Computing graph properties by randomized subcube partitions

Ehud Friedgut Jeff Kahn Avi Wigderson

January 30, 2003

Abstract

We prove a new lower bound on the randomized decision tree complexity of monotone graph properties. For a monotone property $\mathcal{A}$ of graphs on n vertices, let $p = p(\mathcal{A})$ denote the threshold probability of $\mathcal{A}$, namely the value of p for which a random graph from $G(n,p)$ has property $\mathcal{A}$ with probability $1/2$. Then the expected number of queries made by any decision tree for $\mathcal{A}$ on such a random graph is at least $\Omega(n^2/\max\{pn, \log n\})$.

Our lower bound holds in the subcube partition model, which generalizes the decision tree model. The proof combines a simple combinatorial lemma on subcube partitions (which may be of independent interest) with simple graph packing arguments. Our approach motivates the study of packing of "typical" graphs, which may yield better lower bounds.

1 Introduction

A decision tree is the most basic computational device. It has unlimited power, and attempts to compute a function of an input string by successively (and adaptively) querying its values in different coordinates. The complexity of such an algorithm is the maximum number of queries used for a worst case input.

A basic problem studied for this model is the complexity of monotone graph properties. Here the input is a graph (on n vertices), and the function is the indicator of a monotone (i.e. upward closed) family of such graphs $\mathcal{A}$.

A fundamental result of Rivest and Vuillemin [**?**] is that *every* such function requires $\Omega(n^2)$ queries. In other words, the trivial algorithm of querying the whole input is not worse than the best algorithm by more than a constant factor. (We assume throughout that properties $\mathcal{A}$ are nontrivial, meaning $\mathcal{A}$ contains some, but not all graphs.) Moreover, further results [**?**, **?**] show that for some values of n every such function is *evasive*, namely the trivial algorithm is best possible—no algorithm can save even one query.

When randomization enters, the story becomes more interesting. A randomized decision tree for a given function is simply a probability distribution over (deterministic) decision trees for that function. Thus, we deal here only with errorless algorithms. The complexity of such an algorithm is the maximum over inputs of the expectation (with respect to the given distribution) of the number of queries. Randomized decision trees can be much more efficient than their deterministic counterparts. The largest gap known is for a function on N bits, which is evasive (hence its deterministic complexity is exactly N), while its randomized complexity is $\Theta(N^{.753..})$. Moreover, this function is transitive (that is, invariant under some group acting transitively on the coordinates).

The question of the randomized complexity of monotone graph properties was raised by several people, and a conjecture attributed to Karp is that for these functions randomization *does not* help—namely that every such property of graphs on n vertices has randomized complexity $\Omega(n^2)$. For some properties, e.g. that of having an isolated vertex, it is easy to see that the randomized complexity is about $n^2/4$, roughly half the trivial bound, but no better upper bound (even for algorithms that are allowed to make a small error) is known for any property. For some specific properties a quadratic lower bound is known.

The first nontrivial lower bounds on the randomized complexity of general graph properties—$\Omega(n(\log n)^{1/12})$—were given by Yao [**?**]. The basic methods introduced by Yao were significantly improved by [**?**, **?**, **?**] to yield the current best lower bound $\Omega(n^{4/3}(\log n)^{1/3})$, still a far cry from the (conjectured tight) upper bound. A slightly better bound of $\Omega(n^{3/2})$ has been proved for the class of all properties defined by a single minterm (that is, the property holds iff the input graph contains a (copy of a) specified graph as a subgraph) [**?**].

In this paper we give a lower bound that depends on the threshold probability of the property. Let $\mathcal{A}$ be a graph property, and let p be such that

$$\Pr[G(n,p) \in \mathcal{A}] = 1/2,$$

where $G(n,p)$ is the usual random graph on n vertices with edge probability p. Then our lower bound is $\Omega(\min\{n^2/\log n,\, n/p\})$. Note that this is nearly quadratic for many known properties (e.g. connectivity, Hamiltonicity, absence of isolated vertices, containing a triangle), and improves the best bounds above as long as $p << n^{-1/3}$. On the other hand, we get no improvement for general properties.

