NORTHWESTERN UNIVERSITY

BANDWIDTH AND PROBABILISTIC COMPLEXITY

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Electrical Engineering

and Computer Science

By

Jonathan S. Turner

Evanston, Illinois

June 1982

BANDWIDTH AND PROBABILISTIC COMPLEXITY

Jonathan Turner

ABSTRACT

We study the probabilistic performance of heuristic algorithms for the NP-complete bandwidth minimization problem. Let $G=(V,E)$ be a graph with $V=\{1, \ldots, n\}$. Define the ***bandwidth*** of G by

$$\phi(G)=\min_{\tau} \max_{1 \le u < v \le n} |\tau(u)-\tau(v)|$$

where τ ranges over all permutations on V. Let A be a bandwidth minimization algorithm and let $A(G)$ denote the bandwidth of the layout produced by A on the graph G. We say that A is a *level algorithm* if for all graphs $G=(V,E)$ the layout τ produced by A on G satisfies

$$\forall u,v \in V \;\; d(\tau^{-1}(1),u) < d(\tau^{-1}(1),v) \Rightarrow \tau(u) < \tau(v)$$

The level algorithms were first introduced by Cuthill and McKee [27] and have proved quite successful in practice although it is easy to construct examples that cause them to perform poorly. Consequently worst-case analysis provides no insight into the practical success of these algorithms. In this thesis we use probabilistic analysis in order to gain an understanding of these algorithms and to help us design better algorithms.

Let $B_n^{\psi}=(U,F)$ be the graph defined by $U=\{1, \ldots, n\}$, $F=\{\{u,v\} \mid u,v \in U \wedge |u-v| \le \psi\}$, and let G be a random spanning subgraph of B_n^{ψ} in which the vertices have been randomly re-labelled. We show that if A is a level algorithm and $\ln n = o(\psi)$ then $A(G) \le 3(1+\epsilon)\phi(G)$ almost always holds, where ϵ is any positive constant. We also introduce a class of algorithms called the modified level algorithms and show that if A' is a modified level algorithm and

$\ln n = o(\psi)$ then $A'(G) < 2(1+\epsilon)\phi(G)$ almost always holds. A particular level algorithm $MLA1$ is analyzed and we show that when $\ln n = o(\psi)$ and $\psi < n/4$, $MLA1(G) < (1+\epsilon)\phi(G)$. We also study several other properties of random subgraphs of B_n^{ψ}.

CONTENTS

LIST OF FIGURES

1. Introduction

In the decade since its introduction by Cook [1] and Karp [2], the theory of *NP*-completeness has become a major topic in computational complexity. The list of known *NP*-complete and *NP*-hard problems has grown to include several hundred problems from many different contexts [3]. The theory tells us than unless deterministic polynomial time is equivalent to nondeterministic polynomial time ($P=NP$), all of these problems are intractable in the sense that the amount of resources required to solve them grows faster than any polynomial function of the problem size. Since it is considered unlikely that $P=NP$ all of these problems are thought to be intractable.

Despite this conclusion, the occurrence of *NP*-hard problems in many important applications has provided a strong motivation for finding methods of dealing with them effectively in practice. Surprisingly, for some problems algorithms are known which work quite well in practice. An important area of research is to understand why such algorithms work well and find ways to use this insight to design effective algorithms for other hard problems.

One fairly successful class of algorithms is called the *approximation algorithms*. If the problem of interest can be expressed as an optimization problem (and for many *NP*-hard problems this is a natural formulation) one can sometimes find algorithms which are guaranteed to produce solutions that are close to the optimal value, where 'close' means within a constant factor. To define the notion more precisely, we need some definitions. Let Π be some optimization problem, and let D_Π be the domain of Π, that is the set of all instances of Π. For each $I \in D_\Pi$ there is a set of candidate solutions $S_\Pi(I)$ and for each $\tau \in S_\Pi(I)$ there is an associated cost $C_\Pi(I,\tau)$. The cost of an optimal solution for an instance $I \in D_\Pi$ is denoted $OPT(I)$ and is defined as $\min_{\tau \in S_\Pi(I)} C_\Pi(I,\tau)$ if Π is a minimization problem and as $\max_{\tau \in S_\Pi(I)} C_\Pi(I,\tau)$ if Π is a maximization problem. Let A be a particular algorithm for Π and define $A(I)$ to be the value of the solution produced by A on instance I. Next

$$R_A(I) = \begin{cases} \dfrac{A(I)}{OPT(I)} & \text{if } \Pi \text{ is a minimization problem} \\ \dfrac{OPT(I)}{A(I)} & \text{if } \Pi \text{ is a maximization problem} \end{cases}$$

$R_A(I)$ is a measure of how close A comes to the optimal value for the particular instance I. Note that in general $R_A(I)$ is greater than or equal to one. We would like to keep it as close to one as possible. Define

$$R_A = \min\{r \geq 1 | R_A(I) \leq r \ \forall I \in D_\Pi\}$$

R_A is called the *absolute performance ratio* for A. Note that in general R_A may be undefined. That is, there may be no finite value of r that satisfies the conditions of the definition. If R_A is defined, then we say that A is a *worst case approximation* algorithm, since it guarantees that even in the worst case the cost of the solution produced by A will not differ from the optimal cost by more than the factor R_A. Algorithms of this sort have been found for bin packing problems [4][5][6] and the euclidean traveling salesman problem [7] among others.

These methods have three shortcomings. First, the performance bound may be too loose for a particular application. For example, Christofides' traveling salesman algorithm, can in the worst case, produce tours that are 1.5 times the length of an optimal tour. A more serious difficulty is that for some problems it is just as hard to find solutions having cost that is close to the optimal as it is to find an optimal solution. Sahni and Gonzalez [8] show that unless $P{=}NP$ there exists no polynomial time approximation algorithm for the non-euclidean traveling salesman problem. Similar results have been proved for many other problems. The last difficulty is that many algorithms that appear to be quite successful in practice are known not to be worst case approximation algorithms. Worst case analysis yields no insight into the practical success of such algorithms.

The above considerations have led some researchers to analyze algorithms from a probabilistic point of view rather than a worst case point of view. This has led to several interesting results. Karp [9] gives an algorithm for the euclidean traveling salesman problem and shows that if I is selected at random, then with high probability $R_A(I) \leq 1+\epsilon$ for any $\epsilon > 0$. In order to prove such results it is necessary to define a sequence of probability distributions $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n, \ldots,$ where $\mathbf{F}_n$ is a probability distribution over all problem instances of size n. Karp defined such a sequence for the euclidean traveling salesman problem and was able to show that

$$\lim_{n \to \infty} \mathrm{P}(R_A(I) \leq 1+\epsilon) = 1 \qquad (*)$$

when I is drawn at random from $\mathbf{F}_n$ and ϵ is any fixed constant greater than zero. Hence the algorithm has the very nice property that its performance becomes increasingly predictable as the problem size gets large. Any algorithm that satisfies (*) is called a *probabilistic optimization algorithm*. Similarly, we say that A is a *probabilistic approximation algorithm* if there is some constant $c > 1$ such that

$$\lim_{n \to \infty} \mathrm{P}(R_A(I) \leq c) = 1$$

Similar results have been reported in [10][11][12][13][14][15].

Many of the techniques for proving theorems concerning the probabilistic performance of algorithms have come from the theory of random graphs pioneered by Erdös and Renyi [16][17]. This theory permits one to predict many properties of randomly selected graphs with remarkable precision. Let $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n, \ldots,$ be a sequence of probability distributions where $\mathbf{F}_n$ is defined over graphs with n vertices. We say that a property holds for *almost all* graphs drawn from $\mathbf{F}_n$, if the probability of the property holding approaches one as $n \to \infty$. In [16], Erdös and Renyi prove that for any $\epsilon > 0$, almost all graphs with n vertices and $(\frac{1}{2}+\epsilon) n \ln n$ edges are connected, while almost all graphs with $(\frac{1}{2}-\epsilon) n \ln n$ edges are disconnected. In a similar spirit, Posa [14] shows that for large enough c, almost all graphs with $cn \ln n$ edges contain a hamiltonian circuit. Similar results are contained in [18][19][20][21][22][23][24][25].

Bandwidth in a graph is a measure of the locality of the relation represented by the graph. It is best known in the context of matrix bandwidth minimization; in this application, linear reductions in the bandwidth of a matrix can result in quadratic improvements in computational efficiency. Several bandwidth minimization algorithms were proposed in the sixties and early seventies [26][27][28][29][30], before Papadimitriou showed that the problem is *NP*-complete [31]. Garey, Graham, Johnson and Knuth later strengthened this result, showing that the problem remains *NP*-complete when restricted to free binary trees [32]. At the same time they gave a linear time algorithm to determine if a graph has bandwidth two. Saxe [33] then gave a dynamic programming algorithm which for any fixed k, could determine if a graph had bandwidth k in time $O(n^{k+1})$. Monien and Sudborough [34] showed how to reduce the time bound to $O(n^k)$. In this thesis we will be mainly interested in probabilistic analysis of efficient algorithms for reducing the bandwidth of graphs. The use of probabilistic techniques is justified by the observation that none of the polynomial time algorithms appearing in the literature is a worst case approximation algorithm. However there are algorithms which appear to be quite successful in practice. The analysis presented here gives some insight into that success. Note that while the known algorithms for this problem are not approximation algorithms there is no known proof that approximation algorithms for bandwidth minimization do not exist. As with many other graph problems, the techniques used to prove such results seem to break down in this case.

The remainder of this section presents the definitions and terminology we will need. Section 2 discusses some possible choices for probability distributions and attempts to justify the particular selection made here. Section 3 is devoted to proving that the algorithms in a particular class are probabilistic approximation algorithms for bandwidth minimization. It also gives a modification that improves their performance somewhat. Section 4 shows how the algorithms of Section 3 can be refined to yield near optimal results in many cases. It also gives some experimental results that give further insight into the comparative performance of the different algorithms. Section 5 contains several theorems concerning properties of random graphs with small bandwidth. There is also an appendix which contains results from elementary probability theory that are used in the body of the thesis together with informal proofs where appropriate.

A *graph* G consists of a vertex set $V=\{1,2,...,n\}$ and an edge set $E \subseteq \{\{u,v\} \mid u,v \in V\}$. The notation u—v means $\{i,j\}\in E$ and similarly $u \not- v$ means $\{i,j\}\notin E$. For sets of vertices X, Y, X—Y means that there is some edge joining a vertex in X to a vertex in Y. Similarly $X \not- Y$ means there is no such edge. The *neighborhood* of a vertex $v\in V$ is the set $N(v)=\{u \mid v—u\}$. The *degree* of a vertex $v\in V$ is $d(v)=|N(v)|$. The *distance* between two vertices u and v is denoted $d(u,v)$ and is defined as the length of the shortest path connecting u and v. The *diameter* of a graph G is denoted by $D(G)$ and is defined as $\max_{u,v\in V} d(u,v)$. The distance between two sets of vertices S, T is $\min_{u\in S,\, v\in T} d(u,v)$.

A *layout* τ of a graph G is a permutation on V. The *identity layout* is the identity permutation and the inverse of a layout τ is denoted τ^{-1}. The *length* of an edge $\{u,v\}$ with respect to a layout τ is $\delta_\tau(\{u,v\})=|\tau(u)-\tau(v)|$. Similarly, the length of an edge $\{u,v\}$ is $\delta(\{u,v\})=|u-v|$. The *bandwidth* of G with respect to a layout τ is denoted $\phi_\tau(G)$ and is defined as $\max_{\{u,v\}} \delta_\tau(\{u,v\})$. The bandwidth of G is defined as $\phi(G)=\min_\tau \phi_\tau(G)$. These definitions are illustrated in Figure 1. The graph shown has bandwidth 3 as can be seen from the layout shown at the right. In this drawing, the position of a vertex from left, determines the number assigned to it by the layout (thus $\tau(3)=2, \tau(5)=6$, etc.).

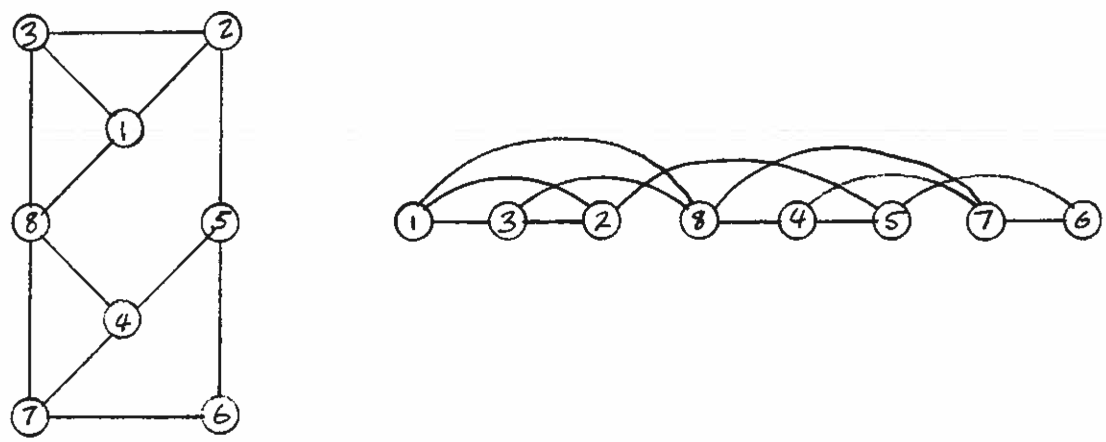

Figure 1. Graph Illustrating Bandwidth Definitions

2. Choice of Probability Distributions

The study of the probabilistic performance of algorithms requires some specification of the relative probabilities of the different instances of the problem for each problem size. When the problem instances are graphs there are several natural probability distributions that are in common use. The simplest choice is the uniform distribution Γ_n which assigns each of the $2^{\binom{n}{2}}$ graphs on n vertices the same probability of being chosen. It is not difficult to see that the following random experiment can be used to generate a graph $G=(V,E)$ in Γ_n.

1. Let $V=\{1,2,\ldots,n\}$, $E=\varnothing$.

2. For each $\{u,v\}$ $1\leq u<v\leq n$, add the edge $\{u,v\}$ to E with probability ½.

This suggests a natural generalization. If each edge is included with probability p, we get the probability distribution $\Gamma_n(p)$, which is probably the most popular choice in the literature on random graphs.

