Errata to the paper
Three Theorems regarding Testing Graph Properties

Oded Goldreich
Department of Computer Science
Weizmann Institute of Science
Rehovot, Israel.
oded.goldreich@weizmann.ac.il

Luca Trevisan
EECS – Computer Science Div.
UC-Berkeley
Berkeley, CA 94720, USA.
luca@eecs.berkeley.edu

August 5, 2005

The following error in our paper [4] was discovered by Asaf Shapira. In [4, Sec. 2.2], testers for graph properties are formally defined as receiving the number of vertices (denoted N) and the proximity parameter (denoted ϵ) as explicit inputs. Unfortunately, the rest of our paper (specifically, Sections 4 and 5) fails to properly account for this convention. In particular, the presentation in [4] presumes that certain actions of any relevant tester are independent of N. Consequently, Theorem 2 is not valid (as stated in [4]), and the proof of Theorem 3 (as presented in [4, Sec. 5]) needs to be slightly refined. Below we sketch the adequate modifications, referring to the terminology and notation of [4]. In particular, we show that a non-uniform relaxation of [4, Thm. 2] is valid, whereas [4, Thm. 3] is valid as stated.

We mention that the issue of dependence of graph property testers on the number of vertices is crucial also in the recent works of Alon and Shapira [1, 2].

1 Modifying Theorem 2

The problem with Theorem 2 (of [4]) is that a canonical tester for property Π was defined as one that uniformly selects a sample of vertices, and accepts iff the subgraph induced by this vertex-set has property Π', *where* Π' *is a fixed graph property that only depends on* Π. Theorem 2 asserts that any property that is testable within query complexity $q(N,\epsilon)$ admits a canonical tester of query complexity $O(q(N,\epsilon)^2)$, and that in the case of one-sided error a sample of $2q(N,\epsilon)$ vertices suffices.

To see that Theorem 2 is not valid as stated, consider the property of being a graph with an odd number of vertices, and the tester that makes no queries but rather accepts the graph iff N (given to it explicitly) is odd. In this case Theorem 2 asserts a canonical tester of query complexity zero, which is impossible (because the canonical tester is supposed to decide based solely on the induced subgraph). This specific counterexample may be eliminated by allowing the canonical tester to sample $2q(N,\epsilon)+1$ vertices (in the case of one-sided error and $O(q(N,\epsilon)+1)$ vertices in general). In this case, we can prove the claim by presenting a canonical tester that selects $N \bmod 2$ vertices, and accepts iff the subgraph induced by this vertex-set has an odd number of vertices. (Needless to say, this is not the canonical tester that emerges from the original proof of Theorem 2.)

We do not know whether the forgoing modification of [4, Thm. 2] is valid, but certainly if this modification is valid then its proof must go beyond the presentation and the ideas of [4]. For example, consider the property of being a bipartite graph with a number of vertices that is a prime number. Building on the tester of [3], on input (N,ϵ), a canonical tester may proceed as follows. If N is prime, it selects a random sample of $2\lceil \mathrm{poly}(1/\epsilon)\rceil + 1$ vertices, and otherwise it selects a sample of two vertices. In either case, the tester accepts iff the induced subgraph is a bipartite graph with an odd number of vertices. Similarly, we can easily handle any graph property that is a union of a constant number of graph properties, each having a canonical tester and referring to graphs with a different number of vertices. The question, however, is what happens in the general case. We suggest this question as an open problem.

In this errata we point out that the proof presented in [4, Sec. 4] established a weaker version of [4, Thm. 2]. The latter version refers to a more relaxed notion of a canonical tester, which in turn refers to an infinite sequence of graph properties, $\{\Pi^{(i)}\}$. On input (N, ϵ), this tester accepts iff the induced subgraph (determined as in the original definition) has property $\Pi^{(N)}$.

> That is, here we call T canonical if, for some function $s : \mathbb{N} \times [0,1] \to \mathbb{N}$ and an infinite sequence of graph properties $\{\Pi^{(i)}\}$, the tester operates as follows: on input (N, ϵ) and oracle access to any N-vertex graph G, the tester T selects uniformly a set of $s(N, \epsilon)$ vertices (in G), and accepts if and only if the corresponding induced subgraph (of G) has property $\Pi^{(N)}$.

Indeed, the definition in [4] requires that $\Pi^{(i)}$ be independent of i (i.e., $\Pi^{(i)} = \Pi^{(j)}$ for all i, j).

Going through [4, Sec. 4], we observe that all assertions regarding the derived testers hold, except that the operation of these testers is allowed to depend arbitrarily on N. The consequence is, indeed, that the decision of the final tester (i.e., the "canonical" one) depends on whether the induced subgraph has a graph property that is allowed to depend on N.

Alternatively, one may restate and prove [4, Thm. 2] while referring to testing graphs of a fixed number of vertices. We merely need to specify a constant for the main part of the theorem (and as in the original [4, Thm. 2] this constant is determined by the number of repetitions that suffices for decreasing the tester's error probability from 1/3 to 1/6).[1] We stress that the entire proof as presented in [4, Sec. 4] refers to testing graphs with a fixed number of vertices; that is, it starts with a tester for N-vertex graphs, and derives a sequence of testers, each referring to N-vertex graphs.