The above bound is actually stronger than stated in two senses. One is that it holds not just in the worst case, but on average for $G(n,p)$ (this is typical of almost all lower bounds). The other, more interesting aspect, is that our lower bound holds in a seemingly stronger model. Note that a decision tree for any function f partitions the inputs into subcubes (determined by its paths), each of which contain inputs with a constant f value. Our model allows *any* such partition, not just one derived from a decision tree. The complexity on a given input is simply the co-dimension of the subcube containing that input (corresponding to the number of positions "read" by the algorithm, or equivalently, to the path length in a decision tree). Then our lower bound on the expected complexity (of an input drawn from $G(n,p)$) holds for any partition as above.

The result is obtained using a simple lemma about such partitions. Given any partition of the Boolean cube $\{0,1\}^N$ to subcubes, pick an input at random using the product distribution assigning a 1 to each coordinate independently with probability p. Then the ratio of the expected number of positions "read" containing 1, and the expected number of those containing 0, is exactly $p/(1-p)$. Like the packing lemmas typically used in such lower bounds, this lemma is used to show (roughly) that if the minterms of the function are small, then the maxterms are large. However, here this trade-off improves as p shrinks. Combining this lemma with standard packing arguments gives our lower bound.

Before going into the proof we wish to raise a natural question, which is probably not relevant to improving the present bounds, but is nonetheless of some theoretical interest. We would like to know whether the random subcube partition model we present is essentially different than the random decision tree model.

Question 1.1 *Are there graph properties, or properties of other families of* *sets, for which the random decision tree complexity has a different asymptotic behavior then the random subcube partition complexity (called RAND(f) below). Are there any "natural" examples of this phenomenon?*

The paper is organized as follows. In section 2 we present the subcube partition model, define basic notions and prove the basic combinatorial theorem for it mentioned above. In section 3 we derive the lower bound on computing graph properties in this model. In section 4 we propose some questions and conjectures on packing "typical" graphs, which seem potentially relevant to improving the current bounds, and also of independent interest.

2 Preliminaries

Consider the Boolean cube $C = \{0,1\}^N$. Let $C = \bigcup C_i$ be a partition of C into subcubes. Every such subcube C_i can be associated with a characteristic vector in $\{0,1,*\}^N$. The 0's and 1's are the coordinates that are fixed in C_i and the *'s are the free coordinates. For every such C_i let $X_1(C_i)$ be the number of 1's in the corresponding vector, and $X_0(C_i)$ the number of 0's. We also set $X(C_i) = X_0(C_i) + X_1(C_i)$, the co-dimension of the subcube C_i.

The deterministic complexity of a partition is simply the maximum value of $X(C_i)$. We will say that a partition of C computes a function f if f is constant on each C_i. Let $DET(f)$ denote the minimum deterministic complexity over all partitions that compute f. Note that this is a lower bound on the deterministic decision tree complexity of f.

The distributional complexity over some probability distribution D on C is defined as follows. Let $x \in C$ be a point chosen at random according to D. Let $i(x)$ be the index of the subcube containing x, i.e. $x \in C_{i(x)}$. Define the random variables

$$X_1 = X_1(C_{i(x)}),\ X_0 = X_0(C_{i(x)}),\ X = X_0 + X_1.$$

Now the distributional complexity for this partition is the expectation of X. The distributional complexity of a function f, denoted $DIST(f, D)$, is the minimum of this value over all partitions computing f. Again, this is a lower bound on the distributional complexity of a decision tree for f on D.

Finally, a randomized partition is a distribution on partitions, and we may consider the expectation of X over such a distribution, for a worst case input x. We let $RAND(f)$ be the minimum of this value over distributions supported on partitions which compute f. Again, this is a lower bound on the randomized decision tree complexity of f.

As observed by Yao for decision trees, the von-Neumann MiniMax theorem [**?**] applies in this context, giving, for every f,

$$RAND(f) = \max_D DIST(f, D).$$

As is common, we will choose a particular D to prove a lower bound on $RAND(f)$. We consider only product distributions $D = \mu_p = \mu$ given by by $\mu(x) = p^{\sum x_i} q^{N - \sum x_i}$, where (here and throughout the paper) we set $q = 1 - p$.

