A Graduate Course on Algorithms **Spring Semester, 1997/8**

Lecture 7: Coloring 3-Colorable Graphs

Lecturer: Uri Zwick *Scribe: Ayal Zaks*

This lecture deals with the problem of proper vertex colorings of graphs. More specifically, we are interested in coloring a given 3-colorable graph with as few colors as we can.

7.1 Introduction

7.1.1 Hardness Results

It is easy to see that it is NP-Hard to find a 3 coloring for any given 3-colorable graph - this would enable deciding whether a general graph is 3-colorable. If a polynomial-time algorithm existed for 3-coloring a 3-colorable graph, run it on a general graph for as the long as the polynom specifies; if a 3-coloring was obtained by that time - the graph is 3-colorable. If not (the algorithm did not terminate, crashed, produced an improper 3-coloring or a coloring with more than 3 colors) then the graph is not 3-colorable.

Khanna, Linial and Safra [KLS93] proved that finding a 4-coloring of a 3-colorable graph is also NP-Hard. It is not known whether finding a 5-coloring of such graphs is also NP-Hard.

7.1.2 Known Results

In 1983 Wigderson [Wig83] showed how a $O(\sqrt{n})$-coloring can easily be obtained. Blum [Blu94] achieved a $O(n^{\frac{3}{8}})$-coloring. Both proofs are combinatorial. In 1994 Karger, Motwani and Sudan [KMS98] obtained a $O(n^{\frac{1}{4}})$-coloring using semi-definite programming, and in 1996 Blum and Karger [BK97] showed how to obtain a $O(n^{\frac{3}{14}})$-coloring.

7.1.3 Extensions

There are extensions of the above results to coloring k-colorable graphs for $k > 3$. Let Δ denote the maximum degree in the graph. Obviously, the graph can be colored using $\Delta + 1$ colors. The result of [KMS98] shows that $O(\Delta^{\frac{1}{3}})$ colors suffice (and in general, for any $k \geq 3$, $O(\Delta^{1-\frac{2}{k}})$ suffice).

7.2 Coloring with O(sqrt(n)) colors

The key observation is that in a 3-colorable graph, the subgraph spanned by the neighbors $N(v)$ of any vertex v, is 2-colorable. The algorithm works as follows: as long as there is a vertex v with $|N(v)| \geq \sqrt{n}$, color $N(v)$ using two new colors and remove the colored vertices from the graph. After no more than $\sqrt{n}$ steps we will remain with a subgraph of at most $\sqrt{n}$ vertices; color the remaining subgraph with $\sqrt{n}$ new colors. This coloring uses at most $2\sqrt{n} + \sqrt{n} = 3\sqrt{n}$ colors.

A more careful analysis shows that $\sqrt{8n}$ colors suffice.

7.3 Using Semi-Definite Programming

There is a connection between 3-coloring and MAXCUT, for which a semi-definite relaxation proved to be powerful. In MAXCUT we look for a partition of the vertices into two groups, with a maximum number of edges connecting the two groups. In 3-coloring, we look for a partition of the vertices into three color-classes, with a minimum number (zero) of edges inside each class. To approximate MAXCUT, we started by defining a variable for each vertex that could be either +1 or -1, and relaxed this integer constraint by letting each variable be a real vector of norm 1 in R^n. For 3-coloring, we will assign a vector to each vertex, and we would like for vectors of two adjacent vertices to be at an angle of $\frac{2\pi}{3}$ apart. The hope is that the semidefinite program will yield vectors that could be clustered together into a small number of clusters, each of which will represent a color class.

As opposed to MAXCUT, here it is known that the graph is 3-colorable, so the optimum is known - three clusters suffice. However, assigning to each vertex one of 3 vectors such that any pair of adjacent vertices are assigned to different vectors is NP-Hard. The SDP relaxation for 3-coloring is intended only to find a feasible solution, and not to determine what the optimal value is.

Definition *vector 3-coloring* of a graph $G = (V, E)$ is an assignment of a vector for each vertex such that $(i, j) \in E \rightarrow v_i \cdot v_j = -\frac{1}{2}$.

Corollary 7.1 *Any 3-colorable graph is also vector-3-colorable; the vectors can all belong to one plane. Note that the converse is not true.*

Remark. In a similar way, a *vector 2-coloring* can be defined where instead of $-\frac{1}{2}$ appears -1. It is easy to see that a graph is vector-2-colorable iff it is 2-colorable.