Unfortunately, $\Gamma_n(p)$ is not suitable for studying the performance of bandwidth minimization algorithms for the simple reason that almost all graphs in $\Gamma_n(p)$ have bandwidth close to n, and consequently even the most trivial algorithms qualify as probabilistic optimization algorithms. This claim is substantiated by the following results.

Lemma 2.1. Let $G=(V,E)$ be a graph on n vertices. $\phi(G)\leq n-2k \implies \exists V_1,V_2\subset V$ such that $|V_1|=|V_2|=k$ and $V_1 \not\!\!— V_2$.

proof. If $\phi(G)\leq n-2k$ then there is a layout τ such that $u — v \implies |\tau(u)-\tau(v)|\leq n-2k$. Let $V_1=\{\tau^{-1}(1),\ldots,\tau^{-1}(k)\}$ and $V_2=\{\tau^{-1}(n-k+1),\ldots,\tau^{-1}(n)\}$. If $V_1 — V_2$ then there are vertices $u\in V_1$ and $v\in V_2$ such that $u — v$. But by the definition of V_1 and V_2, $\tau(u)\leq k$ and $\tau(v)\geq n-k+1$, hence $|\tau(u)-\tau(v)|>n-2k$. This contradicts the definition of τ. □

Lemma 2.2. Let $0<p<1$ and $G=(V,E)\in\Gamma_n(p)$. $\mathrm{P}(\phi(G)\leq n-2k)\leq\left(\frac{en}{k}(1-p)^{k/2}\right)^{2k}$.

proof. By Lemma 2.1, $P(\phi(G) \le n-2k) \le P(\exists V_1, V_2$ such that $|V_1|=|V_2|=k \wedge V_1 \not\leftrightarrow V_2)$. Since there are k^2 'potential edges' joining V_1 and V_2, all of which must be absent if $V_1 \not\leftrightarrow V_2$, this last probability is

$$\le \binom{n}{k}\binom{n-k}{k}(1-p)^{k^2} \le \left(\frac{en}{k}\right)^{2k}(1-p)^{k^2} = \left(\frac{en}{k}(1-p)^{k/2}\right)^{2k} \qquad \square$$

Define $\lambda_n(c) = \frac{\ln n}{\ln c}$. Note that $\lambda_n(c) > 0$ when $0 < c < 1$ and $n > 1$, $c^{\lambda_n(c)} = \frac{1}{n}$ and $\lim_{n \to \infty} \lambda_n(c) = \infty$ for c fixed $0 < c < 1$. We will usually write $\lambda(c)$ for $\lambda_n(c)$.

Theorem 2.1. Let $0<p<1$ be fixed. For almost all $G \in \Gamma_n(p)$, $\phi(G) > n - 4\lambda(1-p)$.

proof. Applying Lemma 2.2 with $k = 2\lambda(1-p)$ gives

$$P(\phi(G) \le n - 4\lambda(1-p)) \le \left(\frac{en}{2\lambda(1-p)}(1-p)^{\lambda(1-p)}\right)^{4\lambda(1-p)} = \left(\frac{e}{2\lambda(1-p)}\right)^{4\lambda(1-p)} \to 0 \qquad \square$$

Theorem 2.2. Let $\epsilon > 0$, $c > 0$ be fixed, $p = c\frac{\ln n}{n}$. For almost all $G \in \Gamma_n(p)$, $\phi(G) \ge n(1-\epsilon)$.

proof. Applying Lemma 2.2 with $k = \epsilon n/2$ gives

$$P(\phi(G) \le n(1-\epsilon)) \le \left(\frac{en}{\epsilon n/2}(1-p)^{\epsilon n/4}\right)^{\epsilon n} \le \left(\frac{2e}{\epsilon}e^{-\epsilon np/4}\right)^{\epsilon n} = \left(\frac{2e}{\epsilon}n^{-\epsilon c/4}\right)^{\epsilon n} \to 0 \qquad \square$$

Theorems 2.1 and 2.2 show that even when the edge probability p is very small, almost all $G \in \Gamma_n(p)$ are such that $\frac{n}{\phi(G)} \to 1$. Consequently, even the most trivial bandwidth minimization algorithms (for example, the algorithm that always outputs the identity layout) qualify as probabilistic optimization algorithms. As a result, some other distribution is required if we are to make meaningful distinctions among algorithms on the basis of their probabilistic performance.

There are no hard and fast rules we can apply in selecting a probability distribution, however there are a few guiding principles that we will find useful. In general we want to avoid distributions that are 'too easy'; any distribution that classifies trivial algorithms as probabilistic optimization algorithms would fall into this category. This suggests looking for distributions that are hard in some sense. Ideally any results we prove for a hard distribution would have implications for other distributions as well. The following definitions are an attempt to arrive at a useful notion of hard distribution for the bandwidth minimization problem.

Let Λ_n be a probability distribution over n vertex graphs. We say that Λ_n is *unbiased* with respect to bandwidth if for all integers ψ, $0 \le \psi < n$, all graphs with bandwidth ψ are equally likely. We say that Λ_n is *uniform* with respect to bandwidth if for $G \in \Lambda_n$ $P(\phi(G)=\psi)=1/n$ $(0 \le \psi < n)$. Note that while $\Gamma_n(½)$ is unbiased with respect to bandwidth (in fact it is unbiased with respect to every graph invariant), it is extremely non-uniform. It is this lack of uniformity that makes it unsuitable for comparing the performance of bandwidth minimization algorithms.

Let $\Phi_n(\psi)$ be the unbiased distribution over n vertex graphs with bandwidth ψ, and let A be a bandwidth minimization algorithm. We say that A is a *universal* probabilistic optimization algorithm if

$$\lim_{n \to \infty} P(R_A(I) < 1+\epsilon) = 1$$

when I is drawn at random from any one of $\Phi_n(\psi)$ and ϵ is any fixed constant greater than zero. The adjective 'universal' is justified by the observation that any algorithm that satisfies the above condition is a probabilistic optimization algorithm for all distributions that are unbiased with respect to bandwidth. Universal probabilistic approximation algorithms can be defined similarly.

The concept of universal algorithms is an attractive one, but it is difficult to see how to apply it directly. $\Phi_n(\psi)$ doesn't appear to be amenable to analysis precisely because we have no simple characterization of bandwidth ψ graphs. Consequently we find it necessary to approximate $\Phi_n(\psi)$ with something more tractable. Let ψ be an integer $(0 \le \psi < n)$ and let $G=(V,E)$ be generated by the following random experiment.

1. Let $V=\{1,2, \ldots, n\}$. For each $\{u,v\}$ $1 \le u < v \le n$ such that $|u-v| \le \psi$ include the edge

$\{u,v\}$ in E with probability ½.

2. Randomly renumber the vertices so that each of the $n!$ possible labelings is equally likely.

We can generalize the above distribution by generating edges with probability p. The resulting distribution is denoted $\Omega_n(\psi,p)$. Note that for $G \in \Omega_n(\psi,p)$, $\phi(G) \leq \psi$. It is shown below that if $\epsilon > 0$, $0 < p < 1$ are fixed then for almost all $G \in \Omega_n(\psi,p)$, $\phi(G) \geq (1-\epsilon)\psi$. Hence $\Omega_n(\psi,p)$ approximates the uniform characteristic of $\Phi_n(\psi)$. However, strictly speaking it is not uniform, since it slightly favors graphs with small bandwidth. Nor is it unbiased since it tends to favor graphs having many minimum bandwidth layouts. One other distribution will be useful in the analysis. $\Psi_n(\psi,p)$ is defined by the first step of the random experiment that defines $\Omega_n(\psi,p)$. Obviously, if Q is any graph invariant (that is, a property that depends only on the structure of a graph) then Q occurs with exactly the same probability in $\Omega_n(\psi,p)$ and $\Psi_n(\psi,p)$.

Theorem 2.3. Let $0 < p < 1$ be fixed, $\ln n \leq \psi < n$. For almost all $G \in \Psi_n(\psi,p)$, $\phi(G) > \psi - 4\lambda_\psi(1-p)$.

proof. Let $G' \subseteq G$ be the subgraph induced by vertices $\{1,2, \ldots, \psi\}$. Note that G' is a random graph with distribution $\Gamma_\psi(p)$. Applying Theorem 2.1

$$\phi(G') > \psi - 4\lambda_\psi(1-p)$$

The theorem follows from the fact that $\phi(G) \geq \phi(G')$. □

An immediate consequence of this result is that if $\psi \to \infty$ then $\psi < (1+\epsilon)\phi(G)$ for any fixed $\epsilon > 0$. Theorem 2.3 is sufficient for the other results appearing in the body of this thesis. It is interesting however to consider tightening the relationship between ψ and $\phi(G)$.

Conjecture. Let $0 < p < 1$ be fixed. There is some constant $c > 0$ such that $c\ln n \leq \psi \leq n - c\ln n \Rightarrow$ for almost all $G \in \Psi_n(\psi,p)$, $\phi(G) = \psi$.

3. Probabilistic Approximation Algorithms for Bandwidth Minimization

A variety of heristic algorithms for bandwidth minimization appeared in the late sixties and early seventies [26][27][28][29][30]. The most successful algorithms were all members of a class of algorithms which I will refer to as *level algorithms.* The idea behind the level algorithms is to order the vertices according to their distance from some fixed vertex u. Formally, we say that an algorithm is a level algorithm if the layout τ produced for a graph $G=(V,E)$ satisfies

$$\forall u,v \in V \quad d(\tau^{-1}(1),u) < d(\tau^{-1}(1),v) \Rightarrow \tau(u) < \tau(v)$$

The various types of algorithms differ mainly in how $\tau^{-1}(1)$ is chosen, and how vertices equidistant from $\tau^{-1}(1)$ are ordered relative to each other. In this section we will consider only algorithms that try all possible choices for $\tau^{-1}(1)$. This increases the complexity of the algorithm by at most a factor of n and simplifies the analysis. In section 5 more sophisticated techniques will be considered. We also ignore the relative orderings of vertices equidistant from $\tau^{-1}(1)$ in this section, focusing instead on properties that hold for the entire class of level algorithms.

The first question to address is whether any of the level algorithms are worst case approximation algorithms. It is in fact easy to show that none of them is. Consider the tree in Figure 2. It is not difficult to see that no matter what choice is made for $\tau^{-1}(1)$, a level algorithm will always produce a layout with bandwidth ≥ 4, whereas the actual bandwidth of this tree is 2. The example is easily extended. For any value k, we can construct a tree with bandwidth 2, such that the layout produced by any level algorithm has bandwidth $\geq k$. This implies that even for the very restricted case of binary trees, the level algorithm can produce layouts with bandwidths that differ from the optimal by an arbitrarily large factor. This set of examples is due to I. H. Sudborough.

In this section we show that in spite of their poor worst case performance, the level algorithms work quite well for random graphs. This provides some theoretical basis for their success in

Figure 2. Tree Demonstrating Poor Worst-case Performance of Level Algorithms

practice.

Lemma 3.1. Let $\epsilon>0$, $0<p<1$ be fixed, $\alpha=(1+\epsilon)\lambda(1-p^2)\le\psi<n$. For almost all $G\in\Psi_n(\psi,p)$, there exists a path of length two between every pair of vertices u,v such that $|u-v|\le\psi$.

proof. The probability that two specific vertices u,v with $|u-v|\le\psi$ are not connected by a 2-path is $\le(1-p^2)^{\psi-1}$. There are $<n\psi$ such pairs. Hence, the probability that any such pair is not connected by a 2-path is

$$\le n\psi(1-p^2)^{\psi-1}\le n\alpha(1-p^2)^{\alpha-1}=n\alpha n^{-(1+\epsilon)(1-1/\alpha)}\to 0 \qquad \square$$

Before proceeding we need the following definitions. Let $G=(V,E)$ and define $V_i(u)=\{u|d(u,v)=i\}$ for all $u\in V$. Also let $V_i=V_i(1)$. Next, define $l_i(u)=\min v\in V_i(u)$ and $r_i(u)=\max v\in V_i(u)$. Let $l_i=l_i(1)$ and $r_i=r_i(1)$. Note that $|V_i|\le r_i-l_i$.

Lemma 3.2. Let $\epsilon>0$, $0<p<1$ be fixed, $(1+\epsilon)\lambda(1-p^2)\le\psi<n$. For almost all $G\in\Psi_n(\psi,p)$, $r_i-3\psi\le r_{i-3}<l_i-\psi$ for all $i\ge3$.

proof. The shortest path from 1 to r_i must pass through some $u\in V_{i-3}$. Clearly $r_i-u\le3\psi$, hence $r_i-3\psi\le u\le r_{i-3}$. To see that $r_{i-3}<l_i-\psi$, assume otherwise. Then there is some vertex v on the shortest path from 1 to r_{i-3} such that $l_i-\psi\le v<l_i$ and $d(1,v)\le i-3$. By Lemma 3.1 there is a 2-path from v to l_i, giving $d(1,l_i)\le i-1$, which is a contradiction. $\square$

We are now ready to show that with high probability $|V_i|<2\psi$, for all $i\ge0$.

Theorem 3.1. Let $\epsilon>0$, $0<p<1$ be fixed, $(1+\epsilon)\lambda(1-p^2)\le\psi<n$. For almost all $G\in\Psi_n(\psi,p)$,

$r_i - l_i < 2\psi$, for all $i \geq 0$.

proof. The result is trivial for $i \leq 2$. For $i \geq 3$, Lemma 3.2 tells us that $r_i \leq r_{i-3} + 3\psi$ and $l_i > r_{i-3} + \psi$ Hence,

$$r_i - l_i < (r_{i-3} + 3\psi) - (r_{i-3} + \psi) = 2\psi \qquad \square$$

Define

$$LEVEL(G) = \min_{u \in V} \max_{i \geq 0} |V_i(u) \cup V_{i+1}(u)|$$

and note that $LEVEL(G)$ is an upper bound on the bandwidth of the layout produced by any level algorithm on G. We are now ready to relate $LEVEL(G)$ to $\phi(G)$.

