplus_{r+1} [1],[i_{j+1}],\ldots,[i_d]} \right). \tag{6.2}$$

Figure 6.4 shows an example of a hypercube topology of a 3-cube of radius 10, where

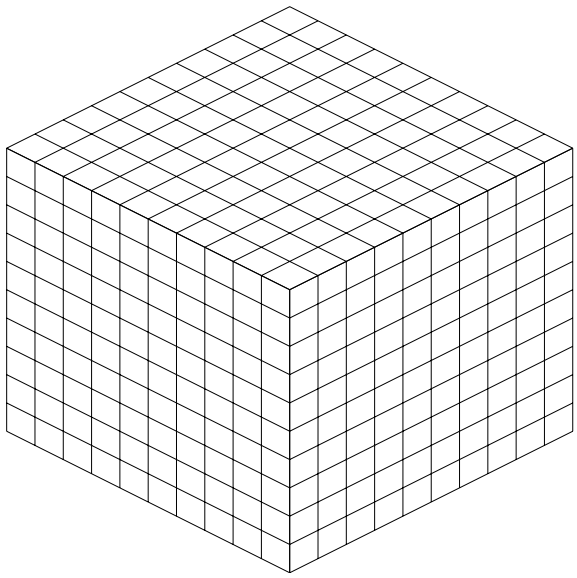
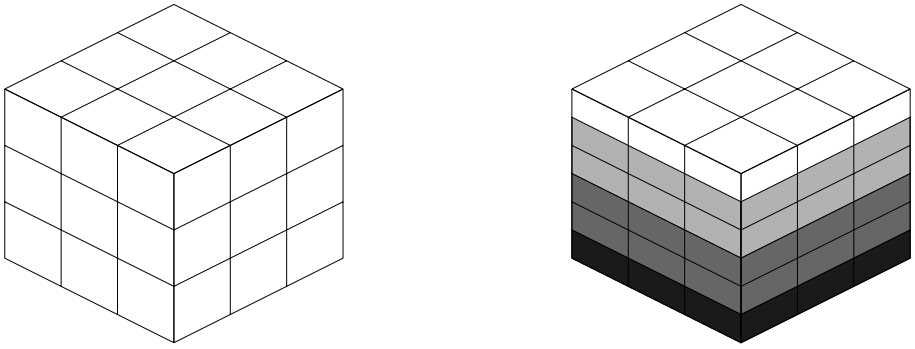


Figure 6.4: Example of a hypercube topology. The figure is a 3-cube of radius 10, where each intersection of line segments is a bit, and each line segment between two bits represents a sub-function.



doi:10.6342/NTU201902839

each intersection of line segments is a bit, and each line segment between two bits represents a sub-function. The main difference between Equations 6.2 and 6.1 is that the sub-functions in Equation 6.2 do not form any cycle along one axis, meaning that there is no homogeneity within fitness functions of the same form as in Equation 6.2. In Figures 6.1(b) and 6.3, how can a torus be separated into a layered structure by separating along one axis is demonstrated. By applying a similar technique to hypercubes, a fitness function with a topology of a $d$-cube of radius $r$ can be viewed as a layered structure of $r+1$ layers with a maximal size in the FOS being $r^{d-1}$. Figure 6.5 shows an example of how it can be applied to a 3-cube. Based on this intuition, the following theorems are derived.



(a) A 3-torus of radius 3. (b) Layered structure of Figure 6.5(a).

Figure 6.5: Example of forming a layered structure from a hypercube topology. Figure 6.5(a) is an illustration of a fitness function with a topology of a 3-cube of radius 3, where each intersection of line segments is a bit, and each line segment between two bits represents a sub-function. Figure 6.5(b) shows a layered structure that the fitness function in Figure 6.5(a) can be viewed as. The indexes of the bits in the same colored region belong to the same mask.