Remark. Vector colorings are closely related to orthonormal representations of graphs. A collection of vectors $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal representation of graph $G = (V, E)$ if for $(i, j) \in E \rightarrow v_i \cdot v_j = 0$.

Lemma 7.2 *The neighbors $N(i)$ of any vertex i of a vector-3-colorable graph G, induce a 2-colorable subgraph.*

Proof: Let v_i be the vector that corresponds to vertex i, and let $\{v_{j_1}, v_{j_2}, \ldots\}$ be the vectors that correspond to the neighbors of i. Project each v_j on the hyperplane perpendicular to v_i:

$$v'_j \leftarrow v_j - (v_i \cdot v_j)v_i = v_j + \frac{1}{2}v_i$$

and normalize it: $v''_j \leftarrow \frac{2}{\sqrt{3}}v'_j$. It is easy to see that $v''_{j_1} \cdot v''_{j_2} = -1$ for any pair of adjacent vertices j_1, j_2 which are both adjacent to i. This shows that the subgraph induced on $N(i)$ is vector-2-colorable, which in turn implies that it is 2-colorable, as desired. ■

Definition h alf-coloring of a graph $G = (V, E)$ is a proper assignment of colors to at least half of the vertices of V.

Remark. If a half coloring of G can be found in polynomial time using n^α colors, where $n = |V|$ and α is a constant, then a complete coloring of G can be found (in polynomial time) using $n^\alpha + (\frac{n}{2})^\alpha + \ldots = O(n^\alpha)$ colors.

A randomly chosen hyperplane will separate between two vectors v_i, v_j that belong to adjacent vertices i, j with probability $\frac{2}{3}$. So such a hyperplane can be used to color all the vertices by two colors, where one-third of the edges are expected to turn out monochromatic. We will therefore randomly choose r hyperplanes, that will partition the vectors into at most 2^r parts; this partition will be used to color the vertices with 2^r colors. The probability that v_i, v_j (where $(i,j) \in E$) are colored with the same color is clearly $(\frac{1}{3})^r$. Therefore the expected number of monochromatic edges is $|E| \cdot (\frac{1}{3})^r = \frac{d}{2}n(\frac{1}{3})^r < \frac{n}{4}$ if we take $r \geq \log_3 d + 1$ where d is the average degree in G. Thus with high probability the number of monochromatic edges will be less than $\frac{n}{2}$. By erasing the color of one vertex from each monochromatic edge we will obtain a half-coloring of G. By choosing carefully the set of $\frac{n}{2}$ vertices whose color is erased, we can ensure that the average degree d does not increase. The number of colors used is at most $2^r = O(d^{\log_2 3}) = O(d^{0.631})$.

7.4 Finding Large Independent Sets

An alternative approach to finding half-colorings using as few colors as possible, is to try to find a large independent set, for instance of size at least $\frac{n}{\Delta^{1/3}\sqrt{\ln \Delta}}$. If we succeed, a complete coloring can be obtained by finding such independent sets and coloring each one successively; this will require $O(\Delta^{\frac{1}{3}}\sqrt{\ln \Delta})$ which is the currently best known bound on Δ.

Note. When an independent set is removed from the graph, the average degree may increase. Therefore in the above case we consider the maximum degree Δ instead of the average degree d.

7.4.1 Finding A Large Independent Set

Let r be an arbitrary vector from R^n (not necessarily of norm 1), chosen from an n-dimensional normal distribution, and let c be a constant which will be determined later. Define $U = \{i | r \cdot v_i \geq c\} \subseteq V$, and let $n' = |U|$, and $m' = |\{i, j | i, j \in U, (i,j) \in E\}|$. We will show that

$$E[n' - m'] = \Omega\left(\frac{n}{\Delta^{1/3}\sqrt{\ln \Delta}}\right),$$

and therefore if we remove one vertex from every edge in U, we will obtain an independent set of the desired size.

