Electronic Colloquium on Computational Complexity, Report No. 89 (2006)

A Note on Adaptivity in Testing Properties of Bounded Degree Graphs

Sofya Raskhodnikova* Adam Smith*

July 6, 2006

Abstract

We show that in the bounded degree model for graph property testing, adaptivity is essential. An algorithm is *non-adaptive* if it makes all queries to the input before receiving any answers. We call a property *non-trivial* if it does not depend only on the degree distribution of the nodes. We show that every tester for a non-trivial property that makes $o(\sqrt{n}/d)$ queries to the input graph on n vertices of degree at most d has to be adaptive.

1 Introduction

Property Testing [RS96, GGR98] is a relatively new framework for approximation algorithms that has received a lot of attention (see [Fis01, Ron01] for surveys). Unlike in the dominant approach to approximation algorithms, where the goal is to obtain an approximation to the cost of the optimal solution, in property testing the goal is to approximate the distance of a given instance to an instance that has a desired property. This framework has yielded *sublinear algorithms* for many problems. Sublinear algorithms run in less time than it would take to read the input. Such efficiency, even at the cost of accuracy, is crucial when processing massive datasets.

Two main models have been proposed for studying properties of graphs in this framework: the *adjacency matrix model* [GGR98], suitable for dense graphs, and the *adjacency lists model*, suitable for graphs of bounded degree [GR02][1]. The adjacency matrix model has been studied much more thoroughly than the adjacency lists model. In a recent breakthrough, Alon et al. [AFNS06] gave a combinatorial characterization of graph properties that can be tested in the adjacency matrix model in time independent of the size of the input.

Our current understanding of the adjacency lists model, proposed by Goldreich and Ron [GR02], is less complete. This note takes a small step towards painting a more complete picture. In this model, a graph on n vertices of degree at most d is represented by a function of vertex v and integer i, where $v \in \{1, \ldots, n\}$ and $i \in \{1, \ldots, d\}$. The value of the function at (v, i) is the ith neighbor of v or a special symbol signifying that v does not have an ith neighbor. The distance between two n-vertex graphs is the number of vertex pairs which are connected by an edge in exactly one of the two graphs divided by the total possible number of edges, dn. A graph is *ϵ-far* from having the desired property if the distance from the graph to any graph satisfying the property is at least ϵ. An algorithm for a particular property gets n, d, the approximation parameter ϵ and oracle

*Weizmann Institute of Science, Rehovot, Israel. Email: `<firstname>.<lastname>@weizmann.ac.il`

[1] A hybrid model, suitable for general graphs has also been considered [AKKR06], but it is beyond the scope of this note.

ISSN 1433-8092

access to the function that defines the input graph. The algorithm has to accept graphs with the desired property (with probability at least $\frac{2}{3}$) and reject graphs that are ϵ-far from the property (with probability at least $\frac{2}{3}$). In one step, the algorithm can query the value of the function at one location (v, i). The *query complexity* of an algorithm is the number of queries it makes in the worst case. The query complexity of an algorithm is a lower bound on its running time.

In the adjacency matrix model, the role of adaptivity is well understood. Goldreich and Trevisan [GT01] proved that in that model, without loss of generality, one can consider only very simple algorithms, which inspect a random induced subgraph. In particular, these algorithms are *non-adaptive*, that is, their queries do not depend on answers to previous queries. More specifically, an arbitrary q-query algorithm for a graph property in the adjacency matrix model can be simulated by a $2q^2$-query non-adaptive algorithm. Thus, adaptivity is not essential in this model.

We show that unlike in the adjacency matrix model, in the adjacency lists model adaptivity is essential. So, it is not a coincidence that the two techniques that were successfully employed in this model – the breadth-first search from a random point [GR02] and a random walk [GR00] – are adaptive. We prove that for all *non-trivial* properties in the bounded degree model, algorithms that make $\Omega\left(\sqrt{n}/d\right)$ queries must be adaptive.