Theorem 3.2. Let $\epsilon > 0$, $0 < p < 1$ be fixed, $\ln n = o(\psi)$, $\psi < n$. For almost all $G \in \Omega_n(\psi, p)$, $LEVEL(G) < 4(1+\epsilon)\phi(G)$.

proof. By Theorem 3.2, $LEVEL(G) < 4\psi$. By Theorem 2.3 $\psi < (1+\epsilon)\phi(G)$. $\square$

From Theorem 3.2 we can conclude that the level algorithms are all probabilistic approximation algorithms for $\Omega_n(\psi, p)$ when ψ is not too small, and they usually produce layouts with bandwidth at most four times $\phi(G)$. It is also easy to show that a simple level algorithm can be implemented to run in time $O(n|E|) = O(n^2\phi(G))$. The performance bound in Theorem 3.2 can be improved by tightening the bounds on $|V_i|$.

Lemma 3.3. Let $\epsilon > 0$, $0 < p < 1$ be fixed, $\alpha = (1+\epsilon)\lambda(1-p^2) \leq \psi < n$. For almost all $G \in \Psi_n(\psi, p)$, there exists a path of length two between every pair of vertices u, v such that $|u - v| \leq 2\psi - \alpha$.

proof. Let $u, v \in V$ with $|u - v| \leq 2\psi - \alpha$. Let $i = 2\psi - |u - v|$. The probability that $d(u, v) > 2$ is $\leq (1-p^2)^i$. Since for each i there are $\leq n$ such pairs, the probability that any pair is not joined by a 2-path is

$$\leq \sum_{i=\lceil\alpha\rceil}^{2\psi-1} n(1-p^2)^i < n(1-p^2)^\alpha \sum_{i=0}^{\infty} (1-p^2)^i = p^{-2} n^{-\epsilon} \to 0 \qquad \square$$

Lemma 3.4. Let $\epsilon > 0$, $0 < p < 1$ be fixed, $\alpha = (1+\epsilon)\lambda(1-p^2) \leq \psi < n$. For almost all $G \in \Psi_n(\psi, p)$, $r_i - 3\psi \leq r_{i-3} < l_i - (2\psi - \alpha)$, for all $i \geq 3$.

proof. The shortest path from 1 to r_i must pass through some $u \in V_{i-3}$. Clearly $r_i - u \le 3\psi$, hence $r_i - 3\psi \le u \le r_{i-3}$. To see that $r_{i-3} < l_i - (2\psi - \alpha)$, assume otherwise. Then there is some vertex v on the shortest path from 1 to r_{i-3} such that $l_i - (2\psi - \alpha) \le v < l_i$ and $d(1,v) \le i-3$. By Lemma 3.3 there is a 2-path from v to l_i, giving $d(1,l_i) \le i-1$, which is a contradiction. □

Theorem 3.3. Let $\epsilon > 0$, $0 < p < 1$ be fixed, $\alpha = (1+\epsilon)\lambda(1-p^2) \le \psi < n$. For almost all $G \in \Psi_n(\psi,p)$, $r_i - l_i < \psi + \alpha$, for all $i \ge 3$.

proof. By Lemma 3.4 $r_i \le r_{i-3} + 3\psi$ and $l_i > r_{i-3} + (2\psi - \alpha)$ Hence,

$$r_i - l_i < (r_{i-3} + 3\psi) - (r_{i-3} + (2\psi - \alpha)) = \psi + \alpha \qquad \square$$

Note that Theorem 3.3 says nothing about the size of V_1 and V_2. As we shall see, these cases differ from the rest and will be handled in the next theorem. First however, we need a lemma concerning the binomial distribution, $\mathbf{B}(n,p)$. By definition if $x \in \mathbf{B}(n,p)$ then

$$\mathrm{P}(x=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

The mean and variance of x are np and $np(1-p)$ respectively. The following lemma is from Angluin and Valiant [13].

Lemma 3.5. If $x \in \mathbf{B}(n,p)$ then for any ϵ, $0 < \epsilon < 1$, $\mathrm{P}(x \le (1-\epsilon)np) < e^{-\epsilon^2 np/2}$ and $\mathrm{P}(x \ge (1+\epsilon)np) < e^{-\epsilon^2 np/3}$.

Theorem 3.4. Let $c > 0$, $\epsilon > 0$, $0 < p < 1$ be fixed, $\alpha = c \ln n \le \psi < n$. For almost all $G \in \Psi_n(\psi,p)$, $(1-\epsilon)p\psi < |V_1| < (1+\epsilon)p\psi$ and $|V_2| < (1+\epsilon)(2-p)\psi$. Also, if $\psi < n/2$ then $(1-\epsilon)(2-p)\psi - \alpha < |V_2|$.

proof. $|V_1|$ is a binomial random variable in $\mathbf{B}(\psi,p)$. By Lemma 3.5,

$$\mathrm{P}(|V_1| < (1-\epsilon)p\psi) < e^{-\epsilon^2 p\psi/2} \to 0$$

$$\mathrm{P}(|V_1| > (1+\epsilon)p\psi) < e^{-\epsilon^2 p\psi/3} \to 0$$

This establishes the bounds on $|V_1|$. Since $|V_2| \le 2\psi - |V_1|$,

$$|V_2| < 2\psi - (1-\epsilon)p\psi < (1+\epsilon)(2-p)\psi$$

Applying Lemma 3.3 gives

$$|V_2| \geq (2\psi-\alpha) - |V_1| > (2-(1+\epsilon)p)\psi - \alpha > (1-\epsilon)(2-p)\psi - \alpha$$ □

Theorem 3.5. Let $\epsilon>0$, $0<p<1$ be fixed, $\psi<n$, $\ln n = o(\psi)$. For almost all $G \in \Omega_n(\psi,p)$, $LEVEL(G) < (1+\epsilon)(3-p)\phi(G)$.

proof. By Theorems 3.3 and 3.4, $LEVEL(G) < (1+\epsilon')(3-p)\psi$ for any fixed $\epsilon'>0$. By Theorem 2.3, $\psi < (1+\epsilon')\phi(G)$. Selecting ϵ' so that $(1+\epsilon')^2 = (1+\epsilon)$ yields the theorem. □

Notice that $|V_2|$ is a lower bound on the bandwidth. From Theorem 3.4 we expect that $|V_2| \approx (2-p)\psi$. Consequently, even the best level algorithms are likely to produce layouts that differ from the optimal by a factor of roughly $(2-p)$. We can improve the performance of the level algorithms by using a different strategy for ordering the vertices in $V_1 \cup V_2$. Define

$$V'_2(u) = \begin{cases} V_2(u) & \text{if } V_3(u) = \varnothing \\ V_2(u) \cap \{v \mid v - V_3(u)\} & \text{if } V_3(u) \neq \varnothing \end{cases}$$

$$V'_1(u) = (V_1(u) \cup V_2(u)) - V'_2(u)$$

$$V'_i(u) = V_i(u) \qquad i=0,\ i \geq 3$$

Also, let $V'_i = V'_i(1)$, $l'_i(u) = \min V'_i(u)$, $r'_i(u) = \max V'_i(u)$, $l'_i = l'_i(1)$, $r'_i = r'_i(1)$. We can now define the *modified level algorithms*. Formally, A is a modified level algorithm, if the layout τ produced for the graph $G=(V,E)$ satisfies

$$\forall u,v \in V \quad u \in V'_i(\tau^{-1}(1)) \wedge v \in V'_{i+1}(\tau^{-1}(1)) \Rightarrow \tau(u) < \tau(v)$$

Let $LEVEL'(G) = \min_{u \in V} \max_{i \geq 0} |V'_i(u) + V'_{i+1}(u)|$. Clearly, $LEVEL'(G)$ is an upper bound on the bandwidth of any layout produced by a modified level algorithm on G. We will show below that for almost all $G \in \Psi_n$, $\dfrac{LEVEL'(G)}{\phi(G)} < 2(1+\epsilon)$ for any fixed $\epsilon>0$.

Lemma 3.6. Let $\epsilon>0$, $0<p<1$ be fixed, $(1+\epsilon)\lambda(1-p^2) \leq \psi < n$. For almost all $G=(V,E) \in \Psi_n(\psi,p)$, $u \in V \wedge |\{u-\psi, \ldots, u+\psi\} \cap V_i| \geq (1+\epsilon)\lambda(p) \Rightarrow u - V_i$.

proof. Note that by Lemma 3.1, for each vertex u there at most five sets V_i such that $|\{u-\psi, \ldots, u+\psi\} \cap V_i| \geq 1$. Hence, the probability that for any $G \in \Psi_n(\psi, p)$, the assertion is not true is

$$\leq 5n(1-p)^{(1+\epsilon)\lambda(1-p)} = 5n^{-\epsilon} \to 0 \qquad \square$$

Lemma 3.7. Let $\epsilon > 0$, $0 < p < 1$ be fixed, $\alpha = (1+\epsilon)\lambda(1-p^2) \leq \psi < n$. For almost all $G \in \Psi_n(\psi, p)$ there exists a path of length three between every pair of vertices u, v such that $|u-v| \leq 3\psi - \alpha$.

proof. Let $u, v \in V$ be such that $i = 3\psi - |u-v| \geq \alpha$. Let $x_j = u + \psi - j$ for $0 \leq j \leq i$, as illustrated in Figure 3. Clearly any 3-path connecting u and v must pass through one of $x_0, \ldots, x_i$. Applying A.8, the probability that no three path joins u and v is

$$= \mathrm{P}(\text{no 3-path} \wedge u \not- x_0 \wedge \cdots \wedge u \not- x_i)$$

$$+ \mathrm{P}(\text{no 3-path} \wedge u \not- x_0 \wedge \cdots \wedge u \not- x_{i-1} \wedge u - x_i)$$

$$+ \mathrm{P}(\text{no 3-path} \wedge u \not- x_0 \wedge \cdots \wedge u \not- x_{i-2} \wedge u - x_{i-1})$$

$$+ \cdots$$

$$+ \mathrm{P}(\text{no 3-path} \wedge u - x_0)$$

$$= (1-p)^{(i+1)} \mathrm{P}(\text{no 3-path} | u \not- x_0 \wedge \cdots \wedge u \not- x_i)$$

$$+ (1-p)^{(i+1)} \mathrm{P}(\text{no 3-path} | u \not- x_0 \wedge \cdots \wedge u \not- x_{i-1} \wedge u - x_i)$$

$$+ (1-p)^{i} \mathrm{P}(\text{no 3-path} | u \not- x_0 \wedge \cdots \wedge u \not- x_{i-2} \wedge u - x_{i-1})$$

$$+ \cdots$$

$$+ (1-p) \mathrm{P}(\text{no 3-path} | u - x_0)$$

$$< (1-p)^{(i+1)} [\, \mathrm{P}(\text{no 3-path} | u \not- x_0 \wedge \cdots \wedge u \not- x_i)$$

$$+ \mathrm{P}(\text{no 3-path} | u \not- x_0 \wedge \cdots \wedge u \not- x_{i-1} \wedge u - x_i)$$

$$+ (1-p)^{-1} \mathrm{P}(\text{no 3-path} | u \not- x_0 \wedge \cdots \wedge u \not- x_{i-2} \wedge u - x_{i-1} \wedge u \not- x_i)$$

$$+ \cdots$$

$$+ (1-p)^{-i} \mathrm{P}(\text{no 3-path} | u - x_0 \wedge u \not- x_1 \wedge \cdots \wedge u \not- x_i)]$$

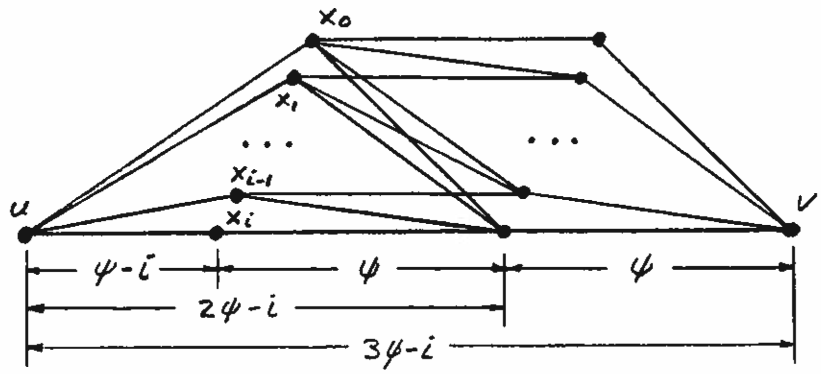

Figure 3. Definition of x_is

$$=(1-p)^{i+1}\left[1+(1-p^2)+\frac{(1-p^2)^2}{(1-p)}+\frac{(1-p^2)^3}{(1-p)^2}+\cdots+\frac{(1-p^2)^{i+1}}{(1-p)^i}\right]$$

$$=(1-p)^{i+1}\left[1+(1-p^2)\sum_{j=0}^{i}(1+p)^j\right]$$

$$=(1-p)^{i+1}\left[1+(1-p^2)\frac{(1+p)^{i+1}-1}{p}\right]$$

$$=\frac{1}{p}\left[(1-p^2)^{i+2}+(p^2+p-1)(1-p)^{i+1}\right]$$

Since for each value of i there are at most n vertex pairs u,v such that $|u-v|=i$, the probability that any pair u,v with $|u-v|\leq 3\psi-\alpha$ is not connected by a 3-path is

$$\leq \sum_{i=|\alpha|}^{3\psi-1}\frac{n}{p}\left[(1-p^2)^{i+2}+(p^2+p-1)(1-p)^{i+1}\right]$$

$$<\frac{n}{p}\left[\left\{(1-p^2)^{\alpha}\sum_{i=0}^{\infty}(1-p^2)^i\right\}+\left\{(p^2+p-1)(1-p)^{\alpha}\sum_{i=0}^{\infty}(1-p)^i\right\}\right]$$

$$<\frac{n}{p}\left[\frac{1}{p^2}n^{-(1+\epsilon)}+\frac{p^2+p-1}{p}n^{-(1+\epsilon)}\right]<\frac{2}{p^3}n^{-\epsilon}\rightarrow 0$$

□

Theorem 3.6. Let $\epsilon>0$, $0<p<1$ be fixed, $\alpha=(1+\epsilon)\lambda(1-p^2)$, $\beta=(1+\epsilon)\lambda(1-p)$ and $\max(\alpha,2\beta)\leq\psi\leq(n-\beta)/2$. For almost all $G\in\Psi_n(\psi,p)$ $|V'_i|\leq\psi+\alpha$ for $i\geq 0$.

proof. The result follows from Theorem 3.3 for $i\geq 3$ and is immediate for $i=0$. Before proving the theorem for $1\leq i\leq 2$ we first need to show that $|V_3\cap\{\psi+2,\ldots,2\psi+\beta\}|\geq\beta$. Let $A=\{\psi+2,\ldots,2\psi+1\}$. By Lemmas 3.3 and 3.7 $A\subseteq V_2\cup V_3$. Let $x=|A\cap V_3|$. Clearly if $x\geq\beta$ then we're done. Assume then that $x<\beta$ and let $B=\{2\psi+2,\ldots,2\psi+(\beta-x)+1\}$ and let

$y=|B|$. Note that since $\psi \geq 2\beta$ that $\forall u \in B \ |\{u-\psi, \ldots, u+\psi\} \cap V_2| \geq \beta$. Thus by Lemma 3.6 $B \subseteq V_3$. Since $x+y=\beta$ we have that $|V_3 \cap \{\psi+2, \ldots, 2\psi+\beta\}| \geq \beta$.