Since r is picked from an n-dimensional normal distribution and v_i are unit vectors, $r \cdot v_i$ have one-dimensional normal distribution: $N(x) = \int_x^\infty \phi(y)dy$ where $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. We will use the well-known bounds of normal distributions, that for every positive x:

$$\left(\frac{1}{x} - \frac{1}{x^3}\right)\phi(x) \leq N(x) \leq \frac{1}{x}\phi(x). \tag{7.1}$$

Now,

$$\begin{aligned} E[n'] &= n \cdot Pr[v \cdot r \geq c] = n \cdot N(c) \\ E[m'] &= \sum_{(i,j)\in E} Pr[v_i \cdot r \geq c, v_j \cdot r \geq c] \\ & \quad Pr[v_i \cdot r \geq c, v_j \cdot r \geq c] \leq Pr[(v_i + v_j)r \geq 2c]. \end{aligned}$$

Since $||v_i + v_j||^2 = (v_i + v_j)(v_i + v_j) = v_i^2 + v_j^2 + 2v_iv_j = 1 + 1 + 2(-\frac{1}{2}) = 1$, it follows that $v_i + v_j$ has a standard normal distribution, so $Pr[(v_i + v_j) \geq 2c] = N(2c)$. Therefore,

$$E[n' - m'] \geq n \cdot N(c) - |E| \cdot N(2c) \geq n(N(c) - \frac{\Delta}{2} N(2c)).$$

We choose c large enough so that $N(c) \geq \Delta \cdot N(2c)$ as follows. Using (7.1):

$$\frac{N(c)}{N(2c)} \geq \frac{(\frac{1}{c} - \frac{1}{c^3})\mathrm{e}^{-c^2/2}}{\frac{1}{2c}\mathrm{e}^{-2c^2}} \geq 2\left(1 - \frac{1}{c^2}\right)\mathrm{e}^{3c^2/2}.$$

In order to obtain $\mathrm{e}^{3c^2/2} \geq \Delta$ we set $c = \sqrt{\frac{2}{3}\ln\Delta}$. If c turns out to be smaller than $\sqrt{2}$, which happens when $\Delta < \mathrm{e}^3$, we color G naively with $\Delta + 1$ colors. Otherwise we will have $\frac{N(c)}{N(2c)} \geq \Delta$ as desired.

Therefore,

$$E[n' - m'] \geq n(N(c) - \frac{1}{2}N(c)) = \frac{1}{2}N(c) \cdot n \geq \frac{1}{2}\frac{1}{\sqrt{2\pi}}\left(\frac{1}{c} - \frac{1}{c^3}\right)\mathrm{e}^{-c^2/2} = \Omega\left(\frac{n}{\Delta^{1/3}\sqrt{\ln\Delta}}\right).$$

This proves the following theorem.

Theorem 7.3 *A 3-colorable graph can be polynomially colored using $\tilde{O}(\Delta^{1/3})$ colors.*

Combining this result with Wigderson's method, we immediately obtain the following result.

Theorem 7.4 *A 3-colorable graph can be polynomially colored using $\tilde{O}(n^{1/4})$ colors.*

Proof: As long as there exists a vertex v with $|N(v)| > d$ (d will be determined soon), color $N(v)$ with two new colors. If all the vertices have degree $\geq d$, color them with $\tilde{O}(d^{1/3})$ colors using theorem 7.3. The total number of colors used is at most $2\frac{n}{d} + \tilde{c}d^{1/3}$. This sum is minimized when $\frac{n}{d} = d^{1/3} \rightarrow d = n^{3/4}$. Using such a d we obtain a coloring with $\tilde{O}(n^{1/4})$ colors. ∎

7.4.2 Generalized Algorithm

Theorem 7.5 *In order to color a 3-colorable graph with $O(n^\alpha)$ colors, it suffices to be able to find one of the following:*

1. *A 2-colorable set of $n^{1-\alpha}$ vertices.*
2. *A 2-colorable set S having at most $|S| \cdot n^\alpha$ neighbors.*
3. *A pair of vertices u, v having at least $n^{1-2\alpha}$ common neighbors.*

Proof: If case 1 occurs, we color the $n^{1-\alpha}$ vertices with two colors. This can occur at most $\frac{n}{n^{1-\alpha}} = n^\alpha$ times. Notice that case 1 is a special case of 2, and that if case 2 occurs all the time then case 1 holds. Indeed, let S_1 be a 2-colorable set having a neighborhood T_1 of at most $|S_1| \cdot n^\alpha$ vertices. Let S_{i+1}, T_{i+1} be such sets in the graph induced on $V \setminus \{\cup_{j=1}^{i} S_j \cup_{j=1}^{i} T_j\}$ for $i = 1, 2, \ldots, k-1$. Denote $S = \cup_{i=1}^{k} S_i$ and $T = \cup_{i=1}^{k} T_i$, so that $V = S \cup T$, and T is the neighborhood of S. Clearly S is 2-colorable, being a disjoint union of 2-colorable sets, and $|T| \leq |S| \cdot n^{\alpha}$ since $|T_i| \leq |S_i| \cdot n^{\alpha}$ for all i. Therefore $|S| \geq n^{1-\alpha}$ and case 1 holds.

