COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA **30** (2) (2010) 217–224
DOI: 10.1007/s00493-010-2483-5

A FAST ALGORITHM FOR EQUITABLE COLORING

HENRY A. KIERSTEAD*, ALEXANDR V. KOSTOCHKA†,
MARCELO MYDLARZ, ENDRE SZEMERÉDI‡

Received February 5, 2008

A proper vertex coloring of a graph is *equitable* if the sizes of color classes differ by at most one. The celebrated Hajnal–Szemerédi Theorem states: For every positive integer r, every graph with maximum degree at most r has an equitable coloring with $r+1$ colors. We show that this coloring can be obtained in $O(rn^2)$ time, where n is the number of vertices.

1. Introduction

An *equitable k-coloring* of a graph G is a proper k-coloring, in which any two color classes differ in size by at most one. Equitable colorings naturally arise in some scheduling, partitioning, and load balancing problems [2,12, 13]. In 1970 Hajnal and Szemerédi [3] proved the following theorem, which had been conjectured by Erdős.

Theorem 1. *Every graph with maximum degree at most r has an equitable $(r+1)$-coloring.*

Alon and Füredi [1] and Janson and Ruciński [4] used this theorem to derive deviation bounds for sums of random variables that exhibit limited dependence. Rödl and Ruciński [11] used it to give a simpler proof of the

Mathematics Subject Classification (2000): 05C15, 05C85

* Research of this author is supported in part by the NSA grant MDA 904-03-1-0007.

† Research of this author is supported in part by NSF grant DMS-06-50784 and by grant 06-01-00694 of the Russian Foundation for Basic Research.

‡ Research of this author is supported in part by NSF grant DMS-0902241.

Blow-Up Lemma [9]. The proof of Theorem 1 was surprisingly long and complicated, and did not yield a polynomial time algorithm. Recently, several new results related to Theorem 1 have appeared. Mydlarz and Szemerédi [10] and independently Kierstead and Kostochka [5] found polynomial time algorithms for such coloring. Apart from this, Kierstead and Kostochka [6] proved the following Ore-type theorem strengthening a conjecture in [8]:

Theorem 2. *If for every edge xy of a graph G, the sum $d(x)+d(y)$ of the degrees of its ends does not exceed $2r+1$, then G has an equitable $(r+1)$-coloring.*

They have also obtained partial results for equitable versions of Brooks' Theorem [7] and equitable list coloring. This work has led to new insights into the methods of both groups (Mydlarz–Szemerédi and Kierstead–Kostochka). The purpose of this paper is to present a faster algorithm than either group had obtained previously. This requires at least two new methods and the recombination of several old ideas. In particular, Lemma 3 and the use of the maximal independent set in Case 2 of its proof are new. An effective method for determining the case is crucial for the algorithm. Searching for an algorithm also led us to an easier proof. In Section 2 we present the proof of Theorem 1 on which our algorithm is based. The discharging arguments of previous proofs are replaced by more direct counting arguments. In Section 3 we show that our new proof yields a simple $O(rn^2)$ time algorithm.

Most of our notation is standard; possible exceptions include the following. For a vertex y and a set of vertices X, $N_X(y) := N(y) \cap X$ and $d_X(y) := |N_X(y)|$. The set of edges with one end in X and the other end in Y is denoted by $E(X,Y)$; $E(y)$ is the set of edges incident with y. Let $|G|$ be *the order* (the number of vertices) and $\|G\|$ be *the size* (the number of edges) of a graph G. We write $A-x$ for $A \setminus \{x\}$.

2. Proof of Theorem 1

Let G be a graph with $\Delta(G) \leq r$. We may assume that $|G|$ is divisible by $r+1$: If $|G| = s(r+1)-p$, where $1 \leq p \leq r$, then set $G' := G + K_p$. Then $|G'|$ is divisible by $r+1$ and $\Delta(G') \leq r$. Moreover, the restriction of any equitable $(r+1)$-coloring of G' to G is an equitable $(r+1)$-coloring of G. So we may assume $|G| = (r+1)s$.