Now, by Lemma 3.4, $l_3 \geq 2\psi-\alpha+1$. This implies that $l'_2 \geq \psi-\alpha+1$ and since $r'_2 \leq 2\psi+1$, it follows that $|V'_2| \leq \psi+\alpha$ as claimed.

Finally, note that if $u \in A$ and $u \geq \psi+\beta$ then by Lemma 3.5 $u - V_3$ and hence $u \notin V'_1$. Thus $|V'_1| \leq \psi+\beta < \psi+\alpha$ as claimed. □

Theorem 3.7. Let $\epsilon>0$, $0<p<1$ be fixed, $\psi<n$, $\ln n = o(\psi)$. For almost all $G \in \Omega_n(\psi, p)$ $LEVEL'(G) < 2(1+\epsilon)\phi(G)$.

proof. If $\psi > (n-\lambda(1-p))/2$ then since $LEVEL'(G) < n$

$$LEVEL'(G) < 2\psi+\lambda(1-p) < (2+\epsilon')\psi$$

for any fixed $\epsilon'>0$. If $\psi \leq (n-\lambda(p))/2$ then we can apply Theorems 3.3 and 3.4 giving

$$LEVEL'(G) < 2\psi + \lambda(1-p^2) < (2+\epsilon')\psi$$

By Theorem 2.3, $\psi < (1+\epsilon')\phi(G)$. Selecting ϵ' so that $(2+\epsilon')(1+\epsilon') = 2(1+\epsilon)$ yields the theorem. □

Hence the modified level algorithms are probabilistic approximation algorithms that almost always produce layouts having bandwidth at most $2(1+\epsilon)\phi(G)$. Furthermore, as will be shown in the next section, certain of the modified level algorithms are capable of near optimal performance, whereas the level algorithms are not.

4. Obtaining Nearly Optimal Layouts with Modified Level Algorithms

The main conclusions of section 3 were,

- That the level algorithms are probabilistic approximation algorithms for $G \in \Omega_n(\psi,p)$.
- That because $|V_2| \approx (2-p)\psi$ none of the level algorithms can produce solutions that are close to optimal for $G \in \Omega_n(\psi,p)$.
- That the modified level algorithms are capable of better performance for $G \in \Omega_n(\psi,p)$.

As was observed in section 3 the modified level algorithms reduce to the level algorithms if ψ is much larger than $n/2$. Consequently in this range we cannot expect near optimal performance from the modified level algorithms either. In this section we study a specific modified level algorithm and show that it produces near optimal layouts when $\psi < n/4$. Experimental evidence presented below suggests that the modified level algorithms are capable of good performance for ψ just slightly less than $n/2$. Completely different techniques are needed for $\psi > n/2$.

Let $G=(V,E)$ and define for all $u,v \in V$

$$gc_u(v) = V_2(v) \cap V'_{i+2}(u) \quad \forall v \in V_i(u)$$

$$gp_u(v) = V_2(v) \cap V'_{i-2}(u) \quad \forall v \in V_i(u)$$

Also let $gc(v) = gc_1(v)$, $gp(v) = gp_1(v)$. The algorithm we will analyze is based on the observation that for $G \in \Psi_n(\psi,p)$ if $u,v \in V'_i$ and $v-u$ is not too small, then with high probability $|gc(u)| < |gc(v)| \wedge |gp(u)| > |gp(v)|$.

Modified Level Algorithm 1 (MLA1)

For each $u \in V$

Let τ be any layout that satisfies the following conditions for all $x, y \in V$.

(a) $x \in V'_i(u) \wedge y \in V'_{i+1}(u) \Rightarrow \tau(x) < \tau(y)$

(b) $1 \le i \le 2 \wedge x, y \in V'_i(u) \wedge |gc_u(x)| < |gc_u(y)| \Rightarrow \tau(x) < \tau(y)$

(c) $i \ge 3 \wedge x, y \in V'_i \wedge |gp_u(x)| > |gp_u(y)| \Rightarrow \tau(x) < \tau(y)$

Output the layout having minimum bandwidth.

Define $MLA1(G)$ as the bandwidth of the layout produced by $MLA1$ on graph G. The following results show that under appropriate conditions $MLA1(G) \le (1+\epsilon)\phi(G)$, for almost all $G \in \Omega_n(\psi, p)$.

Lemma 4.1. Let $\epsilon > 0$, $0 < p < 1$ be fixed, $\alpha = (1+\epsilon)\lambda(1-p^2)$, $2\alpha \le \psi < n$. For almost all $G \in \Psi_n(\psi, p)$, $r'_{i-1} - 2\alpha < l'_i \le r'_{i-1} + 1$ and $r'_{i-1} + \psi - \alpha \le r'_i \le r'_{i-1} + \psi + \alpha$, for all $i \ge 1$. For $i > 1$, $r'_i \le r'_{i-1} + \psi$, and for $i \ne 2$, $r'_{i-1} - \alpha < l'_i$.

proof. Figure 4 illustrates the assertions being made. For $1 \le i \le 2$ the result is implicit in the proof of Theorem 3.6. For $i \ge 3$, Lemma 3.4 gives $l'_i > r'_{i-3} + 2\psi - \alpha$. Since $r'_{i-1} \le r'_{i-3} + 2\psi$, $l'_i > r'_{i-1} - 2\alpha$. By Lemmas 3.3 and 3.7 $\{r'_{i-3}+\psi+1, \ldots, r'_{i-3}+2\psi-\alpha\} \subseteq V'_{i-1}$ and $\{r'_{i-3}+2\psi+1, \ldots, r'_{i-3}+3\psi-\alpha\} \subseteq V'_i$ and $\{r'_{i-3}+2\psi-\alpha+1, \ldots, r'_{i-3}+2\psi\} \subseteq V'_{i-1} \cup V'_i$. This implies $r'_{i-1}+1 \in V'_i$, hence $l'_i \le r'_{i-1}+1$. For $i > 2$ it is clear that $r'_i \le r'_{i-1} + \psi \le r'_{i-1} + \psi + \alpha$. $r'_i \ge r'_{i-1} + \psi - \alpha$ follows from Lemma 3.7 and $r'_{i-1} \le r'_{i-3} + 2\psi$. □

A consequence of Lemma 4.1 is that at least $\psi - 3\alpha$ of the vertices in V'_i are found in a region containing only vertices in V'_i. These regions are shown as the solid areas in Figure 5. The regions associated with V'_i and V'_{i+1} are separated by a transition region containing at most 2α vertices.

Lemma 4.2. Let $\epsilon > 0$, $0 < p < 1$ be fixed, $\alpha = (1+\epsilon)\lambda(1-p^2)$, $4\alpha \le \psi \le n/4$. For almost all $G \in \Psi_n(\psi, p)$ $1 \le i \le 2 \wedge u, v \in V'_i \wedge u - v \ge 4\alpha \Rightarrow |gc(u)| > |gc(v)|$.

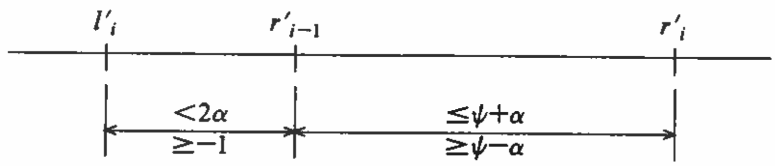

Figure 4. Illustration for Lemma 4.1

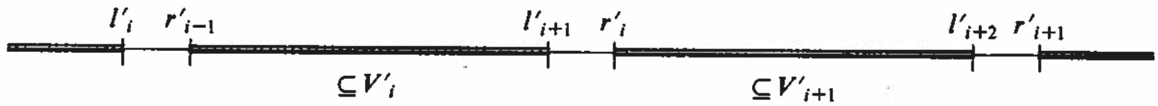

Figure 5. Separation of Vertices into Regions According to Distance from 1

proof. Lemma 3.3 implies that if $u, v \in V'_i$ and $u-v \geq \alpha$ then $gc(v) \subseteq gc(u)$. It remains only to show that there is some vertex x such that $x \in gc(u) - gc(v)$. Let $x = u+2\psi-\lceil\alpha\rceil$. If $u-v \geq 4\alpha$, Lemma 4.1 yields,

$$r'_{i+1} \leq r'_{i-1}+2\psi < l'_i+2\psi+2\alpha \leq v+2\psi+2\alpha < u+2\psi-2\alpha < x$$

Thus $x \notin V'_{i+1}$, and since by Lemma 3.3 there is a 2-path from u to x, $x \in gc(u)$. Since $x > v+2\psi$, $x \notin gc(v)$. □

Lemma 4.3. Let $\epsilon>0$, $0<p<1$ be fixed, $\alpha=(1+\epsilon)\lambda(1-p^2)$, $4\alpha\leq\psi\leq n/4$. For almost all $G \in \Psi_n(\psi,p)$ $i\geq 3 \wedge u,v \in V'_i \wedge u-v \geq 4\alpha \Longrightarrow |gp(u)| < |gp(v)|$.

proof. Lemma 3.3 implies that if $u, v \in V'_i$ and $u-v \geq \alpha$ then $gp(u) \subseteq gp(v)$. It remains to show that there exists some vertex x in $gp(v)-gp(u)$. Let $x = v-2\psi+\lceil\alpha\rceil$. If $u-v \geq 4\alpha$, Lemma 4.1 yields

$$l'_{i-1} > r'_{i-2}-2\alpha \geq r'_i-2\psi-2\alpha \geq u-2\psi-2\alpha \geq v-2\psi+2\alpha > x$$

Thus $x \notin V'_{i-1}$ and since by Lemma 3.3 there is a 2-path from v to x, $x \in gp(v)$. Since $x < u-2\psi$, $x \notin gp(u)$.

Theorem 4.1. Let $\epsilon>0$, $0<p<1$ be fixed, $\ln n = o(\psi)$, $\psi \leq n/4$. For almost all $G \in \Omega_n(\psi,p)$ $MLA1(G) < (1+\epsilon)\phi(G)$.

proof. Recall the random experiment by which G is generated. First, a graph $G' \epsilon \Psi_n(\psi, p)$ is generated, then the vertices of G' are randomly renumbered to yield G. Let τ_1 be the permutation that reverses the renumbering. That is, τ_1 is the permutation which when applied to the vertices of G converts it back to G'. Also, let τ_2 be the layout computed by *MLA*1 which satisfies $\tau^{-1}{}_1(1)=\tau^{-1}{}_2(1)$. Now consider any two vertices u, v in G. By Lemmas 4.1 to 4.3, if $\tau_1(u)-\tau_1(v) > 4\alpha$ then $\tau_2(v) < \tau_2(u)$, where $\alpha=(1+\epsilon)\lambda(1-p^2)$. Consequently, for any u there can be at most 4α vertices v such that $\tau_1(u)>\tau_1(v) \wedge \tau_2(u)<\tau_2(v)$. Similarly there can be at most 4α vertices w such that $\tau_1(u)<\tau_1(w) \wedge \tau_2(u)>\tau_2(w)$. This means that for all u, $|\tau_1(u)-\tau_2(u)| \leq 4\alpha$ and hence if u is adjacent to v then $|\tau_2(u)-\tau_2(v)| \leq \psi+8\alpha$ $\leq (1+\epsilon')\psi \leq (1+\epsilon')^2\phi(G)$ for any fixed $\epsilon'>0$. Choosing ϵ' so that $(1+\epsilon')^2=(1+\epsilon)$ yields the theorem. □

There are other possible strategies for arranging the vertices within each level. Cuthill and McKee [27] who first suggested the level algorithms, arranged the vertices in the order in which they were visited by a depth first search algorithm. This results in an arbitrary ordering of the first level and arranges each vertex in subsequent levels based on the position of its 'leftmost' neighbor in the previous level. Cheng [29][35] proposed several modifications of the Cuthill-McKee strategy. He suggested that the vertices in the first level be ordered in increasing order of the number of neighbors they have in the next level. He suggested a similar strategy for breaking ties between vertices in subsequent levels. The Cuthill-McKee algorithm with Cheng's refinements is easily adapted to the modified level strategy. I will refer to this algorithm as *MLA*2.