Definition 1 *We call a graph property* non-trivial *if it does not depend only on the degree distribution of the nodes: namely, for each n there is some degree sequence $d_1, \ldots, d_n \in \{0, 1, \ldots, d\}$ such that there is at least one graph, G_1, with node degrees $d_1, \ldots, d_n$ with the property and at least one, G_2, that is ϵ-far.*

To the best of our knowledge, all properties that have been studied in the literature are *non-trivial*. Examples include connectedness (i.e., is the input graph connected?) [GR02], bipartiteness (i.e., is the input graph bipartite?) [GR99], and being an expander [GR00].

Given the right definition of a *non-trivial* graph property, our result is not hard to prove. The main idea is that one cannot distinguish a random isomorphic copy of G_1 from a random isomorphic copy of G_2 with $o\left(\sqrt{n}/d\right)$ non-adaptive queries and error probability $\leq 1/3$. Therefore, by Yao's minimax principle [Yao77] all ϵ-testers for non-trivial graph properties in the bounded degree model with query complexity $o\left(\sqrt{n}/d\right)$ must be adaptive.

2 Adaptivity in the exploration of bounded degree graphs

Before we state and prove our result, we introduce notation used in the rest of the paper. Recall that the statistical distance between distributions $\mathcal{D}_1$ and $\mathcal{D}_2$ is defined as:

$$CD(\mathcal{D}_1, \mathcal{D}_2) = \max_{S=\text{support}(\mathcal{D}_1)\cup\text{support}(\mathcal{D}_2)} \left(\left| \Pr_{x \leftarrow \mathcal{D}_1}[x \in S] - \Pr_{x \leftarrow \mathcal{D}_2}[x \in S] \right| \right).$$

In what follows, $\mathcal{D}_1 \approx_\delta \mathcal{D}_2$ denotes that the statistical difference between the distributions $\mathcal{D}_1$ and $\mathcal{D}_2$ is at most δ.

Theorem 2 *Every non-adaptive tester for a* non-trivial *property in the* adjacency lists model *requires $\Omega\left(\sqrt{n}/d\right)$ queries.*

To simplify the argument, we will give the tester a little more power: every time it queries a neighbor i of vertex v, it will get the entire adjacency list of vertex v, i.e., it will find out a "star" portion of the graph with v in the center and its neighbors connected to it. The main idea is that with $q = o(\sqrt{n}/d)$ queries even the enhanced (non-adaptive) tester will see only a disjoint collection of stars with high probability. Therefore, the only information the tester will be able to collect is the degrees of q vertices.

To prove the lower bound formally, we will use Yao's principle (the version with two distributions, see Claim 5). Let G_1 and G_2 be as stated in Definition 1. Let $\mathcal{P}$ be a random isomorphic copy of G_1 and $\mathcal{N}$ be a random isomorphic copy of G_2. Without loss of generality assume that the (deterministic) tester queries the adjacency lists of nodes $1, \ldots, q$. Let $a_1 \ldots a_q(G)$ be the answers to the queries on input G. Define $\mathcal{P}$-view to be the distribution on $a_1 \ldots a_q(G)$ when G is selected according to $\mathcal{P}$. Similarly, define $\mathcal{N}$-view. As proved in the Appendix, it is enough to show that $\mathcal{P}$-view $\approx_{\frac{1}{4}}$ $\mathcal{N}$-view.

Let BAD denote the event that two stars centered at query points intersect, namely, that for some pair of queries v and u there is some vertex s such that both (u, s) and (v, s) are edges in the input graph. Let I be a random variable that denotes the number of such intersecting pairs. If v_i is the node that got mapped to node i under random isomorphism of graph G then the set containing v_i, the neighbors of v_i and the neighbors of neighbors of v_i has at most $d^2 + 1$ nodes. Under random isomorphism, the probability that one of these nodes is mapped to node j is at most $\frac{d^2+1}{n}$. Therefore, for both distributions $\mathcal{P}$-view and $\mathcal{N}$-view,

$$E[I] \le \binom{q}{2} \frac{d^2 + 1}{n} \le \frac{1}{17}, \text{ for sufficiently large } n.$$