**Theorem 6.7.** *The population size required for binary-encoded GAs using OM with an oracle to solve a problem with fitness function of the same form as in Equation 6.2 with a success probability of* $1-\Theta\left(\ell^{-\alpha}\right)$*, where* $\ell$ *is the chromosome length and* $\alpha$ *is a positive constant, is* $O\left(\alpha 2^{\ell^{\frac{d-1}{d}}}\log\ell\right)$.


doi:10.6342/NTU201902839

*Proof.* Consider the following FOS $\mathcal{F} = \langle M^1, M^2, \cdots, M^{r+1} \rangle$, where $M^i = \{[j_1], [j_2], \ldots, [j_d] | j_1 = i - 1\}$. A fitness function of the same form as in Equation 6.1 can be viewed as a layered structure of $r + 1$ layers with FOS $\mathcal{F}$ with

$$g_i\left(\boldsymbol{x}_{\cup_{j=1}^{i} M^j}\right) = \sum_{\substack{0 \le s_1 \le j-1, \\ 0 \le s_2, \ldots, s_d \le r, \\ t \in \{1,2,\ldots,d\},\ s_t \ne r}} \cdots \sum f^{\text{sub-cube}}_{s_1, s_2 \ldots, s_d, t}\left(x_{[s_1],\ldots,[s_d]}, x_{[s_1],\ldots[s_{t-1}],[s_t]\oplus_{r+1}[1],[i_{t+1}],\ldots,[s_d]}\right)$$

and

$$h_i\left(\boldsymbol{x}_{\cup_{j=i}^{r+1} M^j}\right) = \sum_{\substack{j \le s_1 \le r, \\ 0 \le s_2, \ldots, s_d \le r, \\ t \in \{1,2,\ldots,d\},\ s_t \ne r}} \cdots \sum f^{\text{sub-cube}}_{s_1, s_2 \ldots, s_d, t}\left(x_{[s_1],\ldots,[s_d]}, x_{[s_1],\ldots[s_{t-1}],[s_t]\oplus_{r+1}[1],[i_{t--+1}],\ldots,[s_d]}\right),$$

where $\boldsymbol{x}$ is a chromosome, and $\boldsymbol{x}_M$ is the part of $\boldsymbol{x}$ indicated by a mask $M$. The maximum size of $M^i$ in $\mathcal{F}$ is $(r+1)^{d-1} = \left((r+1)^d\right)^{\frac{d-1}{d}} = \ell^{\frac{d-1}{d}}$. By applying Theorem 6.3, the population size required for binary-encoded GAs using OM with an oracle to solve a problem with fitness function of the same form as in Equation 6.1 with a success probability of $1 - \Theta\left(\ell^{-\alpha}\right)$, where $\ell$ is the chromosome length and $\alpha$ is a positive constant, is $\mathrm{O}\left(\alpha 2^{\ell^{\frac{d-1}{d}}} \log \ell\right)$. □

**Theorem 6.8.** *With a probability of* $1 - O\left(\ell^{-\alpha}\right)$*, where* $\ell$ *is the chromosome length and* $\alpha$ *is a positive constant, for binary-encoded GAs using OM with an oracle to solve a problem with fitness function of the same form as in Equation 6.2 with population size being* $O\left(\alpha 2^{\ell^{\frac{d-1}{d}}} \log \ell\right)$*, the nFE is* $O\left(r+1\right)$.

*Proof.* The result can be directly inferred from Theorems 6.4 and 6.7. □

Theorems 6.5 and 6.7 derive bound on the population size required for binary-encoded GAs using OM with an oracle to solve a problem with fitness function of the same form as in Equation 6.1 and in Equation 6.2 respectively. The only difference between the topologies is homogeneity. However, the order of the upper bound can be reduced to nearly half. Also, for more complicated problem structures that result in a higher dimension of


doi:10.6342/NTU201902839

certain structures or topologies, the upper bound will be closer to the exponential of the chromosome length, and this phenomenon fits the intuition.

