On the Treewidth of NK Landscapes

Yong Gao and Joseph Culberson

Department of Computing Science, University of Alberta
Edmonton, Alberta, Canada, T6G 2E8
{ygao, joe}@cs.ualberta.ca

Abstract. The concepts of treewidth and tree-decomposition on graphs generalize those of the trees. It is well established that when restricted to instances with a bounded treewidth, many NP hard problems can be solved polynomially. In this paper, we study the treewidth of the NK landscape models. We show that the NK landscape model with adjacent neighborhoods has a constant treewidth, and prove that for $k \geq 2$, the treewidth of the NK landscape model with random neighborhoods asymptotically grows with the problem size n.

1 Introduction

NK landscapes have been widely used in the study of genetic algorithms and computational biology [1]. There are basically two classes of NK landscape models: the NK landscape model with adjacent neighborhood and the NK landscape model with random neighborhood. Both of the two models have been analyzed and characterized from the perspectives of statistics and computational complexity [2,3,4,5].

In [2], it was shown that even though the NK landscape model with adjacent neighborhoods can be solved polynomially and the NK landscape model with the random neighborhood is usually NP complete, the two classes of NK models share almost identical statistical characteristics such as the average number of local minima and the average height of the local minima. This has puzzled researchers in this field for a while. In [4], it was shown that the decision versions of NK landscapes with random neighborhoods are easy to solve with probability asymptotic to 1 under two commonly used probabilistic settings. This is more or less in contrast to the common observation that NK landscapes with random neighborhoods are usually hard for genetic algorithms.

In the study of constraint satisfiability problems and algorithmic graphs [6,7], it is well-known that problems with an underlying tree structure can be solved linearly. The concepts of treewidth and tree-decomposition of graphs generalize the concept of trees and measure the degree to which a graph behaves like a tree [7]. It is well established that many NP complete problems, when restricted to instances with a bounded treewidth structure, can be solved polynomially via dynamic programming or some other deterministic algorithms [8].

In this paper, we study the treewidth of NK landscapes in an effort to further understand the differences between the two classes of NK models and the reasons why they appear to be hard for genetic algorithms.

E. Cantú-Paz et al. (Eds.): GECCO 2003, LNCS 2723, pp. 948–954, 2003.

In Section 2, we introduce the NK landscape model, its interaction graph, and the concepts related to the treewidth and tree decomposition of graphs. In Section 3, we study the treewidth of the NK landscape models and prove that the adjacent neighborhood NK landscape model has a fixed treewidth and the treewidth of the random neighborhood NK landscape model grows linearly with the problem size n. In Section 4, we discuss the implications of our results and future work.

2 NK Landscape Models

An NK landscape

$$f(x) = \sum_{i=1}^{n} f_i(x_i, \Pi(x_i)), \tag{1}$$

is a real-valued function defined on binary strings of fixed length, where $n > 0$ is a positive integer and $x = (x_1, \cdots, x_n) \in \{0,1\}^n$. It is the sum of n *local fitness functions* f_i, $1 \leq i \leq n$. Each local fitness function $f_i(x_i, \Pi(x_i))$ depends on the *main variable* x_i and its neighborhood

$$\Pi(x_i) \subset \mathcal{P}_k(\{x_1, \cdots, x_n\} \backslash \{x_i\}) \tag{2}$$

where $\mathcal{P}_k(X)$ denotes the set of all subsets of size k from X. The most important parameters of an NK landscape are the number of variables n, and the size of the neighborhood $k = |\Pi(x_i)|$.

In an NK landscape, the neighborhood $\Pi(x_i)$ can be chosen in two ways: the *random neighborhood*, where the k variables are randomly chosen from the set $\{x_1, \cdots, x_n\} \backslash \{x_i\}$, and the *adjacent neighborhood*, where k variables with indices nearest to i (modulo n) are chosen. To simplify the discussion, we assume in this paper that the adjacent neighborhoods are defined as follows: for each i,

$$\Pi(x_i) = (x_{max(0,i-\lceil \frac{k}{2} \rceil)}, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{min(n,i+\lceil \frac{k}{2} \rceil)}). \tag{3}$$

We use $A(n,k)$ to represent the NK landscape model with adjacent neighborhood and $N(n,k)$ to represent the NK landscape model with the random neighborhood.