Our approach is based on the following simple observation.

Lemma 2.1 *For any partition* $\{C_i\}$ *of* C *and* x *drawn from* μ_p, $\frac{E(X_1)}{E(X_0)} = \frac{p}{q}$.

We give two proofs of this simple fact.

First proof. It is tempting to try to prove this easy lemma by induction on the number of subcubes. However this does not work, since not every partition into subcubes can be achieved from a coarser one by splitting one of the cubes involved into two; indeed this is what makes the partition model stronger than the decision tree model, in which such a coarser partition always exists—simply merge two maximal paths which diverge only at the last query).

Instead, we will work our way up by reverse induction, starting with the partition into 2^n subcubes and showing that merging two neighboring subcubes does not change $\frac{E(X_1)}{E(X_0)}$.

To begin notice that the lemma is trivial in the case when there are 2^n subcubes (=points). In this case X_1 is the number of 1's in a point chosen at random, $E(X_1) = np$ and $E(X_0) = nq$ so the lemma holds.

Now assume the lemma holds for a given partition: Let $E(X_1)/E(X_0) = p/q$. Let C_1 and C_0 be two neighboring subcubes, i.e. their characteristic vectors are identical except for one coordinate where the C_1-vector has a 1 and the C_0-vector has a 0. Merging the two subcubes into one results in a subcube whose characteristic vector is the same as the two aforementioned vectors except for a * replacing the 1 and the 0 at the coordinate where they disagreed. Note that $\mu(C_1)/\mu(C_0) = p/q$. Let X_1' and X_0' be the new random variables corresponding to the new partition. It is easy to see that $E(X_i') = E(X_i) - \mu(C_i)$ for $1 = 1, 2$ so that

$$\frac{E(X_1')}{E(X_0')} = \frac{E(X_1) - \mu(C_1)}{E(X_0) - \mu(c_0)} = \frac{p}{q}.$$

□

Second proof. For $b \in \{0, 1\}$ and $j \in [N]$ let X_b^j be 1 if the jth coordinate of the characteristic vector of $C_{i(x)}$ is b and 0 otherwise (it is $1-b$ or *). Note that $X_b = \sum_{j\in[N]} X_b^j$, so by additivity of expectation, all we have to prove is that for every j, $E(X_1^j)/E(X_0^j) = p/q$. But in fact this holds even if we condition on an arbitrary setting of all coordinates other than x_j: either this setting determines the subcube containing our input, in which case there is no contribution to either numerator or denominator; or it does not, in which case the subcube containing the input is determined by the jth bit, and there is a contribution of p to the numerator and q to the denominator.

□

3 Graph properties

A graph is associated with its characteristic vector of 1's (edges) and 0's (non-edges). Graphs on n vertices are thus represented by binary vectors of length $N = \binom{n}{2}$.

Let $\mathcal{A}$ be a (nontrivial) monotone graph property of graphs on n vertices. "Graph property" means that as a Boolean function on $C = \{0, 1\}^N$, $\mathcal{A}$ is invariant under the action induced on edges of the symmetric group S_n (acting on the vertices). Slightly abusing notation, we use $\mathcal{A}$ to denote the set of graphs that have the property, and also the corresponding characteristic function: $\mathcal{A}(G)$ is 1 or 0 depending on whether or not $G \in \mathcal{A}$.

The product measure μ_p on C now becomes the standard measure defining $G(n, p)$, the random graph on n vertices with edge probability p.

Let $p = p(\mathcal{A})$ be the threshold probability for $\mathcal{A}$; that is,

$$\Pr[G(n, p) \in \mathcal{A}] = 1/2.$$

Note that there exists a (unique) such p because $\mu_{\mathcal{A}}(p) = \Pr[G(n, p) \in \mathcal{A}]$ is a continuous, strictly increasing function of $p \in [0, 1]$ with $\mu_{\mathcal{A}}(0) = 0, \mu_{\mathcal{A}}(1) = 1$. By replacing $\mathcal{A}$ by its dual (the family of graphs whose complements are not in $\mathcal{A}$) if necessary, we may assume that $p(\mathcal{A}) \le 1/2$. We will in fact assume in what follows that $p(\mathcal{A}) = o(1)$, since for constant p our result is much weaker than known bounds.