## 6.4 Results on small-world topologies

In Sections 6.2 and 6.3, upper bounds on the number of chromosomes required for binary-encoded GAs using OM with an oracle to solve a problem with torus or hypercube topology is derived. However, in both cases, for $d \geq 2$, the population size required is greater than the nFE needed, meaning that one will initialize a chromosome and use it without knowing the fitness of the chromosome. This is possible for a GA when using an oracle, but not in practical use. Also, in 2-D Ising spin-glass problems, which have two-dimensional torus structures, the derived bound is sub-exponential. However, in practice, the population size required need not be as large, meaning the derived bounds are still loose. Thus, in this section, bounds on the population size required for binary-encoded GAs using OM with an oracle to solve a problem with small-world topologies, a class of modified ring topologies, are estimated to find results more applicable in practice. To begin with, a fitness function with small-world topology of size $\ell$ with $\iota$ bridges, $f^{\text{small-world},\ell,\iota}$, is defined as follows:

$$f^{\text{small-world},\ell,\iota}(\boldsymbol{x}) = \sum_{i=0}^{\ell-1} f_i^{\text{sub-small-world}}\left(x_{[i]}x_{[i]\oplus[1]}\right) + \sum_{j=0}^{\iota-1} f_{\ell+j}^{\text{sub-small-world}}\left(x_{[s_{j,1}]}x_{[s_{j,2}]}\right), \quad (6.3)$$

where $\forall\, j \in \{0,\ldots,\iota-1\}, s_{j,1} < s_{j,2} \wedge s_{j,1} \neq s_{j,2}\oplus 1 \wedge s_{j,1} \neq s_{j,2}\ominus 1$ ,and $(s_{j_1,1}, s_{j_1,2}) \neq (s_{j_2,1}, s_{j_2,2})$ if $j_1 \neq j_2\ \forall\ j_1, j_2 \in \{0,\ldots,\iota-1\}$. Figure 6.6 shows an example of a small-world topology of size $20$ with $3$ bridges, where each circle is a bit, and each line segment between two bits represents a sub-function. The main difference between Equations 5.1 and 6.3 is that the homogeneity that ring topology holds is destroyed by the bridges. By adding bridges, the ring can be separated into a part containing all bits used in bridges and different chains. Figure 6.7 shows an example of how it can be applied to a small-world topology of size $6$ with $2$ bridges. Based on this observation, the following theorem holds.


doi:10.6342/NTU201902839

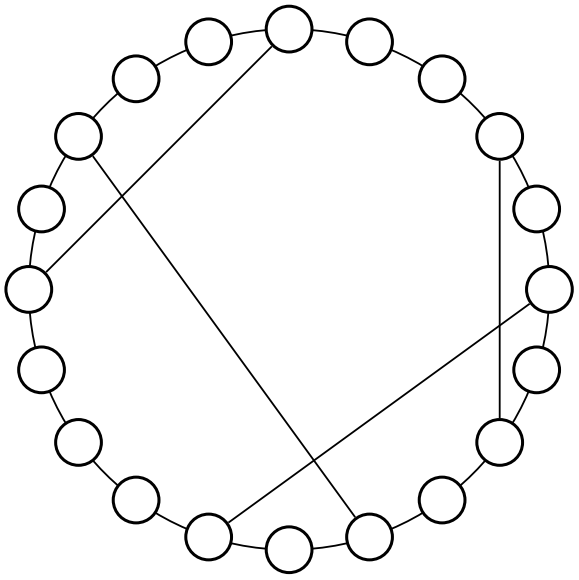


Figure 6.6: Example of a small-world topology. The figure is a small-world topology of size 20 with 3 bridges, where each circle is a bit, and each line segment between two bits represents a sub-function.