Definition 1. *The interaction graph of an NK landscape model is a graph $G(V,E)$ where the vertex set $V = \{x_1 \cdots, x_n\}$ corresponds to the set of variables in the NK landscape and $(x_i, x_j) \in E$ if and only if x_i and x_j both appear in a local fitness function.*

The interaction graph of an NK landscape model captures all the interactions among the variables in the NK landscapes. A knowledge of these interactions is critical in understanding the complexity and designing appropriate algorithms to solve the problems. For example, if the interaction graph is a tree, then a linear time algorithm readily exists to solve the problem. As yet another example, if the underlying graph can be decomposed into several connected components, then

a viable approach to the problem is to first solve the subproblems represented by each connected component and then combine the partial solutions together.

The concept of treewidth and tree decomposition generalizes the above ideas further. Let us start with the definition of the l-tree.

Definition 2. *([7]) l-Trees are defined recursively as follows:*

1. *A clique with l+1 vertices is an l-tree;*
2. *Given an l-tree T_n with n vertices, an l-tree with $n+1$ vertices is constructed by adding to T_n a new vertex which is made adjacent to an l-clique of T_n and non-adjacent to the rest of the vertices.*

Definition 3. *([7]) A graph is called a partial l-tree if it is a subgraph of an l-tree. The treewidth of a graph G is the minimum value l for which G is a partial l-tree.*

The treewidth of a graph has an equivalent definition based on the concept of tree decomposition.

Definition 4. *([7]) A tree decomposition of a graph $G = (V, E)$ is a pair $D = (S, T)$ where $S = \{X_i, i \in I\}$ is a collection of subsets of vertices of G and $T = (I, F)$ is a tree with one node for each subset of S, such that*

1. *$\bigcup_{i \in I} X_i = V$,*
2. *for all the edges $(v, w) \in E$ there exists a subset $X_i \in S$ such that both v and w are in X_i, and*
3. *for each vertex v, the set of nodes $\{i, v \in X_i\}$ forms a subtree of T.*

The width of the tree decomposition $D = (S, T)$ is $max_{i \in I}(|X_i| - 1)$. And the treewidth of a graph is the minimum width over all tree decompositions of the graph.

3 The Treewidth of the NK Landscape Models

In this section the treewidth of the NK landscapes models is studied. We start with the treewidth of the NK landscape model with adjacent neighborhoods. In [2], it has been shown that the NK landscape model with adjacent neighborhoods can be solved by dynamic programming in linear time. The following theorem shows that the interaction graph of the NK landscape model with adjacent neighborhoods has a treewidth independent of n.

Theorem 1. *Let $A(n, k)$ be the NK landscape model with adjacent neighborhoods with the underlying graph G. Then, the treewidth of G is at most $2k$.*

Proof. By direct construction, we can get a tree decomposition with a width k if the cyclic interactions at the boundaries are ignored. When taking into account the cyclic interactions at the boundaries, we can get a tree decomposition with a width $2k$. □

We now turn to the NK landscape model with random neighborhoods. Since the problem is in general NP hard, we do not expect its interaction graph to have a bounded treewidth because that will mean the problem is polynomially solvable. Instead, we are interested in how the treewidth changes as the problem size n and the interaction size k increase, and the probability with which the treewidth remains small enough for algorithms making use of treewidth-related information to work efficiently. Our result below, however, shows that the treewidth asymptotically grows linearly with n.

Definition 5. *([7]) Let $G(V, E)$ be a graph with $|V| = n$. A partition (S, A, B) of V is a balanced l-partition if the following conditions are satisfied:*

1. $|S| = l + 1$;
2. $\frac{1}{3}(n - l - 1) \leq |A|, |B| \leq \frac{2}{3}(n - l - 1)$; *and*
3. *S separates A and B, i.e., there are no edges between vertices of A and vertices of B.*