Consequently,

$$\Pr[BAD] = \Pr[I > 0] = \sum_{i=1}^{\infty} \Pr[I = i] \le \sum_{i=1}^{\infty} i \Pr[I = i] = E[I] \le \frac{1}{17}.$$

We will show that conditioned on BAD not occurring, (1) the tester learns only the degrees of the queried nodes, and (2) the distributions on the degree list seen by the tester are similar under $\mathcal{P}$ and $\mathcal{N}$. For a graph G, let $d_1, ..., d_q(G)$ be the degrees of the vertices of G queried by the tester. Let $\mathcal{P}$-degs be the distribution $d_1, ..., d_q(G)$ when G is selected according to $\mathcal{P}$. Similarly, define $\mathcal{N}$-degs. Observe that $\mathcal{P}$-degs $= \mathcal{N}$-degs because of our condition on the degrees of G_1 and G_2.

To show that knowing adjacency lists does not give any advantage over knowing only the degree list when BAD does not occur, define a randomized algorithm $\mathcal{A}$ that converts a degree list to a possible set of answers. On input $d_1, ..., d_q$, $\mathcal{A}$ picks $d = \sum_{i=1}^{q} d_i$ random numbers from $\{q+1, ..., n\}$ *without replacement* and outputs those numbers in order as elements of the adjacency lists of the nodes $1, \ldots, q$. Note that $\mathcal{A}$ always produces non-intersecting adjacency lists (i.e., $\mathcal{A}$ simulates a world where BAD never happens).

Claim 3 Conditioned on BAD not occurring, *the output of $\mathcal{A}$ is distributed according to*

- $\mathcal{P}$-view *when its input is distributed according to* $\mathcal{P}$-degs*;*
- $\mathcal{N}$-view *when its input is distributed according to* $\mathcal{N}$-degs.

That is, in symbols, $\mathcal{A}(\mathcal{P}\text{-degs}|_{\overline{BAD}}) = \mathcal{P}\text{-view}|_{\overline{BAD}}$ *and* $\mathcal{A}(\mathcal{N}\text{-degs}|_{\overline{BAD}}) = \mathcal{N}\text{-view}|_{\overline{BAD}}$.

Proof We prove the claim only for distribution $\mathcal{P}$. The same proof works for $\mathcal{N}$. First, observe that the distribution on lists of degrees in $\mathcal{P}\text{-view}|_{\overline{BAD}}$ and in $\mathcal{A}(\mathcal{P}\text{-degs}|_{\overline{BAD}})$ is the same: in both cases it is $\mathcal{P}\text{-degs}|_{\overline{BAD}}$, by definition. Thus, it is sufficient to prove that for each possible degree list $d_1, ..., d_q$, the distribution on neighbor lists $a_1, ..., a_q$ is the same in both distributions.

Consider any two non-intersecting sequences of adjacency lists $a_1, ..., a_q$ and $a'_1, ..., a'_q$ which correspond to the same degree list $d_1, \ldots, d_q$. Since $\mathcal{A}(d_1, ..., d_q)$ selects a non-intersecting sequence uniformly at random (from the set of non-intersecting sequences with degrees $d_1, \ldots, d_q$), it outputs both sequences with the same probability. We will show that they also arise with same probability under $\mathcal{P}$-view. Note that there exists some permutation π of G_1, such that the nodes in $a'_1, ...a'_q$ are the images of the nodes in $a_1, ..., a_q$ under π (since, in both cases, no node appears twice). We can now set up a 1-to-1 correspondence between permutations that give rise to $a_1, ..., a_q$ and permutations that give rise to $a'_1, ..., a'_q$: for any permutation σ of $\{1, ..., n\}$ such that $\sigma(G_1)$ has adjacency lists $a_1, ..., a_q$, the permutation $\pi \circ \sigma$ produces $a'_1, ..., a'_q$; similarly, if σ led to $a'_1, ..., a'_q$, then $\pi^{-1} \circ \sigma$ would lead to $a_1, ..., a_q$. This correspondence is 1-to-1 since we can never have $\pi \circ \sigma_1 = \pi \circ \sigma_2$ unless $\sigma_1 = \sigma_2$.

