On Modeling a Complex System with Interacting Components

Chartchai Leenawong

Department of Mathematics and Computer Science
Faculty of Science
King Mongkut's Institute of Technology Ladkrabang
Chalongkrung Road, Ladkrabang
Bangkok, 10520 THAILAND
email: klchartc@kmitl.ac.th

Abstract

This paper presents an approach for modeling a system with complex interactions among the components, in which applications can be found in various fields such as biology, physics, business, and elsewhere. In the general setting, a *system* consists of a finite number of *parts*. For each part, it is necessary to choose one of a finite number of interchangeable *components* so as to maximize the *performance* of the system, which depends also on the interactions among the chosen components. The proposed models include controllable parameters whose values reflect the system size and the amount of interaction among the components. Computer simulation and analytical arguments are used to obtain results about the expected performance of such a system and the effects of the interactions among the components on that quantity.

1 Introduction

It is not difficult to find examples of a complex system with interacting components in various fields of study. This paper discusses combinatorial optimization models for studying such systems. In general, a **complex system** is a *system* composed of a finite number of parts. For each *part*, it is necessary to choose one of a finite number of interchangeable *components* that then interact with each other in complex ways that often cannot be measured. One objective in designing such a system is to choose, for each part, one of the available components in such a way that the resulting system is the *best*, according to a specific measure of *performance*. The performance of a complex system is assumed to be a combination of the contributions of each component which, in turn, depend on how the components interact with each other.

One example of such a complex system arises in the study of *chromosome evolution* in biology [Kauffman and Levin (1987)], where one particular mathematical model was first developed. In this setting, a *chromosome* (the " system") consists of a finite number of gene locations. The various gene locations on the chromosome are called *loci* (the "parts"). At each locus, one of a finite number of different possible versions of the gene, called *alleles* (the "components"), is present. A key question in this biological setting is how evolution selects the allele at each locus so as to obtain a chromosome with the best *fitness* (the "performance"), where the fitness of a chromosome is based on how the selected alleles interact with each other.

Another example arises in physics in the study of spin glasses [Derrida (1981)]. Here, the system consists of a number of contiguous *atoms* (the "parts"). For each atom, it is possible to select *a spin up* or *a spin down* (the "components"). The total *energy* (the "performance") of the atoms depends on how the selected spins interact with each other. The objective is to determine the spin of each atom so that the resulting ensemble has the least total energy.

Such a system also arises in the study of building a *team* (the "system") for performing a task in an organization. The team consists of a number of *job positions* (the "parts"), each of which can be filled with one of a number of *qualified individuals* (the "components") who then interact with each other. These interactions could be stimulation from other individuals, learning from discussion, or competing with each other. The *performance* of such a team is based on how the selected individuals interact with each other. The objective of the problem is to determine who to choose in each position so that the resulting team has the best performance.

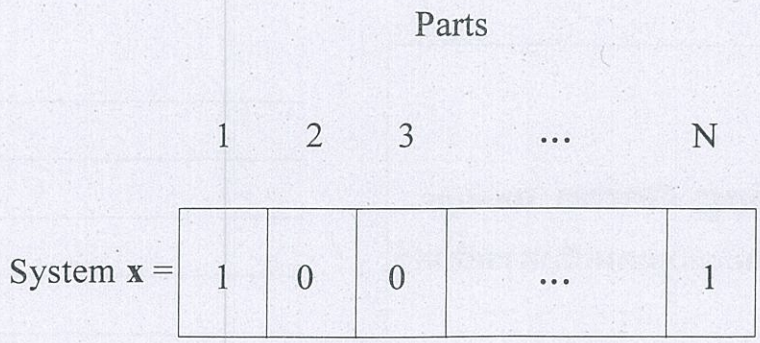
To avoid introducing new terminology, all subsequent discussions are described in the context of general complex systems.

2 Determining The System Performance

This problem of finding the complex system with the best performance is one of the combinatorial optimization problems. It seems easy to solve at first glance—just assign the component that contributes the most to each part. However, due to the interactions among the components on the system, assigning the best component to each part does not guarantee that the system achieves the best performance. An important first step in addressing the problem is to specify how the system performance is measured.

An alternative approach for computing the performance of a general complex system is modified from a model originated in the study of chromosome evolution and hereafter called *the NK model*. For a better understanding, some simplifications on the model are identified as follows:

1. There are only two components available for each part in the system.



$$0 \leq p(\mathbf{x}) = \text{Performance of system } \mathbf{x} \leq 1$$

Figure 1: A System $\mathbf{x}$ as a Binary N-vector.

2. All components interact with each other with the same strength. That is, if the performance of the component in part i depends on the components in parts j and k, then the influence of component j on component i is the same as the influence of component k on component i.