**Theorem 6.9.** *The population size required for binary-encoded GAs using OM with an oracle to solve a problem with fitness function of the same form as in Equation 6.3 with* $\log_2 \Theta(\ell) - \log_2 \log \ell + \Theta(1)$ *bits used in bridges with a success probability of* $1 - \Theta(\ell^{-\alpha})$*, where* $\ell$ *is the chromosome length and* $\alpha$ *is a positive constant, is* $O(\alpha\ell)$*, and the nFE* $O(\ell)$*.*

*Proof.* Consider the partition of a small-world topology mentioned above. Suppose the fitness function is partitioned into $a$ chains with length $b_i$, $1 \leq i \leq a$. Thus, $\log_2 \Theta(\ell) - \log_2 \log \ell + \Theta(1) + \sum_{i=1}^{a} b_i = \ell$. First, by Theorem 4.3, a population of $\Theta(\alpha\ell)$ chromosomes suffices to provide a chromosome having all the alleles of the bits used in the bridges being the same as the ones of the global optimum with probability $1 - \Theta(\ell^{-\alpha})$. Note that since each chain belongs to the ring, Lemma 5.2 holds and therefore, one can donate a chain with all the alleles the same as the ones of the global optimum to the part having all the alleles of the bits used in the bridges being the same as the ones of the global optimum. By Theorems 6.7 and 6.8, for a chain with length $b_i$ a population size of $\alpha \log \ell$ can guarantee existing a set of chromosomes that can use the technique used in layered


doi:10.6342/NTU201902839

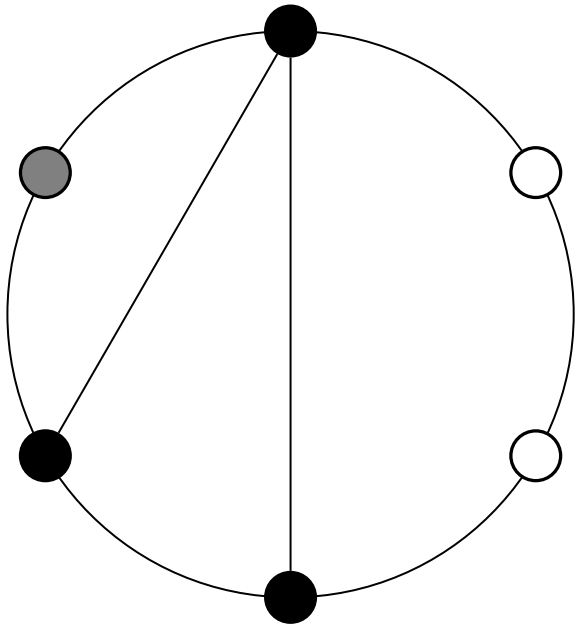


Figure 6.7: Example of a partition of a small-world topology. Circles with the same color filled belond to the same part.

structures to form the segment the global optimum holds in $\mathrm{O}(b_i)$ nFEs with probability $1-\Theta(\ell^{-\alpha})$ with the aid of oracles. Therefore, with a population of $\Theta(\alpha\ell)$ chromosomes, the probability of not existing a set of chromosomes that can use the technique used in layered structures to form the segment the global optimum holds in $\mathrm{O}(b_i)$ nFEs with the aid of oracles is $\omega(\ell^{-\alpha})$. Since each parts are all distinct and by using union bound, the population can be shared to find all required schemas simultaneously, a population of $\Theta(\alpha\ell)$ chromosomes suffices to provide a chromosome having all the alleles of the bits used in the bridges being the same as the ones of the global optimum and for all chains to find a set of chromosomes that can use the technique used in layered structures to form the segment the global optimum holds in a total of $\mathrm{O}(\ell)$ nFE with tha aid of an oracle with probability $1-\Theta(\ell^{-\alpha})$. Since there may exist a more method requiring less chromosomes, the bound on the population is an upper bound, and the theorem statement holds. □