*MLA*2 is more difficult to analyze than *MLA*1 because decisions made in ordering each level affect the ordering of subsequent levels. Consequently one might expect that errors made in ordering the early levels could accumulate and cause large errors further on. Experimental evidence discussed below suggests that in fact the errors do not accumulate, that the process is self-limiting. However straightforward analytical techniques for bounding the error yield

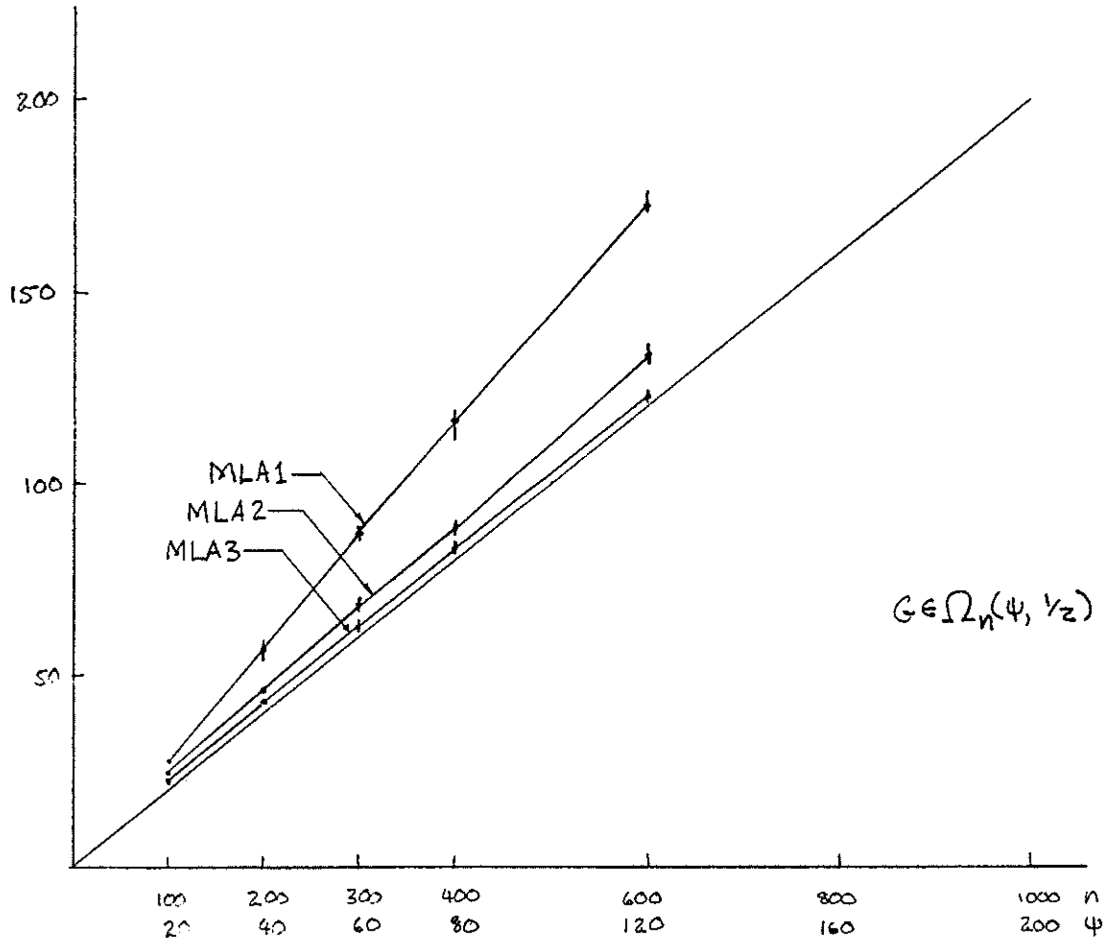

Figure 6. Monte Carlo Simulations with $\frac{n}{\psi}=5$

unsatisfactory results. Consequently, the only available results on these methods are experimental.

A series of Monte Carlo simulations was performed comparing *MLA*1, *MLA*2 and a third algorithm (denoted *MLA*3) which uses the technique of *MLA*1 to arrange the vertices in the first level and reverts to the Cuthill-McKee technique on subsequent levels. The results are summarized in Figures 6-7. For each data point shown, 25 random graphs were generated and each of the algorithms was run. The vertical lines show the range of the bandwidth of the layouts produced by each algorithm. The points are the average bandwidths. There are several points worth noticing. In general *MLA*3 has the best performance followed by *MLA*2 and

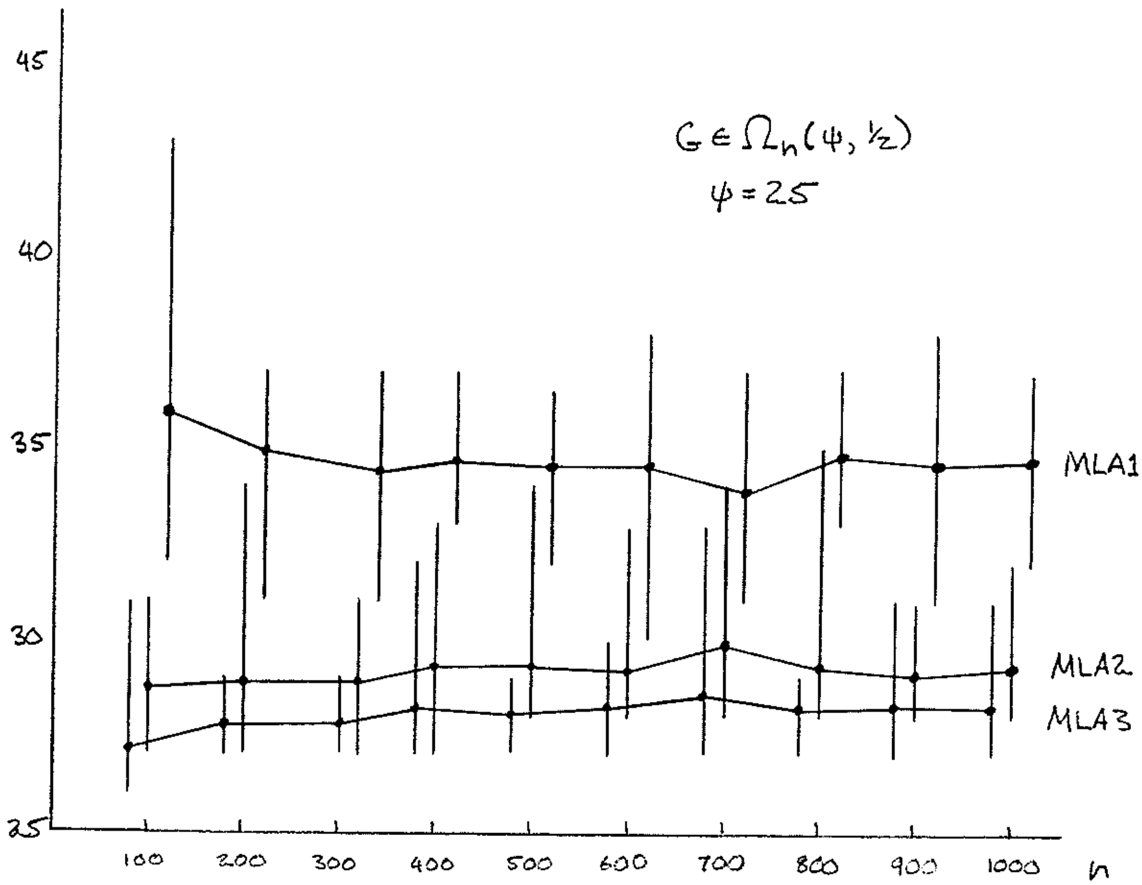

Figure 7. Monte Carlo Simulations with ψ=25

MLA 1. One exception to this appears in Figure 7 where *MLA* 3 begins to perform poorly when ψ exceeds $n/3$. A similar degradation is observed in *MLA* 2 when ψ exceeds $n/2$. Both effects are predicted by the analysis. *MLA* 3 performs poorly for $\psi > n/3$ since at this point most vertices u in V'_1 are such that $gc(u) = V'_3$. The degradation of *MLA* 2 is due to the fact that when $\psi > n/2$, $V'_3 = \varnothing$ and so the algorithm reverts to a level algorithm.

The running time of the three algorithms is roughly comparable. The original Cuthill-McKee strategy trying all choices for the initial vertex results in a time worst-case time complexity of $O(n|E|) = O(n^2\phi(G))$. Including Cheng's refinements increases the running time to $O(n^3)$ in the worst case because of the need to compute for each vertex the intersection of its neighborhood with the next level. For graphs in $\Omega_n(\psi,p)$ however one expects the running

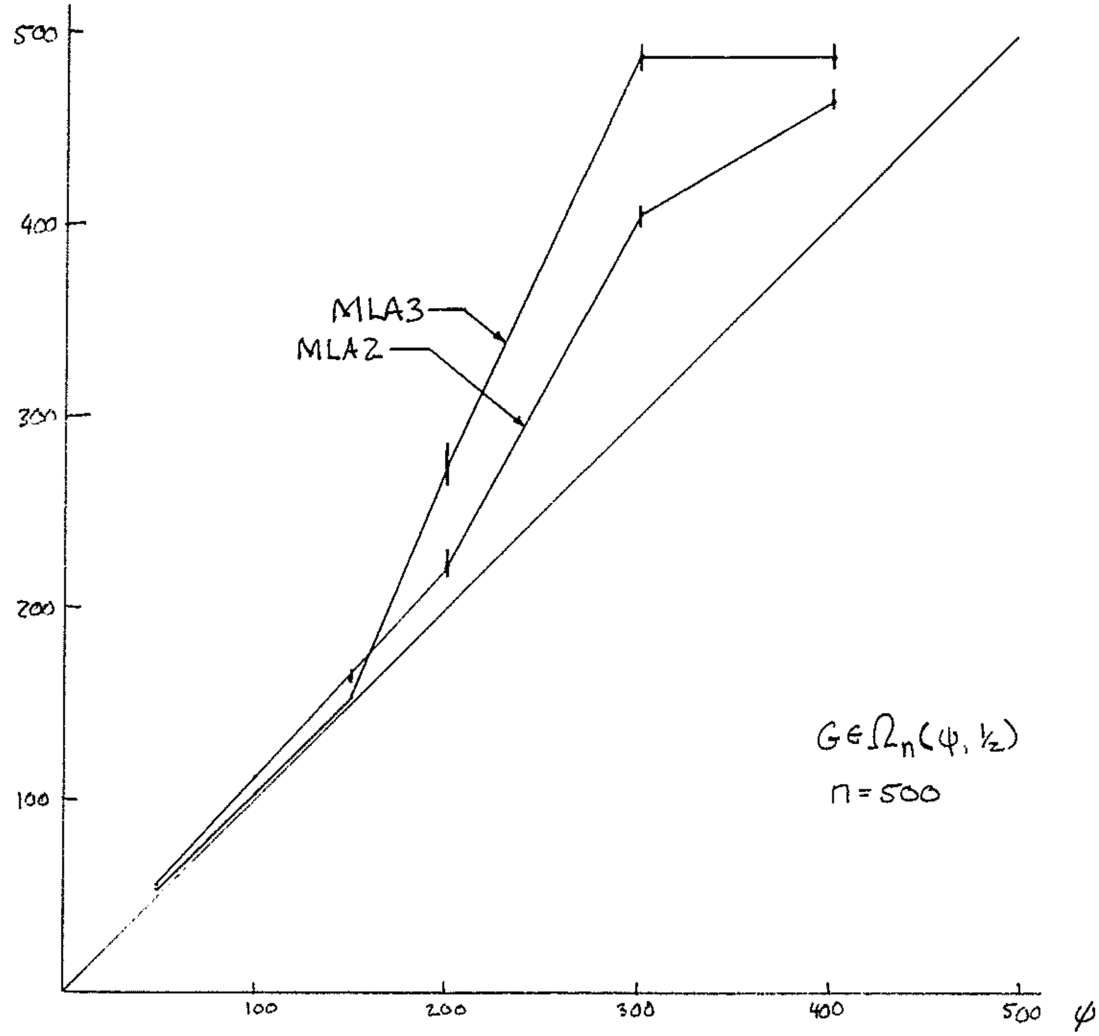

Figure 8. Monte Carlo Simulations with n=500

time to be closer to $O(n^2\phi(G))$. $MLA3$ also has time complexity $O(n^3)$ and expected running time $O(n^2\phi(G))$. $MLA1$ is somewhat worse, having worst-case complexity of $O(n^4)$ and expected running time of $O(n^2\phi^2(G))$.

Up until this point it has been assumed that the bandwidth minimization algorithms tried all possible choices for the initial vertex u that defines the levels. However it is clear that for $G \in \Psi_n(\psi,p)$ the degree of vertex 1 is likely to be smaller than that of most other vertices. Consequently it suffices to try only low degree vertices to define the levels. Cuthill and McKee were the first to notice that the initial vertex should be one of small degree, using heuristic arguments. The following lemma puts a probabilistic upper bound on the number of low degree vertices that need be tried to obtain near optimal performance.

Define $ld(G)=\{v \in V \mid d(v) \le d(1)\}$.

Lemma 4.4. Let $\epsilon>0$, $0<p<1$ be fixed, $12(1+\epsilon)(1/p)\ln n \le \psi < n$. For almost all $G \in \Psi_n(\psi,p)$, $|ld(G)| < 4\left[(3/p)(1+\epsilon)\psi \ln n\right]^{1/2}$.

proof. Let $0<\alpha\le\psi p/2$. By Lemma 3.5

$$P(d(1)\ge\psi p+\alpha)<e^{-\alpha^2/3\psi p}$$

For $v \in V$ such that $(2\alpha/p)<v<n-(2\alpha/p)$

$$P(d(v)\le\psi p+\alpha) < e^{-\alpha^2/2\psi p}$$

Letting $\alpha=\left[3(1+\epsilon)\psi p \ln n\right]^{1/2}$ yields

$$P(d(v)<d(1)) < 2e^{-\alpha^2/3\psi p} = 2n^{-(1+\epsilon)}$$

Since there are $<n$ such vertices v,

$$P(\exists v:(2\alpha/p)<v<n-(2\alpha/p)\wedge d(v)<d(1)) < 2n^{-\epsilon} \to 0$$

Consequently there are at most $4\alpha/p$ vertices in $ld(G)$. □

Lemma 4.4 implies that vertex 1 is almost always one of the first $4\alpha/p$ vertices of smallest degree. To make use of the result, a bandwidth reduction algorithm requires an estimate of ψ, but the bandwidth of any layout can be used for this purpose. The potential time savings by this method is a factor of $(np/4\alpha) \propto \frac{n}{\sqrt{(\psi/p)\ln n}}$.

5. Properties of Random Graphs

Connectivity of Random Graphs in $\Psi_n(\psi,p)$

The following theorem is a special case of a result proved by Erdös and Renyi in [16].

Theorem 5.1. Let $-1<\epsilon<1$ be fixed, $p=(1+\epsilon)\frac{\ln n}{n}$, $G\in\Gamma_n(p)$. If $\epsilon>0$, G is almost always connected. If $\epsilon<0$, G is almost always disconnected.