Theorem 2. *Let $w(n, k)$ be the treewidth of the interaction graph of the NK landscape model with random neighborhoods. Then, for $k \geq 2$, there is a fixed constant $\delta > 0$ such that*

$$\lim_{n} Pr\{w(n, k) \leq \delta n\} = 0. \tag{4}$$

Proof. Let $l = w(n, k)$. It is well-known that if a graph has a treewidth l, then the graph must have a balanced l-partition [7]. Consider the interaction graph $G = G(V, E)$ of the NK landscape with random neighborhoods. Let $\mathcal{P}$ be the set of all the partitions of the vertex set V that satisfies the first two conditions in Definition 5. For a given $P = (S, A, B) \in \mathcal{P}$, define a random variable I_P as follows:

$$I_P = \begin{cases} 1, \text{ if } P \text{ is a balanced partition;} \\ 0, \text{ otherwise.} \end{cases} \tag{5}$$

and let O be the event that I_P is 1, i.e., that there are no edges between vertices of A and vertices of B. Recall that $\Pi(x_i)$ is the set of neighbors of the i-th local fitness function. For each $1 \leq i \leq n$ with $x_i \in A$ (or $x_i \in B$), let O_i be the event that $\Pi(x_i) \subset A \bigcup S$ ($\Pi(x_i) \subset B \bigcup S$ respectively). For $x_i \in S$, let O_i be the event that $\Pi(x_i) \subset A \bigcup S$ or $\Pi(x_i) \subset B \bigcup S$. Then, by the definition of the NK landscape model with random neighborhoods and its interaction graph, we have

$$O = \bigcap_{1 \leq i \leq n} O_i. \tag{6}$$

Since each local fitness function selects its neighbors independently, $O_1, \cdots, O_n$ are mutually independent. We have

$$Pr\{O\} = \prod_{i=1}^{n} Pr\{O_i\}. \tag{7}$$

For $x_i \in A$ (or $x_i \in B$), we have

$$P\{O_i\} \leq \frac{\binom{\frac{2}{3}(n-l-1)+l}{k}}{\binom{n-1}{k}} = (\frac{1}{3})^k(2+\frac{l}{n-1})^k. \tag{8}$$

Similarly, for $x_i \in S$, we have

$$Pr\{O_i\} \leq 2(\frac{1}{3})^k(2+\frac{l}{n-1})^k. \tag{9}$$

Then,

$$Pr\{O\} \leq 2^{l+1}(\frac{1}{3})^{kn}(2+\frac{l}{n-1})^{kn}. \tag{10}$$

Let $I = \sum_{P\in\mathcal{P}} I_P$. By its definition, we have

$$\begin{aligned} |\mathcal{P}| &= \binom{n}{l+1} \sum_{\frac{1}{3}(n-l-1)\leq a\leq \frac{2}{3}(n-l-1)} \binom{n-l-1}{a} \\ &\leq \binom{n}{l+1} 2^{n-l-1}. \end{aligned} \tag{11}$$

It follows that the expectation of I satisfies

$$\begin{aligned} \mathcal{E}\{I\} &= \sum_{P\in\mathcal{P}} \mathcal{E}\{I_P\} \\ &\leq \binom{n}{l+1} 2^{n-l-1} 2^{l+1} (\frac{1}{3})^{kn} (2+\frac{l}{n-1})^{kn} \\ &\leq \binom{n}{l+1} 2^n (\frac{2}{3}+\frac{1}{3}\frac{l}{n-1})^{kn}. \end{aligned} \tag{12}$$

Let $0 < y = \frac{l+1}{n} < 1$. We obtain from Stirling's formula that

$$\binom{n}{l+1} \sim \frac{1}{\sqrt{2\pi y(1-y)n}} \left(\frac{1}{y^y(1-y)^{1-y}}\right)^n. \tag{13}$$

And hence,

$$\mathcal{E}\{I\} \leq \frac{1}{\sqrt{2\pi y(1-y)n}} \left(\frac{2}{y^y(1-y)^{1-y}} \cdot (\frac{2}{3}+\frac{1}{3}y)^k\right)^n. \tag{14}$$

Since for $k \geq 2$,

$$\lim_{y\to 0} \left(\frac{2}{y^y(1-y)^{1-y}} \cdot (\frac{2}{3}+\frac{1}{3}y)^k\right) = 2(\frac{2}{3})^k < 1, \tag{15}$$

there exists a $0 < \delta < 1$ such that

$$\lim_n \frac{1}{\sqrt{2\pi\delta(1-\delta)n}} \left(\frac{2}{\delta^\delta(1-\delta)^{1-\delta}} \cdot (\frac{2}{3}+\frac{1}{3}\delta)^k\right)^n = 0. \tag{16}$$

Therefore, we have

$$\lim_n Pr\{w(n,k) \leq \delta n\} \leq \lim_n Pr\{I > 0\} \leq \lim_n \mathcal{E}[I] = 0. \quad (17)$$

This concludes the proof. □

4 Conclusions and Future Work

As we have shown in the previous sections, the treewidth of the NK landscape is bounded by the interaction index for the adjacent neighborhood model, but grows linearly with the problem size for the random neighborhood model. In addition to the NP complete study of the random neighborhood model, our result is the first one that depicts the difference between the two statistically similar NK landscape models.