Since the difficulty of small-world topology is relative to the number of bridges, Theorem 6.9 gives an upper bound on the number of bridges, in most cases, around logarithm of the chromosome length, such that the population does not exceed the nFE with the help of an oracle, meaning it is potentially practical. Also, note that in the proof of Theorem 6.9,


doi:10.6342/NTU201902839

the number of bits used in bridges instead of the number of bridges is the factor affecting the bound of the population size. Thus, even with the same number of bridges, the bound may be different. Another possible implication of the result is that the model builders in modern GAs using OM is near-optimal, since by adding about $\log \ell$ fake sub-functions, the requirement of population size with the aid of an oracle may grow to linear, while in practice, the observed required population size is sub-linear.


doi:10.6342/NTU201902839



doi:10.6342/NTU201902839

# Chapter 7

# Conclusion

In this thesis, population sizing for binary-encoded GAs using OM in the view of supply is discussed. The concept of oracles is adopted to have the success rate of GAs solely relies on supply. For the most basic supply problem, the expected value and the lower bound are derived, and both of them are of the same order as the bound derived by Goldberg *et al.* is (Goldberg *et al.*, 2001). Next, the population size required for binary-encoded GAs using OM with a $c$-composite oracle is estimated. Tight bounds on 1-composite oracle, a special case, and on $c$-composite oracle, which represents a more general case, are derived. Also, a global upper bound on supply is derived with some discussion on the tightness of the bound. For problem dependent cases, bounds on the population as well as on the nFE are derived for problems with their fitness that can be viewed as a layered structure. Fitness functions with torus and cube topologies are estimated as some special cases, showing the applicability of layered structures. Fitness functions with small-world topologies are also estimated as a more realistic example.

To further extend this work, the following approaches are suggested. In this work, supply is estimated on fitness functions being the sum of sub-functions with two bits. Thus, estimating fitness functions being the sum of sub-functions with more than two bits can be viewed as a choice. Also, most results are based on Theorem 4.3, which discusses the bound on the population size when using a 1-composite oracle. A more thorough investigation on general $c$-composite oracles is also an interesting topic. Last but not least, estimations on the supply for more realistic problem structures, such as random


doi:10.6342/NTU201902839

graphs with all nodes having fixed degrees, are also desired.

The contributions of this thesis are listed below. Firstly, this thesis is the first one to discuss the supply model for non-fully separable overlapping problem structures. Secondly, bounds on the supply for problem dependent and problem independent cases are derived. General upper bound and lower bound are both derived. For more restricted problem independent cases, the bounds are tight for binary-encoded GAs using OM with $c$-composite oracles provided that $c$ is a constant and that the largest size of the masks is a constant. Bounds on the supply for fitness functions with ring topologies are tight, while for fitness functions with torus and cube topologies, only upper bounds are derived. Last of all, in this work, the concept of the oracle is adopted to the population sizing of GAs. With the aid of oracles, bounds on the population size and nFEs for GAs using OM on both problem dependent and independent cases are developed. By using this framework, the next step would be solving the complexity issues on the automaton, which corresponds to the convergence time of GAs.


doi:10.6342/NTU201902839

# Bibliography


Bosman, P. A. N. and Thierens, D. (2012). Linkage neighbors, optimal mixing and forced improvements in genetic algorithms. In *Proceedings of the 14th Annual Conference on Genetic and Evolutionary Computation*, GECCO '12, pages 585–592, Philadelphia, Pennsylvania, USA. ACM.

Cover, T. T. and Thomas, J. A. (2006). *Elements of Information Theory*. Wiley-Interscience, New York, NY, USA.

de Moivre, A. (1711). *De mensura sortis: seu, De probabilitate eventuum in ludis a casu fortuito pendentibus*, pages 213–264. Clements, London, England.

Deb, K. and Goldberg, D. E. (1993). Analyzing deception in trap functions. In *Foundations of Genetic Algorithms*, volume 2, pages 93 – 108. Morgan Kaufmann.

Goldberg, D. E., Deb, K., and Clark, J. H. (1992). Genetic algorithms, noise, and the sizing of populations. *Complex Systems*, 6.