So suppose case 3 holds. Let $S = N(v) \cap N(u)$ and $T = \cup_{s \in S} N(s)$. So $|S| \geq n^{1-2\alpha}$, and $u, v \in T$. If u, v are colored with two different colors (in a proper 3-coloring), then all the vertices in S must receive the third color, and T must be 2-colorable. So we check whether T is 2-colorable; if it is not, we will color u and v with the same color, so for now we collapse them into a single vertex w having $N(v) \cup N(u)$ as its neighbors (this can happen at most n times). If T is 2-colorable, and $|T| \geq n^{1-\alpha}$, then case 1 holds. Otherwise, case 2 holds: S is 2-colorable since $S \subseteq N(v)$, and $|T| < n^{1-\alpha} = n^{1-2\alpha} \cdot n^{\alpha} \leq |S| \cdot n^{\alpha}$. This completes the proof. ∎

We conclude this lecture by showing how to find the desired structures of Theorem 7.5 in polynomial time for $\alpha = 2/9$, to obtain the following result.

Theorem 7.6 *A 3-colorable graph cane be polynomially colored with $O(n^{2/9})$ colors.*

Proof: If there exists a vertex whose degree is smaller than $n^{2/3}$, remove it from the graph. Let V_1 denote the set of all the vertices that were removed. If $|V_1| > n/2$, we will show that case 1 (of Theorem 7.5) holds for the graph G_1 induced on V_1. Every removed vertex is connected to less than $n^{2/3}$ vertices of V_1 that were removed after it. Thus G_1 can be colored with $\tilde{O}((n^{2/3})^{1/3}) = \tilde{O}(n^{2/9})$ [1] colors using Theorem 7.3, which creates a half-coloring of G from which we can obtain an independent set (taking a color class of maximal size) of size $\frac{n/2}{n^{2/9}} = \frac{1}{2}n^{7/9}$ as case 1 requires.

Otherwise, the graph G_2 induced on the remaining vertices $V \setminus V_1$ has at least $n/2$ vertices and minimum degree greater than $n^{2/3}$.

Claim 7.7 *For every vertex v of G_2 there exists a vertex u of G_2 such that $|N(v) \cap N(u)| \geq n^{1/3}$.*

Proof: There are at least $(n^{2/3})^2 = n^{4/3}$ paths of length 2 from v, but less than $n/2$ vertices at distance 2 from v. Therefore there are at least $\frac{n^{4/3}}{n} = n^{1/3}$ such paths that reach the same vertex (pigeon-hole principle). ∎

Such a pair of vertices u, v can be found in polynomial time. Denote $S = N(v) \cap N(u)$ and $T = \cup_{s \in S} N(s)$. So $|S| \geq n^{1/3}$ and $u, v \in T$. If T is not 2-colorable, collapse u and v into one vertex. Let S' be an arbitrary subset of S of size $2n^{1/9}$. If there exist $x, y \in S'$ with $\frac{1}{4}n^{5/9}$ common neighbors, case 3 applies. Otherwise

$$|T| = n^{2/3}|S'| - \frac{n^{5/9}}{4}|S'|^2 = 2n^{1/9}(n^{2/3} - \frac{n^{6/9}}{2}) \geq n^{7/9}$$

and T serves as the set required by case 1. ∎

[1] We are using a version in which the degrees are directed! The maximal degree of G_1 can be greater than $n^{2/3}$.

Bibliography

[BK97] A. Blum and D. Karger. An $\tilde{O}(n^{3/14})$-coloring algorithm for 3-colorable graphs. *Information Processing Letters*, 61:49–53, 1997.

[Blu94] A. Blum. New approximation algorithms for graph coloring. *Journal of the ACM*, 41:470–516, 1994.

[KLS93] S. Khanna, N. Linial, and S. Safra. On the hardness of approximating the chromatic number. In *Proceedings of the 2nd Israel Symposium on Theory and Computing Systems, Natanya, Israel*, pages 250–260, 1993.

[KMS98] D. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidefinite programming. *Journal of the ACM*, 45:246–265, 1998.

[Wig83] A. Wigderson. Improving the performance guarantee for approximate graph coloring. *Journal of the ACM*, 30:729–735, 1983.