It is well-known that optimization problems with bounded treewidth can be decomposed into independent sub-problems and solved polynomially using dynamic programming techniques. This is the case for the NK landscapes with adjacent neighborhoods [2]. Other examples include the constraint satisfaction problems and the inference problem for Bayesian networks [6,9] in which the popular tree-clustering method run polynomially if the problems under consideration have a bounded treewidth. For the random neighborhood model, our result shows that algorithms that make use of the information about the structures of the interactions in the same way as the tree-clustering approach cannot solve the problem efficiently.

An interesting question that deserves further investigation is "Do genetic algorithms exploit the treewidth-related structural information? And if so, to what extent do they rely on that information to work?" We suspect that the answer to the first question is affirmative. In fact, this is best illustrated by the recent work on sampling-based genetic algorithms. Instead of using genetic operators to generate new solutions, these sampling-based algorithms generate candidate solutions by sampling some probability distributions on the solution space and update the distributions based on the information gathered as new solutions are evaluated. The probability distributions may be modelled as the product of independent distributions [10], decomposable distributions naturally obtained from the knowledge about the interaction structures[11], or Bayesian networks that are constructed from the existing candidate solutions[12]. All of these models depend on the factorization of a multivariate probability distribution into the form

$$p(x_1, \cdots, x_n) = \frac{1}{Z} \prod_{C \in \mathcal{C}} \psi_C(x_C),$$

where $p(\cdot)$ is the original distribution and $\mathcal{C}$ is a tree decomposition of a graph, the structure of which is defined (explicitly or implicitly) by the designers of the sampling-based algorithms and is believed to be able to capture the interaction

structure of the original optimization problems. The effectiveness and efficiency of these sampling-based algorithms thus depend critically on how well the factorization approximates the real tree-decomposition of the original problem, and on the width of the tree decomposition which is lower bounded by the treewidth of the original problem.

Another direction of future work is to study the treewidth of the NK landscape models by considering the number of local fitness functions as a parameter as well as the interaction index k. By relaxing the requirement that each variable is associated with a local fitness function in the current model, we can consider a more generalized model in which the number of variables that have an associated local fitness function is another tunable parameter. It would be interesting to study the treewidth of such a generalized NK landscape model.

References

1. Kauffman, S.: The Origins of Order: Self-organization and Selection in Evolution. Oxford University Press, Inc. (1993)
2. Weinberger, E.D.: NP completeness of Kauffman's NK model, a tunable rugged fitness landscape. Technical Report Working Papers 96-02-003, Santa Fe Institute, Santa Fe (1996)
3. Wright, A.H., Thompson, R.K., Zhang, J.: The computational complexity of NK fitness functions. Technical report, Department of Computer Science, University of Montana (1999)
4. Gao, Y., Culberson, J.: An analysis of phase transition in NK landscapes. Journal of Artificial Intelligence Research **17** (2002) 309–332
5. Evans, S.N., Steinsaltz, D.: Estimating some features of NK fitness landscapes. Ann. Appl. Probab. **12** (2002) 1299–1321
6. Dechter, R., Fattah, Y.: Topological parameters for time-space tradeoff. Artificial Intelligence **125** (2001) 93–118
7. Kloks, T.: Treewidth: Computations and Approximations. Springer-Verlag (1994)
8. Downey, R.G., Fellows, M.R.: Parameterized Complexity. Springer (1997)
9. Gottlob, G., Leone, N., Scarcello, F.: A comparison of structural CSP decomposition methods. Articial Intelligence **124** (2000) 243–282
10. Leung, Y., Gao, Y., Zhang, W.: A genetic-based method for training fuzzy systems. In: Proc. of the 10th IEEE International Conference on Fuzzy Systems. Volume 1., IEEE (2001) 123–126
11. Mühlenbein, H., Mahnig, T.: Convergence theory and applications of the factorized distribution algorithm. Journal of Computing and Information Technology **7** (1999) 19–32
12. Pelikan, M., Goldberg, D.E., Lobo, F.: A survey of optimization by building and using probabilistic models. Technical Report 99018, IlliGAL, University of Illinois (1999)