Because of this correspondence, the two sequences of adjacency lists arise with same probability under $\mathcal{P}$-view, and so $\mathcal{P}\text{-view}|_{\overline{BAD}} = \mathcal{A}(\mathcal{P}\text{-degs}|_{\overline{BAD}})$. (Note that the equality would not hold without conditioning on $\overline{BAD}$). ■

Conditioning on $\overline{BAD}$ does not significantly change our distributions, as formalized in claim 4.

Claim 4 *Let E be an event that happens with probability at least $1 - \delta$ under the distribution $\mathcal{D}$ and let $\mathcal{B}$ denote distribution $\mathcal{D}|_E$. Then $\mathcal{B} \approx_{\delta'} \mathcal{D}$ where $\delta' = \frac{1}{1-\delta} - 1$.*

Proof It is enough to show that $\Pr_{\mathcal{B}}[S] \le \Pr_{\mathcal{D}}[S] + \delta'$ for every event S.

$$\Pr_{\mathcal{B}}[S] = \Pr_{\mathcal{D}}[S|E] = \frac{\Pr_{\mathcal{D}}[S \wedge E]}{\Pr_{\mathcal{D}}[E]} \le \frac{\Pr_{\mathcal{D}}[S]}{\Pr_{\mathcal{D}}[E]} \le \frac{\Pr_{\mathcal{D}}[S]}{1-\delta} = \Pr_{\mathcal{D}}[S](1+\delta') \le \Pr_{\mathcal{D}}[S] + \delta'. \blacksquare$$

In particular, if $\delta = \frac{1}{17}$ then $\delta' = \frac{1}{16}$. We will apply the claim four times with these parameters to prove that $\mathcal{P}\text{-view} \approx_{\frac{1}{4}} \mathcal{N}\text{-view}$. First, $\mathcal{P}\text{-view} \approx_{\frac{1}{16}} \mathcal{P}\text{-view}|_{\overline{BAD}}$ and $\mathcal{P}\text{-degs} \approx_{\frac{1}{16}} \mathcal{P}\text{-degs}|_{\overline{BAD}}$. The second statement implies that $\mathcal{A}(\mathcal{P}\text{-degs}) \approx_{\frac{1}{16}} \mathcal{A}(\mathcal{P}\text{-degs}|_{\overline{BAD}})$. Putting the two statements together and applying Claim 3 gives $\mathcal{P}\text{-view} \approx_{\frac{1}{8}} \mathcal{A}(\mathcal{P}\text{-degs})$. Similarly, $\mathcal{N}\text{-view} \approx_{\frac{1}{8}} \mathcal{A}(\mathcal{N}\text{-degs})$. It remains to use that $\mathcal{P}\text{-degs} = \mathcal{N}\text{-degs}$ and, consequently, $\mathcal{A}(\mathcal{P}\text{-degs}) = A(\mathcal{N}\text{-degs})$, yielding $\mathcal{P}\text{-view} \approx_{\frac{1}{4}} \mathcal{N}\text{-view}$, as required.

Acknowledgement. The theorem proved in this note was initially given as a homework problem in a course on sublinear algorithms at the Weizmann Institute of Science, but with an incorrect definition of “non-trivial”. As students complained and the definition was refined, it became apparent that the statement was not as easy to prove as intended.

References

[AFNS06] Noga Alon, Eldar Fischer, Ilan Newman, and Asaf Shapira. A combinatorial characterization of the testable graph properties: It's all about regularity. In *ACM Symposium on Foundations of Computer Science*, 2006.

[AKKR06] Noga Alon, Tali Kaufman, Michael Krivelevich, and Dana Ron. Testing triangle-freeness in general graphs. In *ACM Symposium on Discrete Algorithms*, pages 279–288, 2006.

[Fis01] E. Fischer. The art of uninformed decisions: A primer to property testing. *Bulletin of the European Association for Theoretical Computer Science*, 75:97–126, 2001.

[GGR98] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *JACM*, 45(4):653–750, 1998.

[GR99] Oded Goldreich and Dana Ron. A sublinear bipartiteness tester for bounded degree graphs. In *Combinatorica*, volume 19, pages 335–373, 1999.