The following is a similar result for random graphs with small bandwidth. In particular we prove

Theorem 5.2. Let $-1<\epsilon<1$ be fixed, $0<p<1$, $\psi=½(1+\epsilon)\lambda(1-p)$, $\psi\to\infty$. If $\epsilon>0$ then almost all $G\in\Psi_n(2\psi,p)$ are connected. If $\epsilon<0$ then almost all $G\in\Psi_n(\psi,p)$ are disconnected.

To prove Theorem 5.2 we need to introduce another probability distribution and prove two lemmas. Let n and ψ be positive integers, $\psi<n$, $0<p<1$, and let $G=(V,E)$ be a random variable defined by the following experiment.

- Let $V=\{1,2,\ldots,n\}$. For each $\{u<v\}$ $1\le u,v\le n$, $|u-v|\le\psi$ *or* $|u-v|\ge n-\psi$ include the edge $\{u,v\}$ in E with probability p.

The probability distribution defined by this experiment is denoted $\Psi_n^c(\psi,p)$.

Lemma 5.1. Let $-1<\epsilon<1$ be fixed, $0<p<1$, $\psi=½(1+\epsilon)\lambda(1-p)$, $1\le\psi\le\frac{n}{2}$, $G\in\Psi_n^c(\psi,p)$. If $\epsilon>0$ then G almost always contains no isolated vertex. If $\epsilon<0$ then G almost always contains at least one isolated vertex.

Lemma 5.2. Let $0<\epsilon\le p<1$ where ϵ is fixed, and let $G\in\Gamma_n(p)$. Then $P(D(G)>2)\le\binom{n}{2}(1-\epsilon^2)^{n-2}$.

Proof of Theorem 5.2. First part - $\epsilon>0$. G is connected if the first 2ψ vertices induce a connected subgraph and all other vertices have at least one edge to a lower numbered vertex. By Lemma 5.2, if $p>\alpha$ for some $\alpha>0$ then the probability that the first 2ψ vertices induce a subgraph of diameter >2 is $\leq \binom{2\psi}{2}(1-\alpha^2)^{2\psi-2} \rightarrow 0$. Hence, if p is bounded below, the first 2ψ vertices almost always induce a connected subgraph. If on the other hand $p \rightarrow 0$ we must use Theorem 5.1 to establish that the first 2ψ vertices induce a connected subgraph. This requires that we show that there exists some $\gamma > 0$ such that $p \geq (1+\gamma)\frac{\ln(2\psi)}{2\psi}$. From the hypothesis of the theorem

$$\frac{p(2\psi)}{\ln(2\psi)} = (1+\epsilon)\frac{p \ln n}{\ln(1/(1-p)) \ln(2\psi)} > (1+\frac{\epsilon}{2})$$

for large enough n since $p \rightarrow \ln(1/(1-p))$ as $p \rightarrow 0$ and $n > 2\psi$. Now, the probability that any of the remaining vertices have no edges to lower numbered vertices is $< n(1-p)^{2\psi} = n^{-\epsilon} \rightarrow 0$. This completes the proof of the first part of Theorem 5.1.

Now let $\epsilon < 0$ and let $G' \epsilon \Psi_n^c(\psi,p)$. Clearly, $P(G$ is connected$) \leq P(G'$ is connected) and since by Lemma 5.1, G' is almost always disconnected, it follows that G is almost always disconnected. □

Proof of Lemma 5.1. First part - $\epsilon > 0$. Let

$$X_v = \begin{cases} 1 & \text{if } v \text{ is isolated} \\ 0 & \text{if } v \text{ is not isolated} \end{cases}$$

$$X = X_1 + X_2 + \cdots + X_n$$

Then by A.13 and A.15,

$$\mu = E(X) = \sum_{v=1}^{n} E(X_v) = n(1-p)^{2\psi}$$

$$P(X \geq 1) \leq \mu = n(1-p)^{2\psi} = n^{-\epsilon} \rightarrow 0$$

This completes the proof of the first part.

Second part - $\epsilon < 0$. Let $X, X_1, \ldots, X_n$ be defined as before. By A.14,

$$E(X^2)=\sum_{u=1}^{n}\sum_{v=1}^{n}E(X_uX_v)$$

$$=\sum_{u=1}^{n}\sum_{v=1}^{n}P(u \text{ and } v \text{ are both isolated})$$

$$=n(1-p)^{2\psi}+2\psi n(1-p)^{4\psi-1}+n(n-2\psi-1)(1-p)^{4\psi}$$

Using A.17,

$$P(X=0)\le\frac{\sigma^2}{\mu^2}=\frac{E(X^2)-\mu^2}{\mu^2}$$

$$=\frac{1}{\mu}+\frac{2\psi p}{n(1-p)}-\frac{1}{n}$$

$$<n^{\epsilon}+(1+\epsilon)\frac{p\ln n}{n(1-p)\ln(1/(1-p))}$$

The function $\frac{p}{(1-p)\ln(1/(1-p))}$ gets large as $p\to 1$. However, $1\le\psi<(1+\epsilon)\lambda(1-p)\Longrightarrow p\le 1-n^{-(1+\epsilon)}$. Hence,

$$\mathrm{P}(X=0)<n^{\epsilon}+(1+\epsilon)\frac{\ln n}{n^{-\epsilon}\ln n^{1+\epsilon}}=2n^{\epsilon}\to 0 \qquad \square$$

Proof of Lemma 5.2. Let u and v be any two vertices in G. The number of possible 2-paths between them is $n-2$ and the probability that any one of them is absent is $1-p^2$. Hence the probability that u and v are not connected by a 2-path is $(1-p^2)^{n-2}$. Consequently, the probability that any pair of vertices is not connected by a 2-path is $\le\binom{n}{2}(1-p^2)^{n-2}\le\binom{n}{2}(1-\epsilon^2)^{n-2}$. $\square$

Diameter of Random Graphs in $\Psi_n(\psi,p)$

A simple lower bound for the bandwidth of any connected graph is given by

$$\phi(G)\ge\omega(G)=\left\lceil\frac{n-1}{D(G)}\right\rceil \tag{5.1}$$

since the first and last vertices in any optimal layout are connected by a path of length at most $D(G)$ and hence at least one edge in this path has length $\ge\omega(G)$. Chvatal [36] was apparently the first to notice this. A more general lower bound is given by

$$\phi(G) \geq \omega^*(G) = \max_{G'} \omega(G')$$

where G' ranges over all connected subgraphs of G. The graph shown in Figure 9 shows that $\omega^*(G) \neq \phi(G)$ in general. It is natural to ask if there is any constant c such that for all connected graphs $\phi(G) \leq c\,\omega^*(G)$. We can in fact show that this is not the case. To do so we need some additional definitions.

Let τ be any layout of $G=(V,E)$. The *cutwidth* of G with respect to τ is denoted by $\theta_\tau(G)$ and is defined by

$$\theta_\tau(G) = \max_{1 \leq i < n} |\{\{u,v\} \in E \mid \tau(u) \leq i < \tau(v)\}|$$

$\theta_\tau(G)$ counts the maximum number of edges crossing any vertical line drawn through a picture of the layout of G. For example in the layout shown in Figure 10, $\theta_\tau(G)=4$. The cutwidth of G is defined as $\theta(G) = \min_{\tau} \theta_\tau(G)$. The cutwidth minimization problem is NP-complete for general graphs [37][38]. Lengauer [39][40] has studied the problem for trees and has given a worst case approximation algorithm for this case. Recently Chunk, Makedon, Sudborough and Turner [41] have found a polynomial time algorithm for trees with limited vertex degree. Our interest in the problem here is in the relationship of cutwidth to bandwidth.

Lemma 5.3. If $T=(V,E)$ is a tree then $\phi(T) > \frac{1}{2}\theta(T)$.

proof. Let τ be any minimum bandwidth layout for T, let $k=\phi(T)$ and let i be any integer such that $1 \leq i < n$. Let H be the subgraph induced by $\{\tau^{-1}(max(1,i-k+1)), \ldots, \tau^{-1}(min(n,i+k))\}$. Since $\phi(T)=k$, $\{\{u,v\} \in E \mid \tau(u) \leq i \leq \tau(v)\}$ is a subset of the edges of H, and since H is a forest with $2k$ vertices it has at most $2k-1$ edges. Since this holds for any i in the range, $\theta(T) \leq \theta_\tau(T) \leq 2k-1 < 2\phi(T)$. □

Let $G=(V,E)$. G' is defined to be a *subdivision* of G if we can obtain G' from G by replacing edges with vertices of degree two as shown in Figure 11. Notice that the process of subdividing a single edge does not affect the cutwidth of a graph. Hence if G' is a subdivision of G, $\theta(G') = \theta(G)$.

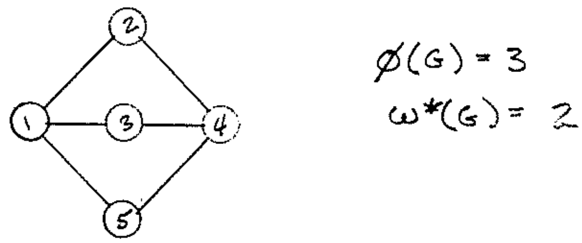

Figure 9. Graph Showing $\omega^*(G) \neq \phi(G)$

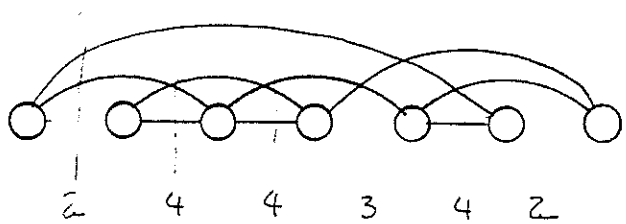

Figure 10. Example of Cutwidth Definitions

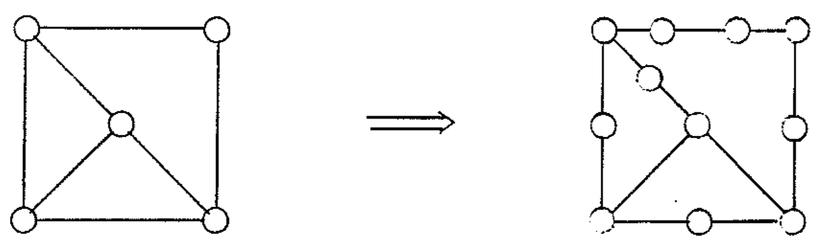

Figure 11. Subdivision of a Graph

Let B_h be a complete binary tree with h levels (2^h-1 vertices). Lengauer [40] showed that $\theta(B_h) = \left\lceil \frac{h-1}{2} \right\rceil + 1$. We are now ready to show that there is no constant c such that $\phi(G) \leq c\,\omega^*(G)$ holds for all graphs. We give a tree T_h with $\omega^*(G)=2$, but which is a subdivision of B_h and consequently by the preceding arguments has $\phi(T_h) > \frac{1}{2}\theta(T_h) > h/4$.

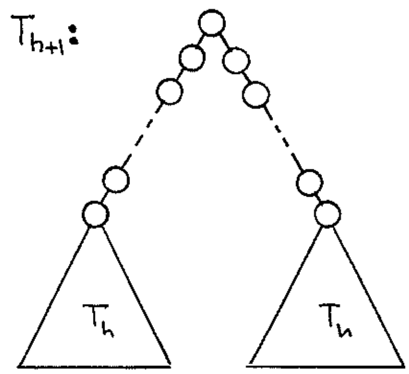

Figure 12. Definition of T_h

T_h is described by a recursive construction. $T_2=B_2$ and T_{h+1} is constructed by joining two copies of T_h with two long chains as shown in Figure 12. The number of vertices in the long chains is twice the number of vertices in T_h. This ensures that $\omega^*(T_h)=2$. Hence we may conclude that even for the restricted case of degree three trees there is no constant c such that $\phi(G)\leq c\,\omega^*(G)$. I am indebted to Ronald Graham for pointing out this elegant argument. In spite of this result however, we can show that for almost all $G\in\Psi_n(\psi,p)$, $D(G)<(1+\epsilon)\frac{n}{\phi(G)}+3$.