Let C be partitioned into subcubes that compute $\mathcal{A}$ (recall this means that the graphs represented by the vectors in each subcube either all belong to $\mathcal{A}$ or all do not). Let X_1, X_0 and X be defined as before with respect to the given partition and the product measure μ_p with $p = p(\mathcal{A})$.

Theorem 3.1 $\max\{E(X_0), E(X_1)\} \geq (1 - o(1)) \min\{\frac{n}{64p}, \frac{n^2}{256 \log n}\}$.

Let MIN_1 denote the number of edges in a smallest graph with property $\mathcal{A}$ (i.e. the size of a smallest minterm of the function $\mathcal{A}$). Let MIN_0 denote the number of edges in the complement of the largest graph not in $\mathcal{A}$ (size of a smallest maxterm of $\mathcal{A}$).

Observation 3.2 $E(X_0) \geq MIN_0/2$

Proof: Note first that if $\mathcal{A}(G) = 0$ for (all) $G \in C_i$ then $X_0(C_i) \geq MIN_0$. Consequently, $E(X_0|x \notin \mathcal{A}) \geq MIN_0$. The result now follows since $\Pr(x \notin \mathcal{A}) = 1/2$)

□

Proof of Theorem ??: If $E(X_1) \geq \frac{n}{64}$ then by Lemma **??** $E(X_0) \geq \frac{qn}{64p}$ and we are done. So let us assume $E(X_1) \leq \frac{n}{64}$. Writing Δ for maximum degree, let

$$\mathcal{B} = \{G : \Delta(G) \leq np + \sqrt{4np \log n}\ \}.$$

According to the Chernoff bound, the probability of a given vertex of $G(n,p)$ having degree more than $np + \sqrt{4np \log n}$ is $O(1/n^2)$; hence $\mu_p(\mathcal{B}) = 1 - O(1/n)$. Let $\mathcal{C} = \mathcal{B} \cap \mathcal{A}$. Then $\mu_p(\mathcal{C}) = 1/2 - O(1/n)$, so that

$$E(X_1|x \in \mathcal{C}) \leq (1/2 - O(1/n))^{-1} E(X_1) \leq (1 + o(1))\frac{n}{32}.$$

In particular, this implies the existence of a graph $G^* \in \mathcal{A}$ on n vertices with at most $(1 + o(1))\frac{n}{32}$ edges and $\Delta(G^*) \leq np + \sqrt{4np \log n}$.

Graphs G and H on vertex set $[n]$ are said to *pack* if there is some permutation $\sigma \in S_n$ for which $\sigma(H) \subseteq \overline{G}$.

Lemma 3.3 *Let G and H be n-vertex graphs with $|G| < (1 + o(1))n/32$, $\Delta(G) \leq np + \sqrt{4np \log n}$, and $|H| < \frac{n^2}{16(np + \sqrt{4np \log n})}$. Then G and H pack.*

Corollary 3.4 *If* $E(X_1) \leq \frac{n}{64}$ *then*

(a) $MIN_0 \geq \frac{n^2}{16(np+\sqrt{4np\log n})}$,

(b) $MIN_0 \geq \left(\min\{\frac{n}{32p}, \frac{n^2}{128\log n}\}\right)$, *and*

(c) $E(X_0) \geq \left(\min\{\frac{n}{64p}, \frac{n^2}{256\log n}\}\right)$.

Of course this gives Theorem **??**.

Proof of corollary. Since $G^* \in \mathcal{A}$, no maxterm H of $\mathcal{A}$ can pack with G; so Lemma **??** gives (a), which immediately gives (b); and (c) then follows from Observation **??**.

$\square$

Proof of Lemma ??. We will use the following packing lemma due to Catlin [**?**] and Sauer and Spencer [**?**].

Lemma 3.5 *For n-vertex graphs G and H,if*

$$\Delta(H)\Delta(G) \leq n/2,$$

then H and G pack.