With these assumptions, a system with N parts, as described above, can be represented as a binary N-vector, $\mathbf{x} = (x_1, \ldots, x_N)$, in which $x_i = 0$ means that one of the two available components is chosen for part i and $x_i = 1$ means that the other component is chosen for that part, as shown in Figure 1. Geometrically, each of the 2^N binary N-vectors corresponds to a corner point of the N-dimensional unit cube, as shown in Figure 2 for $N = 3$.

Each fixed choice of components in the N parts results in a system $\mathbf{x}$ whose relative performance is modeled as a real number, $p(\mathbf{x})$, between 0 and 1. A value close to 0 indicates a system with relatively poor performance and a value close to 1 indicates a system with relatively good performance. It is assumed that the contribution to the system performance of part i, namely, $p_i(\mathbf{x})$, depends on the component in part i and the components in K other parts on the system ($0 \leq K \leq N-1$), say, the $K/2$ parts on either side of part i, wrapping around if necessary. Thus, $K = 0$ indicates that the contribution to the system performance of part i depends only on the component in part i and $K = N-1$ indicates that the contribution to the system performance of part i depends on the component in part i and also on the components in all other $N-1$ parts of the system.

In general, there are 2^{K+1} possible combinations for the components at the $K+1$ parts that affect part i, so the value of $p_i(\mathbf{x})$ is defined to be one of 2^{K+1} uniform

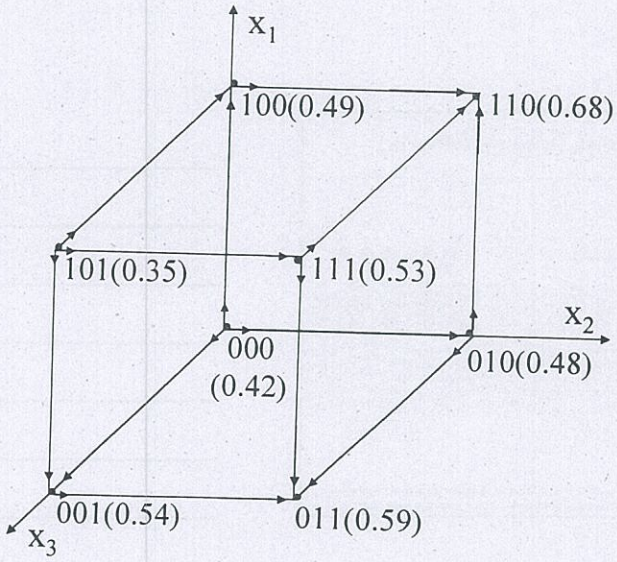


Figure 2: A System $\mathbf{x}$ and Its Performance as a Corner Point of the N-dimensional Unit Cube When $N = 3$.

$0-1$ random numbers—the one that corresponds to the combination of components in part i and the K parts that affect part i. The performance, $p(\mathbf{x})$, of a system $\mathbf{x}$ is then taken to be the average of these contributions:

$$p(\mathbf{x}) = \frac{\sum_{i=1}^{N} p_i(\mathbf{x})}{N} \tag{1}$$

Given values for N, K, and the N tables of 2^{K+1} uniform $0-1$ random numbers, the collection of all 2^N binary N-vectors, together with their performance values, as defined by (1), constitute the NK model. The objective is to find a *global maximum*, that is, a system whose performance is better than the performances of all other systems.

Finding a global maximum in this type of problem has been proved to be NP-complete [Solow et al. (2002)]. However, polynomial algorithms are developed for the special cases when K is independent of N and when K grows with N in such a way that 2^K is a polynomial in N.

To be more specific, when $K = 0$, the algorithm is, for each part, choose the component whose contribution is the largest. For other fixed values of K, the algorithm requires solving 2^K longest paths in an appropriate directed network. Even with these algorithms, finding the best system is possible only for small values of N and K. For other values of N and K, efficient heuristics can be developed to find a *local maximum*, that is, a system whose performance, though not necessarily optimal, is relatively good. An example of such heuristics is presented in the next section.

3 Simulating the Complex Systems

In this section, computer simulations using C++ programming are conducted to generate these systems with complex interactions among the components. First, a heuristic for obtaining a relatively good solution is suggested. Then, the design of the computer experiment is presented, followed by its experimental results.

3.1 A Heuristic

In the context of a complex system, a heuristic modified from mutation process in biology and hereafter called *the one-replacement heuristic* involves searching *one-replacement neighbors*. A **one-replacement neighbor** (or simply a 'neighbor' when the context is clear) of a system $\mathbf{x}$ is a system $\mathbf{y}$ in which the component at exactly one part i of $\mathbf{y}$ is different from the component in part i of $\mathbf{x}$, all other components being the same. For example, in Figure 2, the one-replacement neighbors of system $\mathbf{x}$=100 are system $\mathbf{y}_1 = 000$, $\mathbf{y}_2 = 110$ and $\mathbf{y}_3 = 101$. The performance of a one-replacement neighbor $\mathbf{y}$ may or may not be better than the performance of $\mathbf{x}$. The idea of this one-replacement heuristic proceeds as follows.