We argue by induction on $\|G\|$. The base step $\|G\| = 0$ is trivial, so consider the induction step. Let u be a non-isolated vertex. Let G' be obtained from G by deleting all edges incident with u. By the induction hypothesis,

G' has an equitable $(r+1)$-coloring f. We are done unless some color class V of f contains an edge uv. In this case, since $\Delta(G)<r+1$, some class W contains no neighbors of u. Moving u to W yields an $(r+1)$-coloring of G with all classes of size s, except for one *small* class $V^-:=V-u$ of size $s-1$ and one *large* class $V^+:=W+u$ of size $s+1$. Such a coloring is called *nearly equitable.*

Given a nearly equitable $(r+1)$-coloring, define an auxiliary digraph $\mathcal{H}=\mathcal{H}(G)$, whose vertices are the color classes, so that UW is a directed edge if and only if some vertex $y\in U$ has no neighbors in W. In this case we say that y *witnesses* UW. Let $\mathcal{A}$ be the set of classes that can reach V^- in $\mathcal{H}$, $\mathcal{B}$ be the set of classes not in $\mathcal{A}$ and $\mathcal{B}'$ be the set of classes that V^+ can reach in $\mathcal{H}[\mathcal{B}]$. Call a class $W\in\mathcal{A}$ *terminal*, if every $U\in\mathcal{A}-W$ can reach V^- in $\mathcal{H}-W$, and let $\mathcal{A}'$ be the set of terminal classes. Set $a:=|\mathcal{A}|$, $a':=|\mathcal{A}'|$, $b:=|\mathcal{B}|$, $b':=|\mathcal{B}'|$, $A:=\bigcup\mathcal{A}$, $A':=\bigcup\mathcal{A}'$, $B:=\bigcup\mathcal{B}$ and $B':=\bigcup\mathcal{B}'$. Then $r+1=a+b$. Note that

$$(*)\qquad \begin{array}{lllll} \text{(i)} & \forall W\in\mathcal{A} & \forall y\in B & d_W(y)\geq 1 & \text{and}\\ \text{(ii)} & \forall W\in\mathcal{B}\setminus\mathcal{B}' & \forall z\in B' & d_W(z)\geq 1. & \end{array}$$

The following lemma completes the induction step.

Lemma 3. *If G is a graph with a nearly equitable $(r+1)$-coloring such that every vertex $v\in A'\cup B$ has degree at most r, then G has an equitable $(r+1)$-coloring.*

Proof. Argue by induction on $b=|B|$. If $V^+\in\mathcal{A}$, which includes the base step $b=0$, then there exists a V^+,V^- path $\mathcal{P}=V_1,\dots,V_k$ in $\mathcal{H}$. Moving a witness y_j of V_jV_{j+1} to V_{j+1} for each j yields an equitable $(r+1)$-coloring of G. So we may assume $V^+\in\mathcal{B}$ and $b\geq 1$. Then $|A|=as-1$ and $|B|=bs+1$. If $V^-\in\mathcal{A}'$ then no other class can be in $\mathcal{A}$. So in this case $a=1, b=r$ and $d_B(x)\leq r$ for all $x\in A$. Using (i) of $(*)$, $rs+1\leq|E(V^-,B)|\leq r(s-1)$, a contradiction. So $a'<a$.

An edge wz is *solo* if $w\in W\in\mathcal{A}'$, $z\in B$ and $N_W(z)=\{w\}$. Ends of solo edges are *solo* vertices and *solo neighbors* of each other. If wz is a solo edge and $w\in W\in\mathcal{A}'$ witnesses an edge WX of $\mathcal{H}[\mathcal{A}]$, then moving w to X and z to W yields nearly equitable (or equitable) colorings of $G_1:=G[A+z]$ and $G_2:=G[B-z]$. Since W is terminal, V^- is reachable from $X+w$ in $\mathcal{H}(G_1)-(W+z-w)$. Thus G_1 has an equitable a-coloring. By (i) of $(*)$, $\Delta(G_2)<b$. Since $b(G_2)<b$, the induction hypothesis yields an equitable b-coloring of G_2, and thus an equitable $(r+1)$-coloring of G. Thus we may assume no solo vertex witnesses an edge of $\mathcal{H}[\mathcal{A}]$, and so:

$$(**)\qquad \forall w\in A'(wz \text{ solo} \implies (d_A(w)\geq a-1\ \wedge\ d_{B-z}(w)\leq b-1)).$$