2 Modifying the proof of Theorem 3

Theorem 3 in [4] is valid, but the proof presented in [4, Sec. 5] is inaccurate in that it ignores the possible effect of N on the decision of arbitrary (one-sided error) testers for graph properties (even of the graph partition type). Here we sketch a set of modifications that suffices for removing the aforementioned inaccuracies.

We observe that the justifications presented in [4, Sec. 5] (refer to assertions that) fall into one of three categories.

1. The first category is of (justifications to) assertions that refer to any graph partition property (and do not refer at all to the testability of these properties). Notable examples are Lemmas 5.1 and 5.5, which indeed remain intact (because the operation of potential testers is irrelevant to them).

2. The second category is of (almost immediate) corollaries to previously proved claims, which indeed remain intact. Notable examples are Corollary 5.4 and Theorem 5.8.

3. This leaves us with Claims 5.2 and 5.3 and Lemmas 5.6 and 5.7, which are discussed below. The problem with their current proofs is that they refer to potential testers of property Π, and furthermore to the operation of these testers on graphs of different number of vertices.

The proof of Claim 5.2 is a typical case. The original proof (of Claim 5.2) refers to a blow-up of the original graph, and presupposes that the *canonical* tester's behavior on the blow-up version G' is identical to its behavior on the original graph G. (Note that this assumption is justified for the original version of the notion of canonical testers, but not for the relaxed notion defined above.) This assumption cannot be justified when the tester's decision depend on the number of vertices. Instead, we may verify that the blow-up version G' satisfies the claim's hypothesis (which refers to the original graph G), and work directly with the blow-up version G'. For example, for every $N' > t^2 \cdot N$, we may consider a blow-up of G by a factor of N'/N (allowing fractional vertices and edges as per the conventions in [4, Sec. 5.0]), and argue directly on this blow-up version rather than on G. The same applies to the (main part of the) proof of Claim 5.3, which is analogous to the proof of Claim 5.2.

The proof of Lemma 5.6 refers to any graph partition problem that satisfies Corollary 5.4, and proceeds while referring to the operation of a canonical tester on sufficiently large graphs. However, the argument first

[1] It turns out that 9 repetitions are necessary and sufficient, and so (in the case of two-sided error testers) the canonical tester needs to sample $18q(N, \epsilon)$ vertices.

selects N to be sufficiently large (in a way depending only on Π and the query complexity of the tester), and proceeds by considering the operation of a canonical tester on various N-vertex graphs. Thus, the foregoing problem does not arise, and the proof may remain intact.

Finally, we turn to the proof of Lemma 5.7. The first part of the proof (which establishes Claim 1) presupposes that the tester behavior remains unchanged when considering a blow-up of the graph. This part can be modified as the proof of Claim 5.2. The second part of the proof of Lemma 5.7 (which establishes Claim 2 and the rest) refers to any graph that satisfies the conclusion of Claim 1, and proceeds without reference to any potential tester. Thus it can remain intact.

Comments regarding Corollary 5.9 and Proposition D.2. The proof of Corollary 5.9 refers to Proposition D.2, which in turn is stated in an inaccurate manner. We first note that Corollary 5.9 has a direct proof (see Footnote 16). Alternatively, Corollary 5.9 can be proven based on the following weaker version of Proposition D.2. We call a tester natural if its query complexity is independent of the size of the graph and, on input (N, ϵ) and access to an N-vertex graph, its decision is based solely on the sequence of oracle answers that it has received (while possibly tossing additional fresh coins, provided that their number is independent of N). Note that the testers of [3] (to which the proof of Corollary 5.9 refers) are natural in this sense. The revised version of Proposition D.2 refers only to *natural testers* (for graph properties that are closed under taking induced subgraphs), and is supported by the current proof (which presumes that the tester T is natural). It suffices to note that applying the "canonization process" to a natural tester yields a tester that accepts an N-vertex graph G iff a random induced subgraph (of size $s(\epsilon)$) has property Π' (i.e., $\Pi' = \Pi^{(N)}$ for all N).[2] Using such a tester, the rest of the proof of [4, Prop. D.2] remains intact.

Acknowledgments

We are grateful to Asaf Shapira for calling our attention to the error in [4], and discussing related issues with us.

References

[1] N. Alon and A. Shapira. A Characterization of the (natural) Graph Properties Testable with One-Sided. In *46th FOCS*, to appear, 2005.

[2] N. Alon and A. Shapira. A Separation Theorem in Property Testing. Unpublished manuscript, 2004.

[3] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *Journal of the ACM*, pages 653–750, July 1998.

[4] O. Goldreich and L. Trevisan. Three theorems regarding testing graph properties. *Random Structures and Algorithms*, Vol. 23 (1), pages 23–57, August 2003.

[2]Note that when deriving such a canonical tester, some of the steps in the (revised) proof of [4, Thm. 2] can be avoided. In particular, when applied to natural tests, [4, Clm. 4.2] is trivial.