Step 1: Choose an arbitrary system $\mathbf{x}$ and compute its performance.

Step 2: Search all one-replacement neighbors of $\mathbf{x}$ in an attempt to find a system $\mathbf{x}'$ with better performance than $\mathbf{x}$. If there is no such neighbor $\mathbf{x}'$, stop, the current system $\mathbf{x}$ is a local maximum. Otherwise, select a neighbor $\mathbf{x}'$ randomly whose performance is better than that of the current system and go to Step 3.

Step 3: Set $\mathbf{x} = \mathbf{x}'$ and repeat Step 2.

To illustrate this heuristic, look back at the example in Figure 2. An arrow connecting a system $\mathbf{x}$ to a system $\mathbf{y}$ means that system $\mathbf{y}$ is a one-replacement neighbor of the system $\mathbf{x}$ that has a better performance than $\mathbf{x}$. A system $\mathbf{x}^*$ with no arrow pointing out means that $\mathbf{x}^*$ is a local maximum, that is, the performance of that system is better than the performance of all its neighbors. Thus, starting from the system 001, with performance 0.54, the one-replacement heuristic leads to the local maximum 011, with performance 0.59.

This heuristic selects a neighbor with better performance randomly from among all the neighbors of $\mathbf{x}$ having better performance than $\mathbf{x}$. Thus, if one started from the system 101, with performance 0.35, the heuristic can lead either to the local maximum 011 or to the local maximum 110, which is the global maximum in this example, with performance 0.68.

3.2 Computer Experiments

One interesting question regarding this problem is how the amount of interaction among the components affects the system performance. To answer it, computer experiments using C++ programming are conducted to simulate these complex systems.

More specifically, for the system size $N = 96$ and the amount of interaction K varies from 0 to $N-1$, 500 problems are generated randomly. The objective is to determine the average performance of a local maximum system obtained from the NK model for different values of K. Therefore, for each specific value of K, the one-replacement heuristic is used on the components to obtain a local maximum on each of the 500 randomly-generated problems and then compute an average of these local maxima.

The assumption of the K other parts that affect the contribution of component x_i is defined to be the $K/2$ parts on either side of part i, wrapping around when necessary. The new system $\mathbf{x}'$ is then chosen randomly from all the one-replacement neighbors of the current system $\mathbf{x}$ that yield better performance than $\mathbf{x}$. The results presented in Figure 3 compare the expected performance of a local maximum system in the NK model for different amounts of interaction among the components when $N = 96$. Similar results can also be obtained for other values of N. The following observation and conclusions can be drawn from Figure 3:

- When $K = 0$ or no interaction among the components, the expected performance of a local maximum system is about 0.66 or 2/3.
- When $K = N-1$ or there is full interaction among the components, the expected performance of a local maximum system tends to converge to some certain value.
- When the amount of interaction K is small, the expected performance of a local maximum system exceeds the performance of 2/3 associated with $K = 0$.
- As K keeps increasing, the expected performance of a local maximum system decreases toward the performance of the system when $K = N - 1$. This phenomenon—of decreasing performance associated with increasing interaction—is referred to as the **complexity catastrophe**.

4 The Analysis

In this section, some analytical arguments are provided on two special extreme cases. More precisely, the two cases are when $K = 0$ (no interaction among the components) and when $K = N - 1$ (full interaction). The following theorem provide the ability to find the expected performance of a local maximum system in the NK model when $K = 0$, that is, every component is independent of each other.

Theorem 1 *In the NK model, for any system size N, the expected performance of a local maximum system is 2/3 when $K = 0$.*

It is not difficult to see that this statement is true. Because when $K = 0$, the contribution to system performance of each part only depends on the component chosen from the two available for that part, in which the contribution of each of these