Now let G, H be as in Lemma **??**. Notice that since G has at most $(1+o(1))\frac{n}{32}$ edges it has at most $(1+o(1))\frac{n}{16}$ nonisolated vertices. It thus suffices to pack (say) some $(n/2)$-vertex subgraph G' of G containing all nonisolated vertices of G, with any spanned subgraph of H on $n/2$ vertices. Let H' be the subgraph of H spanned by the (some) $n/2$ vertices of lowest degree. Since the average degree in H is at most $\frac{n}{8(np+\sqrt{4np\log n})}$, we have $\Delta(H') \leq \frac{n}{4(np+\sqrt{4np\log n})}$. Thus

$$\Delta(H')\Delta(G') < n/4 = \frac{n/2}{2},$$

and by Lemma **??** G' and H' do indeed pack.

$\square$

4 Packing Families of Random Graphs

The approach we follow in this paper gives rise to a line of questions that seems to be of great theoretical interest, both since it is essential for exploiting this technique further, and because it is a natural extension of the much-studied topic of graph packing.

As the reader may have noticed, at the end of the previous section we guaranteed the packing of two graphs G^* and H, using the fact that G^* was something like a subgraph of a random graph. The canonical examples showing that the classical graph packing theorems are tight use examples such as the impossibility of packing a complete matching with a star of degree $n-1$. But in our setting—involving graphs that have a reasonable chance of showing up in a random graph—one does not find such high degree vertices, and average degree becomes a more relevant parameter.

Let us begin with a rather bold conjecture that, using the technique in this paper, would immediately imply, up to log factors, the $\Omega(n^2)$ lower bound for randomized decision tree complexity.

Let $\mathcal{A}$ be a monotone graph property, and G a graph. We wish to find a small witness to the fact that $G \in \mathcal{A}$ or $G \notin \mathcal{A}$. Let

$$Witness(G, \mathcal{A}) = \begin{cases} \min\{|H| : H \subseteq G, H \in \mathcal{A}\} \text{ if } G \in \mathcal{A} \\ \min\{|H| : H \subseteq \overline{G}, H \notin \mathcal{A}\} \text{ if } G \notin \mathcal{A} \end{cases}$$

Recall that $p(\mathcal{A})$ is that (threshold) probability p for which

$$\Pr(G(n,p) \in \mathcal{A}) = 1/2.$$

Our conjecture is that at the threshold for $\mathcal{A}$, in the random graph $G = G(n,p)$ typically either a smallest witness for membership in $\mathcal{A}$ uses $\Omega(1/\log(n))$ of the edges of G or a smallest witness for non-membership uses $\Omega(1/\log(n))$ of the edges of $\overline{G}$ (or both):

Conjecture 4.1 *Let $\mathcal{A}$ be a monotone graph property of graphs on n vertices and $p = p(\mathcal{A})$. Then for $G = G(n,p)$ either*

$$E[Witness(G, \mathcal{A}) | G \in \mathcal{A}] = \Omega(n^2 p / \log(n))$$

or

$$E[Witness(G, \mathcal{A}) | G \notin \mathcal{A}] = \Omega(n^2 (1-p) / \log(n)).$$

As a start, this conjecture seems quite fascinating even for balanced graph properties, those for which $p(\mathcal{A}) = 1/2$. In fact in this case we believe the $\log(n)$ factor may be unnecessary, as, for example, we can show to be true when $\mathcal{A}$ is the property of containing a clique of size k ($\approx 2\log_2(n)$). In this case each minimal witness for membership in $\mathcal{A}$ has size of order $\log^2(n)$, but it turns out that a typical witness for non-membership is of size $\Omega(n^2)$; namely we can prove that with high probability *every* subgraph H of G of size less than $n^2/100$ edges packs with a clique of size k. This illustrates the difference between packing arbitrary graphs and packing subgraphs of random graphs. We know there is a Turán-type graph of size $O(n^2/\log n)$ that does not pack with a clique of size k; but such a graph almost never occurs as a subgraph of G.