[GR00] O. Goldreich and D. Ron. On testing expansion in bounded-degree graphs. *Electronic Colloqium on Computational Complexity*, 7(20), 2000.

[GR02] O. Goldreich and D. Ron. Property testing in bounded degree graphs. *Algorithmica*, 32(2):302–343, 2002.

[GT01] Oded Goldreich and Luca Trevisan. Three theorems regarding testing graph properties. In *IEEE Symposium on Foundations of Computer Science*, pages 460–469, 2001.

[Ron01] D. Ron. Property testing. In *Handbook on Randomization, Volume II*, pages 597–649, 2001.

[RS96] R. Rubinfeld and M. Sudan. Robust characterization of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, 1996.

[Yao77] A. C. Yao. Probabilistic computation, towards a unified measure of complexity. In *Proceedings of the Eighteenth Annual Symposium on Foundations of Computer Science*, pages 222–227, 1977.

A Yao's Principle

Yao's Principle states that in order to prove a lower bound for a randomized algorithm, it is enough to give a distribution $\mathcal{D}$ on inputs, for which the lower bound holds for any deterministic algorithm. It is often more convenient to look at positive and negative instances separately. For completeness, we state and prove the resulting version of Yao's Principle.

Claim 5 [Folklore]*: To prove a lower bound q on the worst-case query complexity of a randomized algorithm, it is enough to give two distributions on inputs:*

- *$\mathcal{P}$ on positive instances, and*
- *$\mathcal{N}$ on negative instances*

such that it is hard for any q-query deterministic algorithm to distinguish $\mathcal{P}$ from $\mathcal{N}$. Let $a_1 \ldots a_q(x)$ be the answers to the queries on input x. Define $\mathcal{P}$-view to be the distribution on $a_1 \ldots a_q(x)$ when x is selected according to $\mathcal{P}$. Similarly, define $\mathcal{N}$-view. By "hard to distinguish $\mathcal{P}$ from $\mathcal{N}$", we mean $CD\,(\mathcal{P}\text{-view}, \mathcal{N}\text{-view}) < \frac{1}{3}$.

We will prove this alternative formulation of Yao's Principle, using the mainstream formulation.
Proof Let $\mathcal{A}$ be any (adaptive) deterministic q-query tester. Given distributions $\mathcal{P}, \mathcal{N}$ with $CD\,(\mathcal{P}\text{-view}, \mathcal{N}\text{-view}) < \frac{1}{3}$ for all such testers $\mathcal{A}$, we define a distribution $\mathcal{D}$, as required in the mainstream version of Yao's Principle.

Let S be the set of strings $a_1 ... a_q$ on which $\mathcal{A}$ accepts.

$$\begin{aligned}\left|\Pr_{x\leftarrow\mathcal{P}}[\mathcal{A}(x)=1] - \Pr_{x\leftarrow\mathcal{N}}[\mathcal{A}(x)=1]\right| &= \left|\Pr_{a\leftarrow\mathcal{P}\text{-view}}[a\in S] - \Pr_{a\leftarrow\mathcal{N}\text{-view}}[a\in S]\right| \\ &\le SD(\mathcal{P}\text{-view}, \mathcal{N}\text{-view}) < \frac{1}{3}.\end{aligned}$$

If algorithm $\mathcal{A}$ accepts an input distributed according to $\mathcal{P}$ with probability $< \frac{2}{3}$, we set $\mathcal{D} = \mathcal{P}$. Otherwise, $\mathcal{A}$ accepts an input distributed according to $\mathcal{N}$ with probability $> \frac{2}{3} - \frac{1}{3} = \frac{1}{3}$. In this case, we set $\mathcal{D} = \mathcal{N}$. The algorithm A is correct on inputs distributed according to $\mathcal{D}$ with probability less than $\frac{2}{3}$. In other words, there is no deterministic q-query algorithm that works on distribution $\mathcal{D}$, so we can conclude, by the original formulation of Yao's Principle that there is no q query probabilistic algorithm for the problem. ■

ECCC ISSN 1433-8092
http://eccc.hpi-web.de/