Let $G=(V,E)\in\Psi_n(\psi,p)$ and let $Q=(v_0, \ldots, v_r)$ be the unique path in G that satisfies

$$
\begin{aligned}
&v_0=1\\
&v_i=\max\,\{u\in v\,|\,v—v_{i-1}\}\quad 1\leq i\leq r\\
&v_i<n-\psi\quad 1\leq i<r\\
&v_r\geq n-\psi
\end{aligned}
$$

Next, let $x_i=v_i-v_{i-1}$ $(1\leq i\leq r)$ and let $x=\sum_{i=1}^{r}x_i$. In what follows we will derive a probabilistic upper bound on x which will be used to obtain an upper bound on r. Its is clear that for $1\leq i\leq r$

$$P(x_i=0)=p$$

$$P(x_i=1)=(1-p)p$$

$$\cdots$$

$$P(x_i=k)=(1-p)^k p \qquad 1\le k\le\psi$$

For k larger than ψ, $P(x_i=k)$ depends on whether or not $v_i\le\psi$. For simplicity we will assume that $P(x_i=k)=(1-p)^k p$ for all $k\ge 0$. The error committed by this approximation vanishes as $\psi\to\infty$. We are now ready to prove

Lemma 5.4. For any $\epsilon>0$, $P(x>(1+\epsilon)r(1-p)/p)\le\dfrac{1}{\epsilon^2 r(1-p)}$.

proof. Using the approximation discussed above,

$$E(x_i)=\sum_{j=1}^{\infty} j(1-p)^j p = p\frac{(1-p)}{p^2}=\frac{1-p}{p}$$

$$E(x_i^2)=\sum_{j=1}^{\infty} j^2(1-p)^j p$$

$$=p\sum_{j=1}^{\infty}(1-p)^j(1+3+\cdots+2j-1)$$

$$=p\left[\frac{(1-p)}{p}+\frac{3(1-p)^2}{p}+\frac{5(1-p)^3}{p}+\cdots\right]$$

$$=2\left[(1-p)+2(1-p)^2+3(1-p)^3+\cdots\right]$$

$$-\left[(1-p)+(1-p)^2+(1-p)^3+\cdots\right]$$

$$=2\frac{1-p}{p^2}-\frac{1-p}{p}$$

Let μ be the mean and σ^2 the variance of x.

$$\mu=rE(x_i)=r\frac{1-p}{p}$$

$$\sigma^2=r(E(x_i^2)-E^2(x_i))=r\frac{1-p}{p^2}$$

By Chebyshev's inequality

$$P(x > (1+\epsilon)\mu) \le \frac{\sigma^2}{\epsilon^2\mu^2} = \frac{r(1-p)/p^2}{\epsilon^2 r^2 (1-p)^2/p^2} = \frac{1}{\epsilon^2 r(1-p)} \qquad \square$$

Lemma 5.5. For any $\epsilon > 0$ $P\left(r \ge \frac{n}{\psi - (1+\epsilon)(1-p)/p}\right) \le \frac{1}{\epsilon^2 r(1-p)}$.

proof. From the definition, $v_r - 1 = r\psi - x$, hence

$$n-1 \ge v_r - 1 = r(\psi - \frac{x}{r})$$
$$r < \frac{n}{\psi - x/r}$$

The result now follow from Lemma 5.5. □

Theorem 5.3. Let $\epsilon > 0$, $0 < p < 1$ be fixed, $(1+\epsilon)\lambda(1-p^2) \le \psi < n$. For almost all $G \in \Psi_n(\psi, p)$ $D(G) < (1+\epsilon)\frac{n}{\psi} + 3$.

proof. By Lemma 5.5

$$P\left(r \ge \frac{n}{\psi - (1+\alpha)(1-p)/p}\right) \le \frac{1}{\alpha^2 r(1-p)}$$

holds for any $\alpha > 0$. Letting $\alpha^2 = \ln n$, $\frac{1}{\alpha^2 r(1-p)} \to 0$ and $\frac{n}{\psi - (1+\alpha)(1-p)/p} \to \frac{n}{\psi}$ since $\psi \ge \lambda(1-p^2) \propto \ln n$. Thus for any fixed $\epsilon > 0$, $r < (1+\epsilon)\frac{n}{\psi}$ for large enough n. By Lemma 3.4 there is a 2-path from each $u \in V$ to some v_i $1 \le i \le r$. This implies $D(G) \le r+3$. □

Diameter of Random Graphs in $\Gamma_n(p)$

By Lemma 5.2, if $p \ge \epsilon > 0$ then for almost all $G \in \Gamma_n(p)$, $D(G) = 2$. When p is allowed to approach zero as n gets large the diameter can become larger. By Theorem 5.1, when p is much less than $\frac{\ln n}{n}$ the graph is likely to be disconnected. In this section we consider the probable diameter of random graphs in $\Gamma_n(p)$ when $p = c\frac{\ln n}{n}$ and c is a constant. We do this by examining the probable size of $V_1, V_2, \ldots$. Let $n_i = |V_i|$. Clearly,

$$n_0 = 1$$

$$n_1 \in \mathbf{B}(n-1,p)$$

$$n_2 \in \mathbf{B}(n-(n_1+1), 1-(1-p)^{n_1})$$

$$\cdots$$

$$n_{k+1} \in \mathbf{B}(n-s_k, 1-(1-p)^{n_k})$$

where $s_k = \sum_{j=0}^{k} n_j$. Define $\hat{n}_0=1$, $\hat{n}_{k+1}=(n-\hat{s}_k)(1-(1-p)^{\hat{n}_k})$, where $\hat{s}_k=\sum_{j=0}^{k}\hat{n}_j$. We can use $\hat{n}_k$ as an estimator for n_k. Figure 13 gives values of $\hat{n}_k$ for particular values of n and p. The sequence grows very rapidly until a large fraction of the vertices in the graph has been 'captured'. Then the remaining vertices are taken in the last step. The figure also gives values of the function $(np)^k$. For $k \le 3$, $(np)^k$ gives an excellent estimate for $\hat{n}_k$.

Let k^* be such that $s_{k^*} = n$. In the following we show that for $k \le k^*-2$, $n_k > (np/8)^k$ with high probability. We can use this to get a probabilistic upper bound on k^* and hence on $D(G)$. The main result is

Theorem 5.4. Let $c > 8$ be fixed, $p = c\frac{\ln n}{n}$, $\gamma = np/8$. For almost all $G \in \Gamma_n(p)$, $1 \le k \le k^*-2 \Rightarrow n_k > \gamma^k$.

The proof of Theorem 5.4 is contained in the following lemmas.

Lemma 5.6. Let $c > 8$ be fixed, $p = c\frac{\ln n}{n}$, $\gamma = np/8$. For almost all $G \in \Gamma_n(p)$ $1 \le k \le k^*-2 \wedge n_{k-1} < 1/p \wedge s_{k-1} \le n/2 \Rightarrow n_k > \gamma n_{k-1}$.

proof. Since $n_k \in \mathbf{B}(n-s_{k-1}, 1-(1-p)^{n_{k-1}})$,

$$\bar{n}_k = E(n_k) = (n-s_{k-1})(1-(1-p)^{n_{k-1}}) \ge \frac{n}{2} p n_{k-1}(1-\tfrac{1}{2} p n_{k-1}) > \frac{np}{4} n_{k-1} = 2\gamma n_{k-1}$$

By Lemma 3.5

$$\mathrm{P}(n_k \le \gamma n_{k-1}) \le \mathrm{P}(n_k \le \tfrac{1}{2}\bar{n}_k) < e^{-\bar{n}_k/8} < e^{-\gamma n_{k-1}}$$

Let A_k denote the event $n_k \le \gamma n_{k-1}$. By A.6 the probability that there exists a k satisfying the

k	$\hat{n}_k$	$(np)^k$
0	1	1
1	27.6	27.6
2	763	763
3	20,850	21,096
4	428,450	582,890
5	549,910	-

Figure 13. Comparison of $\hat{n}_k$ with $(np)^k$ for $n=10^6$, $p=2\frac{\ln n}{n}$

hypothesis of the lemma, such that A_k holds is

$$\leq \mathrm{P}(A_1)+\mathrm{P}(A_2|\bar{A}_1)+\cdots+\mathrm{P}(A_{k^*-2}|\bar{A}_1\cdots\bar{A}_{k^*-3})$$

$$\leq e^{-\gamma/4}+e^{-\gamma^2/4}+\cdots+e^{-\gamma^{k^*-2}/4}\rightarrow 0 \qquad \square$$

Lemma 5.7. Let $c>8$ be fixed, $p=c\frac{\ln n}{n}$. For almost all $G\in\Gamma_n(p)$ $1\leq k\leq k^*-2\wedge s_{k-1}\leq n/2$ $\Longrightarrow n_{k-1}<1/p$.

proof. Assume that $n_{k-1}\geq 1/p$. Then since $n_k\in \mathrm{B}(n-s_{k-1},1-(1-p)^{n_{k-1}})$,

$$\bar{n}_k=E(n_k)=(n-s_{k-1})(1-(1-p)^{n_{k-1}})\geq\frac{n}{2}(1-1/e)>n/4$$

By Lemma 3.5

$$\mathrm{P}(n_k\leq n/8)\leq \mathrm{P}(n_k\leq \tfrac{1}{2}\bar{n}_k)<e^{-\bar{n}_k/8}<e^{-n/32}\rightarrow 0$$

Hence, assume $n_k>n/8$. Then the probability that any of the remaining vertices is not adjacent to something in V_k is

$$\leq (n-s_k)(1-p)^{n_k}<ne^{-np/8}=n^{1-c/8}\rightarrow 0$$

This implies that $k^*\leq k+1$ which is a contradiction. $\square$

Lemma 5.8. Let $c>8$ be fixed, $p=c\frac{\ln n}{n}$. For almost all $G\in\Gamma_n(p)$, $1\leq k\leq k^*-2\Longrightarrow s_{k-1}\leq n/2$.

proof. Assume that $s_{k-1}>n/2$ and let k' be the smallest integer such that $s_{k'}>n/2$. By Lemma

n	10^2	10^3	10^4	10^5	10^6	10^{10}	10^{20}
$D(G)$	≤ 6	≤ 8	≤ 8	≤ 10	≤ 12	≤ 16	≤ 24

Figure 14. Growth of Diameter

5.7, $n_{k'-1} < 1/p$ and by Lemma 5.6, for all $k \leq k'$, $n_k > \gamma n_{k-1}$, where $\gamma = np/8$. Since for large n, $\gamma > 2$, we have $s_k > 2s_{k-1}$ for $k \leq k'$. Consequently $n_{k'} = s_{k'} - s_{k'-1} > s_{k'}/2 > n/4$. Now, the probability that any of the vertices in $V - (V_0 \cup V_1 \cup \cdots \cup V_{k'})$ is not adjacent to some vertex in $V_{k'}$ is

$$< (n - s_{k'})(1-p)^{n_{k'}} < ne^{-np/4} = n^{1-c/4} \rightarrow 0$$

This implies that $k^* \leq k'+1$ which is a contradiction. □

This establishes Theorem 5.4. We can now use this to bound $D(G)$.

Theorem 5.5. Let $c>8$ be fixed, $p = c\frac{\ln n}{n}$, $\gamma = np/8$. For almost all $G \in \Gamma_n(p)$, $D(G) \leq 2\left(\left\lceil \frac{\ln(1/p)}{\ln\gamma} \right\rceil + 2\right)$.

proof. Note that $D(G) \leq 2k^*$. Let k' be the smallest integer such that $\gamma^{k'} \geq 1/p$. Clearly $k' = \left\lceil \frac{\ln(1/p)}{\ln\gamma} \right\rceil$. If $k' > k^*-2$, we're done. If $k' \leq k^*-2$ we can apply Theorem 5.4 giving $n_{k'} \geq 1/p$. By the argument used in the proof of Lemma 5.7, this implies $k^* \leq k'+2$. □

Expanding the bound on $D(G)$ gives $D(G) \leq 2\left(\left\lceil \frac{\ln(n/c\ln n)}{\ln\ln n^{c/8}} \right\rceil + 2\right)$. Figure 14 shows how this expression grows as n gets large for $c=10$.

6. Conclusions

In this thesis I have studied the probabilistic analysis of algorithms for *NP*-hard optimization problems, focussing on the bandwidth minimization problem for graphs. I have observed that the non-uniformity of the usual probability distributions for graphs makes them inappropriate for studying the performance of bandwidth minimization algorithms and have introduced the idea of universal probabilistic algorithms to guide the selection of more appropriate distributions.

I have showed that all algorithms in the class of modified level algorithms are probabilistic approximation algorithms for $G \in \Omega_n(\psi,p)$ when p is fixed and $\ln n = o(\psi)$. More precisely, I have shown that these algorithms almost always produce layouts that have bandwidth between $(1-\epsilon)(2-p)\phi(G)$ and $(1+\epsilon)(3-p)\phi(G)$.

I have introduced the class of modified level algorithms and shown that they are probabilistic approximation algorithms for $G \in \Omega_n(\psi,p)$ when p is fixed and $\ln n = o(\psi)$. Specifically I have shown that they almost always produce layouts having bandwidth $\leq (1+\epsilon)2\phi(G)$. In addition I have shown that a specific modified level algorithm is a probabilistic optimization algorithm for $G \in \Omega_n(\psi,p)$ when $\psi \leq n/4$ and $\ln n = o(\psi)$.

I have studied properties of random graphs with limited bandwidth, showing when such graphs are likely to be connected and giving a probabilistic upper bound on their diameter. I've also given a probabilistic upper bound on the diameter of graphs with unrestricted bandwidth.

There are several areas that merit further study. The whole issue of what probability distributions are appropriate for comparing the probable performance of algorithms is extremely important. The concept of universal probabilistic algorithms appears to be a useful one, at least in guiding the search for appropriate distributions. The idea would have greater utility if unbiased distributions such as $\Phi_n(\psi)$ could be related to analytically tractable distributions. If one could show for example that events that occur with high probability in $\Omega_n(\psi,½)$ also occur with high probability in $\Phi_n(\psi)$ it might be possible to establish the existence of a universal probabilistic optimization algorithm for bandwidth minimization.

Probability distributions such as $\Gamma_n(p)$ and $\Psi_n(\psi,p)$ that generate random subgraphs of certain base graphs have very nice analytical properties. There are similar distributions that could also prove useful. For example, one can define distributions that generate random subgraphs of complete k-partite graphs that could be useful for studying properties of random k-chromatic graphs.

Another area worth pursuing is to generalize the main results of this thesis by removing the limitations on ψ and p. The lower bound restrictions on ψ and the restriction of p to constant values appear to be artifacts of the proof techniques. In certain cases I have knowingly sacrificed generality for the sake of clarity. It would be interesting to see how loose the restrictions can be made.

Analytical results on the Cuthill-McKee method of ordering the vertices within levels would be very worthwhile. Another interesting problem is the development of probabilistic optimization algorithms for $\Omega_n(\psi,p)$ when $\psi > n/2$. This appears to require a very different approach.

Finally, there are many other properties of random graphs in $\Psi_n(\psi,p)$ that bear investigating. These include the presence of hamiltonian circuits, the degree sequence, clique number and chromatic number.

Appendix - Summary of Results from Probability Theory*

It is convenient to think of probability theory as the study of random experiments. For example, a random experiment might consist of tossing a coin 100 times and observing the sequence of heads and tails. A particular outcome of a random experiment is called a *sample point* and the set of all sample points is called the *sample space*. In the example, each sequence of heads and tails is a sample point and the set of 2^{100} possible sequences is the sample space. An *event* is a set of sample points. For example, we could define the set of all sequences having heads on the twenty-seventh toss as being an event. A sample space can be either discrete or continuous. We will be concerned only with discrete sample spaces.

Given a discrete sample space with sample points $E_1, E_2, \ldots$, define $P(E_i)$ to be the probability that the event E_i occurs in a particular random experiment. Clearly,

$$P(E_1)+P(E_2)+\cdots=1 \tag{A.1}$$

If the event $A=\{E_{i_1}, \ldots, E_{i_k}\}$ then

$$P(A)=\sum_{j=1}^{k} P(E_{i_j}) \tag{A.2}$$

If A and B are both events then

$$P(A \vee B)=P(A)+P(B)-P(A \wedge B) \tag{A.3}$$

If $A \cap B=\varnothing$ then

$$P(A \vee B)=P(A)+P(B) \tag{A.4}$$

In general,

$$P(A \vee B) \leq P(A)+P(B) \tag{A.5}$$

This fact is used frequently throughout the thesis.