Goldberg, D. E., Deb, K., and Thierens, D. (1993). Toward a better understanding of mixing in genetic algorithms. *Journal of the Society of Instrument and Control Engineers*, 32(1):10–16.

Goldberg, D. E., Sastry, K., and Latoza, T. (2001). On the supply of building blocks. In *Proceedings of the 3rd Annual Conference on Genetic and Evolutionary Computation*, GECCO '03, pages 336–342, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.




doi:10.6342/NTU201902839

Harik, G., Cantú-Paz, E., Goldberg, D. E., and Miller, B. L. (1997). The gambler's ruin problem, genetic algorithms, and the sizing of populations. In *Proceedings of 1997 IEEE International Conference on Evolutionary Computation*, ICEC '97, pages 7–12, Indianapolis, IN, USA. IEEE.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. University of Michigan Press, Ann Arbor, Michigan, USA.

Hsu, S.-H. and Yu, T.-L. (2015). Optimization by pairwise linkage detection, incremental linkage set, and restricted / back mixing: DSMGA-II. In *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, GECCO '15, pages 519–526, Madrid, Spain. ACM.

Liao, Y.-Y., Hsu, H.-W., Juang, Y.-L., and Yu, T.-L. (2019). On the investigation of population sizing of genetic algorithms using optimal mixing. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '19, pages 820–828, Prague, Czech Republic. ACM.

Luong, N. H., Poutré, H. L., and Bosman, P. A. N. (2018). Exploiting linkage information and problem-specific knowledge in evolutionary distribution network expansion planning. *Evolutionary Computation*, 26(3):471–505.

Mühlenbein, H. and Paaß, G. (1996). From recombination of genes to the estimation of distributions i. binary parameters. In Voigt, H.-M., Ebeling, W., Rechenberg, I., and Schwefel, H.-P., editors, *Parallel Problem Solving from Nature — PPSN IV*, pages 178–187, Berlin, Heidelberg. Springer.

Orphanou, K., Thierens, D., and Bosman, P. A. N. (2018). Learning Bayesian network structures with GOMEA. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '18, pages 1007–1014, Kyoto, Japan. ACM.

Pelikan, M., Sastry, K., and Goldberg, D. E. (2006). Sporadic model building for efficiency enhancement of hierarchical BOA. In *Proceedings of the 8th Annual Confer-*


doi:10.6342/NTU201902839

*ence on Genetic and Evolutionary Computation*, GECCO '06, pages 405–412, Seattle, Washington, USA. ACM.

Pelikan, M., Sastry, K., Goldberg, D. E., Butz, M. V., and Hauschild, M. (2009). Performance of evolutionary algorithms on NK landscapes with nearest neighbor interactions and tunable overlap. In *Proceedings of the 11th Annual Conference on Genetic and Evolutionary Computation*, GECCO '09, pages 851–858, Montreal, Québec, Canada. ACM.

Thierens, D. (2010). The linkage tree genetic algorithm. In *Parallel Problem Solving from Nature, PPSN XI*, pages 264–273, Berlin, Heidelberg. Springer.

Thierens, D. and Bosman, P. A. N. (2011). Optimal mixing evolutionary algorithms. In *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, GECCO '11, pages 617–624, Dublin, Ireland. ACM.

Tung, Y.-F. (2015). Theoretical perspective of convergence complexity of evolutionary algorithms adopting optimal mixing. Master's thesis, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan (R.O.C.).

Virgolin, M., Alderliesten, T., Witteveen, C., and Bosman, P. A. N. (2017). Scalable genetic programming by gene-pool optimal mixing and input-space entropy-based building-block learning. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '17, pages 1041–1048. ACM.

Yu, T.-L., Sastry, K., Goldberg, D. E., and Pelikan, M. (2007). Population sizing for entropy-based model building in discrete estimation of distribution algorithms. In *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, GECCO '07, pages 601–608, London, England. ACM.



doi:10.6342/NTU201902839