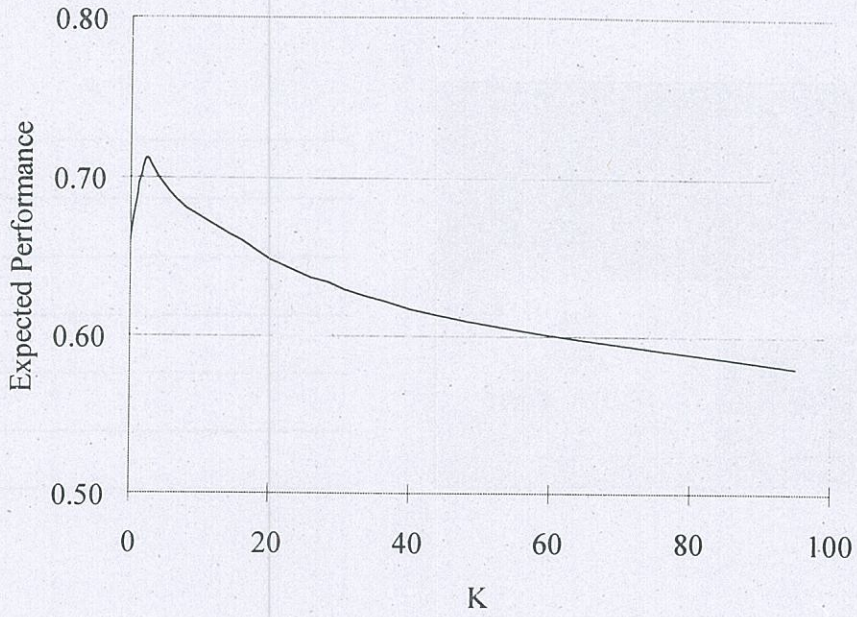


Figure 3: The Expected Performance of Local Maxima in the NK Model When $N = 96$.

two components is generated from a uniform distribution between 0 and 1, the expected contribution of each part is then equal to the expected value of the maximum of the two uniform $0-1$ random numbers, which can be easily shown to be $2/3$. Consequently, when the system performance is computed as an average of all the contributions as in the NK model, the expected system performance is equal to the expected contribution of each part, namely, $2/3$.

The next theorem provides similar conclusion about the expected performance of a local maximum system in the NK model for the other extreme case when $K = N-1$, that is every component depends on every other.

Theorem 2 *In the NK model, for a large system (N approaches ∞), the expected performance of a local maximum system is $1/2$ when $K = N - 1$.*

Here the full interaction among the components is present. When the component in a part is replaced with the other component available for the part, not only the contribution of that part is affected but also the contribution of the component in every other part. Hence, for each replacement made, the contribution of every part is changed as well. Accordingly, for large N, the expected system performance is the expected value of an average of N uniform $0-1$ random numbers, which is $1/2$.

Note here that these two theorems agree with the computer simulation results presented in Figure 3.

5 Conclusions and Future Research

An approach for modeling a general complex system with interacting component has been discussed. More work and applications about this complex system can be found in Macken and Perelson (1989), Macken et al. (1991), Bak et al. (1992), Flyvbjerg and Lautrup (1992), Kauffman (1993), Perelson and Macken (1995), and Levinthal (1997). In addition, in an attempt to make the model more realistic, some future research ideas are suggested as follows.

- **Concept of a Leading Component**: For some applications of this complex system, it might be possible that the system needs an additional component to lead every other component toward the same goal of maximizing the system performance, then that special component will affect the contributions to the system performance of all the remaining components, while they may (or may not) affect the contribution of the leading component. The model presented in this paper does not include the concept of this leading component.

- **The Strength of Interaction**: In this paper, it is assumed that the strength of interaction among different components in the system is all the same, which might not be true in some cases. For such cases, it might be possible to modify the NK model to include the concept of strength of interaction so that the interactions between components can be different. That is, if the contribution to performance of component i is affected by components j and k, then the degree to which component j affects the contribution of component i can be different from that of component k.

References

Bak, P., Flyvbjerg, H., and Lautrup, B. (1992), 'Coevolution in a rugged fitness landscape', *Physical Review A*, **46**, 6724-6730.

Derrida, B. (1981), 'Random-Energy Model: an Exactly Solvable Model of Disordered Systems', *Physical Review B*, **24** 2613-2626.

Flyvbjerg, H. and Lautrup, B. (1992), 'Evolution in a rugged fitness landscape', *Physical Review A*, **46**, 6714-6723.

Kauffman, S. A. (1993), *The Origins of Order*, Oxford University Press, Oxford.

Kauffman, S. A. and Levin, S. (1987), 'Towards a general theory of adaptive walks on rugged landscapes', *Journal of Theoretical Biology*, **128**, 11-45.

Levinthal, D. A. (1997), 'Adaptation on rugged landscapes', *Management Science*, **43**, 934-950.

Macken, C. A., Hagan, P. S., and Perelson, A. S. (1991), 'Evolutionary walks on rugged landscapes', *SIAM Journal on Applied Mathematics*, **51**, 799-827.

Macken, C. A. and Perelson, A. S. (1989), 'Protein evolution on rugged landscapes', *Proceedings of the National Academy of Sciences of the United States of America*, **86**, 6191-6195.

Perelson, A. S. and Macken, C. A. (1995), 'Protein evolution on partially correlated landscapes', *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 9657-9661.

Solow, D., Vairaktarakis, G., Piderit, K. S., and Tsai, M. C. (2002), 'Managerial insights into the system-building problem from a combinatorial optimization model', under revision for *Management Science*.