The probability that a sample point is in A given that it is in B is denoted $P(A|B)$ and is

* The material in this section can be found in any standard textbook on probability theory, for example [42].

defined by

$$P(A|B)=\frac{P(A \wedge B)}{P(B)} \tag{A.6}$$

A and B are said to be independent if $P(A|B)=P(A)$. If A and B are independent we have as an immediate consequence that

$$P(A \wedge B)=P(A)P(B) \tag{A.7}$$

If $H_1, \ldots, H_k$ are events that are mutually exclusive and exhaustive (that is, every sample point is in exactly one of the H_i), then

$$P(A)=\sum_{i=1}^{k} P(A \wedge H_i)=\sum_{i=1}^{k} P(A|H_i)P(H_i) \tag{A.8}$$

A *random variable* is a function that maps the points in the sample space to the integers or some subset of the integers. Referring to the earlier example, the number of heads obtained in the sequence of 100 tosses is a random variable.

A *probability distribution* is a function that assigns probabilities to the values of a random variable. It x is a random variable defined on the sample space S, then we can define the probability distribution f that describes x as follows.

$$f(z)=\sum_{p \in S \| x(p)=z} P(p) \tag{A.9}$$

Less formally, $f(z)$ is simply the sum of the probabilities of all the sample points which x maps to z. It is common to abbreviate the summation on the right of (A.9) by $P(x=z)$. If x is the number of heads obtained in a sequence of 100 coin tosses then $f(z)=\binom{100}{z}2^{-100}$ since there are exactly $\binom{100}{z}$ sequences containing z heads and each sequence occurs with probability 2^{-100}. This is a special case of the binomial distribution $\mathbf{B}(n,p)$ which governs the number of 'successes' in a sequence of n independent binary trials, where the probability of a success on a particular trial is equal to p. The notation $x \in f$ means that the values of the random variable x are distributed according to f.

The *expected value* of a random variable $x \in f$ is denoted $E(x)$ and is defined by

$$E(x)=\sum_{z=-\infty}^{+\infty} zf(z) \tag{A.10}$$

The expected value is also referred to as the *expectation*, *average* or the *mean*, and is often denoted by the symbol μ. The *variance* of a random variable x is a measure of the probability that x differs significantly from the mean. It is denoted by the symbol σ^2 and is defined by

$$\sigma^2=E(x^2)-\mu^2 \tag{A.11}$$

The term $E(x^2)$ is simply the expected value of the random variable x^2 and is defined by applying (A.10).

$$E(x^2)=\sum_{z^2=-\infty}^{+\infty} z^2 g(z^2)$$

where $g(z^2)=\mathrm{P}(x^2=z^2)$, but clearly $g(z^2)=f(z)+f(-z)$. This observation yields the usual definition

$$E(x^2)=\sum_{z=-\infty}^{+\infty} z^2 f(z) \tag{A.12}$$

An alternative definition for the variance is $\sigma^2=E((x-\mu)^2)$. If $x_1, \ldots, x_n$ are random variables then

$$E(x_1+\cdots+x_n)=E(x_1)+\cdots+E(x_n) \tag{A.13}$$

To see this, consider the case of two random variables x and y. By (A.10) and (A.8),

$$\begin{aligned} E(x)+E(y) &= \sum_{z_1=-\infty}^{+\infty} z_1\mathrm{P}(x=z_1)+\sum_{z_2=-\infty}^{+\infty} z_2\mathrm{P}(y=z_2) \\ &= \sum_{-\infty\le z_1,z_2\le+\infty} z_1\mathrm{P}(x=z_1\wedge y=z_2)+\sum_{-\infty\le z_1,z_2\le+\infty} z_2\mathrm{P}(x=z_1\wedge y=z_2) \\ &= \sum_{-\infty\le z_1,z_2\le+\infty} (z_1+z_2)\mathrm{P}(x=z_1\wedge y=z_2) \\ &= E(x+y) \end{aligned}$$

We can now apply (A.13) to obtain

$$E((x_1+\cdots+x_n)^2)=\sum_{i=1}^{n}\sum_{j=1}^{n} E(x_i x_j) \tag{A.14}$$

If x is a non-negative random variable with mean μ then

$$P(x \geq 1) \leq \mu \tag{A.15}$$

The proof is quite simple.

$$P(x \geq 1) = \sum_{z=1}^{+\infty} f(z) \leq \sum_{z=1}^{+\infty} z f(z) = \sum_{z=0}^{+\infty} z f(z) = \mu$$

Chebyshev's inequality states that if x is a random variable with mean μ and variance σ^2 then for any $t > 0$

$$P(|x-\mu| \geq t) \leq \frac{\sigma^2}{t^2}$$

The proof is

$$P(|x-\mu| \geq t) = \sum_{|z-\mu| \geq t} f(z) \leq \frac{1}{t^2} \sum_{|z-\mu| \geq t} (z-\mu)^2 f(z) \leq \frac{\sigma^2}{t^2}$$

We can use Chebyshev's inequality to derive the following useful result.

$$P(x=0) \leq \frac{\sigma^2}{\mu^2} \tag{A.16}$$

In addition to the above results from probability theory we will need the following results for doing asymptotic analysis. First is Stirling's approximation for $n!$

$$\left(\frac{n}{e}\right)^n \sqrt{2\pi n} \leq n! \leq \left(\frac{n}{e}\right)^n \sqrt{2\pi n}\, e^{1/12n} \tag{A.18}$$

The following results can be obtained by expanding the binomial coefficient and applying Stirling's approximation.

$$\frac{1}{2\sqrt{a}} \left(\frac{a}{b}\right)^b \leq \frac{1}{2\sqrt{a}} \left(\frac{a}{a-b}\right)^{a-b} \left(\frac{a}{b}\right)^b \leq \binom{a}{b} \leq \left(\frac{a}{a-b}\right)^{a-b} \left(\frac{a}{b}\right)^b \leq \left(\frac{ea}{b}\right)^b \tag{A.19}$$

The following are consequences of the Taylor Series expansion of e^x.

$$e^x \leq 1+x \tag{A.20}$$

$$\lim_{x \to 0} e^x = 1+x \tag{A.21}$$

We will also make use of the geometric series and the arithmetic-geometric series

$$1+r+r^2+\cdots+r^{n-1}=\frac{1-r^n}{1-r} \tag{A.22}$$

$$1+r+r^2+\cdots=\frac{1}{1-r} \quad (-1<r<1) \tag{A.23}$$

$$r+2r^2+3r^3+\cdots=\frac{r}{(1-r)^2} \quad (-1<r<1) \tag{A.24}$$

Bibliography

[1] Cook, Stephen A. "The Complexity of Theorem-Proving Procedures". In *Proceedings Third Annual ACM Symposium on the Theory of Computing,* 1971, 151-158.

[2] Karp, Richard M. "Reducibility Among Combinatorial Problems". In *Complexity of Computer Computations,* R. E. Miller and J. W. Thatcher (eds), Plenum Press, 1972.

[3] Garey, Michael R., David S. Johnson. *Computers and Intractability - A Guide to the Theory of NP-Completeness.* W. H. Freeman, 1979.

[4] Garey, M. R., David S. Johnson. "Approximation Algorithms for Bin Packing Problems: A Survey". TM-80-1216-19, Bell Labs, 6/9/80.

[5] Johnson, David S. *Near-Optimal Bin Packing Algorithms.* Doctoral Thesis, Dept of Mathematics, MIT, 1973.

[6] Garey, M. R., D. S. Johnson. "On Packing Two-Dimensional Bins". TM-80-1216-53, Bell Labs, 11/12/80.

[7] Christofides, N. "Worst-case Analysis of a New Heuristic for the Traveling Salesman Problem". Technical Report, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA., 1976.

[8] Sahni, S., T. Gonzalez. "P-Complete Approximation Problems". In *Journal of the ACM 23* , 1976, 555-565.

[9] Karp, Richard M. "Probabilistic Analysis of Partitioning Algorithms for the Traveling Salesman Problem in the Plane". In *Mathematics of Operations Research 2* , 8/77, 209-224.

[10] Karp, Richard M. "The Probabilistic Analysis of Some Combinatorial Search Algorithms". In *Algorithms and Complexity,* J. Traub (ed), Academic Press, 1976.

[11] Karp, Richard M. "A Patching Algorithm for the Nonsymmetric Traveling Salesman Problem". In *SIAM Journal of Computing 8,* 1/79, 561-573.

[12] Shapiro, Stephen D. "Performance of Heuristic Bin Packing Algorithms with Segments of Random Length". In *Information and Control 35* , 1977, 146-158.

[13] Angluin, D., L. G. Valiant. "Fast Probabilistic Algorithms for Hamiltonian Circuits and Matchings". In *Journal of Computer and System Sciences 18* , 1979, 155-193.

[14] Posa, L. "Hamiltonian Circuits in Random Graphs". In *Discrete Mathematics 14,* 1976, 359-364.

[15] Babai, Laszlo, Paul Erdös, Stanley M. Selkow. "Random Graph Isomorphism". In *Siam Journal of Computing 9,* 8/80, 628-635.

[16] Erdös, P., A. Renyi. "On Random Graphs I.". In *Publicationes Mathematicae,* 1959, 290-297.

[17] Erdös, P., A. Renyi. "On the Evolution of Random Graphs". In *Magyar Tudomanyos Akademia Matematikai Kutato Intezete Kozlemenyei 5,* 1960, 17-61.

[18] Erdös, P. "On Circuits and Subgraphs of Chromatic Graphs". In *Mathematika 9,* 1962, 170-175.

[19] Erdös, P. "Graph Theory and Probability". In *Canadian Journal of Mathematics 11,* 1959, 34-38.

[20] Erdös, P. "Graph Theory and Probability, II". In *Canadian Journal of Mathematics 13*, 1961, 346-352.

[21] Erdös, Paul, Joel Spencer. *Probabilistic Methods in Combinatorics.* Academic Press, 1974.

[22] Bender, Edward A. "Asymptotic Methods in Enumeration". In *SIAM Review 16*, 10/74, 485-515.

[23] Erdös, P., Robin J. Wilson. "On the Chromatic Index of Almost All Graphs". In *Journal of Combinatorial Theory, Series B 23*, 1977, 255-257.

[24] Grimmet, G. R., C. J. H. McDiarmid. "On Colouring Random Graphs". In *Mathematical Proceedings of the Cambridge Philosophical Society 77*, 1975, 313-324.

[25] Walkup, David W. "Matchings in Random Regular Bipartite Digraphs". In *Discrete Mathematics 31*, 1980, 59-64.

[26] Rosen, Richard. "Matrix Bandwidth Minimization". In *ACM National Conference Proceedings 23*, 585-595, 1968.

[27] Cuthill, E., J. McKee. "Reducing the Bandwidth of Sparse Symmetric Matrices". In *ACM National Conference Proceedings 24*, 157-172, 1969.

[28] Arany, Ilona, Lajos Szoda. "An Improved Method for Reducing the Bandwidth of Sparse Symmetric Matrices". In *Information Processing 71*, 1972, 1246-1250.

[29] Cheng, K. Y. "Minimizing the Bandwidth of Sparse Symmetric Matrices". In *Computing 11*, 1973, 103-110.

[30] Liu, Wai-Hung, Andrew H. Sherman. "Comparative Analysis of the Cuthill-McKee and the Reverse Cuthill-McKee Ordering Algorithms for Sparse Matrices". In *SIAM Journal of Numerical Analysis 13*, 198-213, 4/76.

[31] Papadimitriou, C. H. "The NP-Completeness of the Bandwidth Minimization Problem". In *Computing 16*, 1976, 263-270.

[32] Garey, M. R., R. L. Graham, D. S. Johnson, D. E. Knuth. "Complexity Results for Bandwidth Minimization". In *SIAM Journal of Applied Mathematics 34* , 5/78, 477-495.

[33] Saxe, James B. "Dynamic Programming Algorithms for Recognizing Small-Bandwidth Graphs in Polynomial Time". Carnegie-Mellon University Technical Report, 1980.

[34] Monien, B., I. H. Sudborough. "Bandwidth Problems in Graphs". In *Proceedings 18th Annual Allerton Conference on Communication, Control, and Computing*, 650-659, 1980.

[35] Cheng, K. Y. "Note on Minimizing the Bandwidth of Sparse Symmetric Matrices". In *Computing 11*, 1973, 27-30.

[36] Chvatal, V. "A Remark on a Problem of Harary". In *Czechoslovak Math Journal 20*, 95, 1970.

[37] Stockmeyer, L. J. "The Poynomial Time Hierarchy". In *Theoretical Computer Science 3*, 1-22, .

[38] Gavril, F. "Some NP-complete Problems on Graphs". In *Proceedings 11th Conference on Information Sciences and Systems*, John Hopkins University, , 91-95.

[39] Lengauer, Thomas. "Upper and Lower Bounds on the Complexity of the Min-cut Linear Arrangement Problem on Trees". TM-80-1272-9, Bell Labs, 10/28/80.

[40] Lengauer, Thomas. "Upper and Lower Bounds on the Complexity of the Min-cut Linear Arrangement Problem on Trees". Unpublished memorandum., 1980.

[41] Chung, Moon-Jung, Fillia Makedon, Ivan Hal Sudborough, Jonathan Turner. "Plynomial Algorithms for the Min-Cut Problem on Degree Restricted Trees". , Submitted to IEEE FOCS Conference 82, 4/82.

[42] Feller, William. *An Introduction to Probability Theory and Its Applications.* Wiley, 1968.

Vita

Jonathan S. Turner

Born: November 13, 1953,
Boston, Massachusetts

Current Position

Member of Technical Staff
Bell Laboratories
Naperville, Illinois

Education:

Ph.D. Northwestern University
Evanston, Illinois
June 1982

M.S. Northwestern University
Evanston, Illinois
June 1979

B.S.C.S./B.S.E.E.
Washington University
St. Louis, Missouri
May 1977

A.B. Oberlin College
Oberlin, Ohio
May 1977