Though packing remains in some sense the heart of the matter, real progress (on randomized complexity) probably cannot be based on new packing theorems for pairs of random-like graphs; rather, we need to understand something about packing two *families* of graphs.

Definition 4.2 *Let $\mathcal{A}$, $\mathcal{B}$ be families of graphs. We say $\mathcal{A}$ and $\mathcal{B}$ pack if there exist $A \in \mathcal{A}$ and $B \in \mathcal{B}$ such that A and B pack.*

The relevance of this notion derives from the fact that a family $\mathcal{A}$ and its dual cannot pack. For example Conjecture **??** implies the following statement, which seems to be of independent interest. (Apply the conjecture to the property obtained from $\mathcal{A}$ by adding all graphs with at least $n^2/4$ edges whose complements are not in $\mathcal{B}$.)

Conjecture 4.3 *Let $\mathcal{A}$ and $\mathcal{B}$ be monotone graph properties on n vertices generated by minterms of size $o(n^2/\log n)$, and for which $p(\mathcal{A}), p(\mathcal{B}) > \Omega(1)$. Then $\mathcal{A}$ and $\mathcal{B}$ pack.*

We close with one more question in a similar vein, though now just for pairs of graphs. For a graph G on n vertices, set $\mu(G) = \Pr(G \subset G(n, 1/2))$ (where the containment is up to isomorphism). It is easy to see that if each of $\mu(G)$, $\mu(H)$ is more than $1/2$, then G and H pack; but we expect more is true:

Conjecture 4.4 *For each $\alpha > 0$, if n is sufficiently large and G, H are n-vertex graphs with $\mu(G), \mu(H) \geq \alpha$, then G and H pack.*

(And of course one may ask how quickly $\alpha = \alpha(n)$ can go to zero without falsifying the conjecture.)

Though we have given only a small sample, we hope it is enough to indicate the wealth of possibilities suggested by the present point of view.

References

[B78] B. Bollobás, Extremal Graph Theory, Academic Press, Chapter 8, 1978.

[C74] P. A. Catlin, *Subgraphs of graphs I*, Discrete Math, 10, pp. 225–233, 1974.

[CKS02] A. Chakrabarti, S. Khot and Y. Shi, *Evasiveness of subgraph containment and related properties*, SIAM J. on Computing, 31, pp. 866–875, 2002.

[CK01] A. Chakrabarti and S. Khot, *Improved lower bounds on the randomized complexity of graph properties*, Proc. of the 28th ICALP conference, Lecture Notes in Computer Science 2076, pp. 285–296, 2001.

[G92] D. H. Groğer, *On the randomized complexity of monotone graph properties*, Acta Cybernetica, 10, pp. 119–127, 1992.

[H91] P. Hajnal, *An $\Omega(n^{4/3})$ lower bound on the randomized complexity of graph properties*, Combinatorica, 11, pp. 131–143, 1991.

[K88] V. King, *Lower bounds on the complexity of graph properties*, Proc. of the 20th STOC conference, pp. 468–476, 1988.

[KSS84] J. Kahn, M. Saks and D. Sturtevant, *A topological approach to evasiveness*, Combinatorica 4, pp. 297–306, 1984.

[RV76] R. Rivest and J. Vuillemin, *On recognizing graph properties from adjacency matrices*, TCS 3, pp. 371–384, 1976.

[SS78] N. Sauer and J. Spencer, *Edge-disjoint placement of graphs*, J. of Combinatorial Theory Ser. B, 25, pp. 295–302, 1978.

[SW86] M. Saks and A. Wigderson, *Probabilistic Boolean Decision Trees and the Complexity of Evaluating Game Trees*, Proc. of the 27th FOCS conference, pp. 29–38, 1986.

[vN28] J. von Neumann, *Zur Theorie der Gesellschaftspiele*, Matematische Annalen, 100, pp. 295–320, 1928. English translation *On the theory* *of games and strategy*, in “Contributions to the theory of Games”, IV, pp.13–42, 1959.

[Y87] A.C. Yao, *Lower bounds to randomized algorithms for graph properties*, Proc. of the 28th FOCS conference, pp. 393–400, 1987.