Order $\mathcal{A}$ as $X_0 := V^-, X_1, \dots, X_{a-1}$ so that each X_{i+1} has a previous out-neighbor. If X_l is the last nonterminal class, then some $X_j \in \mathcal{A}'$ with $l < j$ cannot reach V^- in $\mathcal{H} - X_l$. So X_j has no out-neighbor before X_l, and thus $d^+_{\mathcal{A}}(X_j) < a - l$. If $a - l \le b$, then (1) $d^+_{\mathcal{A}}(X_j) < b$; otherwise the last b classes (including X_j) are all terminal, implying (2) $a' \ge b$.

Case 1. $d^+_{\mathcal{A}}(W) < b$ for some $W \in \mathcal{A}'$ (which includes (1)). Every $w \in W$ has neighbors in all classes of A except W and at most $b-1$ other classes. Thus $d_A(w) \ge a - b$ and $d_B(w) < 2b$. Let S be the set of solo vertices in W, and $D := W \setminus S$. If $v \in B - N_B(S)$ then v has no solo neighbor in W, and so has at least two neighbors in D. Thus $2b|D| > 2|B - N_B(S)|$. Using $|S| + |D| = s$ and $r|S| \ge |E(S,A)| + |N_B(S)|$,

$$\begin{aligned} bs + (a-1)|S| = b|D| + r|S| &> |B - N_B(S)| + |E(S,A)| + |N_B(S)| \\ &> bs + |E(S,A)|. \end{aligned}$$

Hence $(a-1)|S| > |E(S,A)|$. So $d_A(w) \le a - 2$, for some (solo) $w \in S$, contradicting $(**)$.

Case 2. $a' \ge b$. For $z \in B'$, let $\sigma(z)$ be the number of solo neighbors of z. Using $(*)$,

$$r \ge d_A(z) + d_B(z) \ge a + a' - \sigma(z) + b - b' + d_{B'}(z) \ge r + 1 + d_{B'}(z) + a' - b' - \sigma(z).$$

So $\sigma(z) \ge a' - b' + d_{B'}(z) + 1$. Let I be a maximal independent set in $G[B']$ with $V^+ \subseteq I$. Then $\sum_{z \in I}(d_{B'}(z) + 1) \ge |B'| = b's + 1$. Since $a' \ge b$,

$$\sum_{z \in I} \sigma(z) \ge \sum_{z \in I} (a' - b' + d_{B'}(z) + 1) \ge (s+1)(a' - b') + b's + 1 > a's = |A'|.$$

So some solo $w \in W \in \mathcal{A}'$ has two solo neighbors z_1 and z_2 in the independent set I. Switch witnesses along a W, V^--path to obtain an equitable coloring of $G_1 := G[A - W + w']$, where w' witnesses the first edge of this path. Since the class Z of z_1 is reachable from V^+, we can equitably b-color $G[B - z_1]$. Let Z' be the new class of z_2. By $(**)$, we can move w to a new class U of $B - z_1$ and then move z_1 to W to obtain a new nearly equitable $(b+1)$-coloring of $G_2 := G - G_1$ with small class $W^* := W - w' - w + z_1$. Since z_2 witnesses that Z' is adjacent to W^*, we have $b(G_2) < b$ and W^* is not terminal. By (i) of $(*)$ and $(**)$, every vertex of $G_2 - W^*$ has at most b neighbors in G_2. So we are done by the induction hypothesis. ∎

3. A fast algorithm

In this section we present an algorithm for finding an equitable $(r+1)$-coloring of any graph G on n vertices with maximum degree at most r and analyze its running time. For this analysis we assume G is received as an $n\times r$ array L, where $L(v,i)$ is the i-th neighbor of v, if it exists, and 0 otherwise. We also assume array entries can be read and written in one step. We will never write a number larger than $O(n)$.

Theorem 4. *Every graph on n vertices with maximum degree at most r can be equitably $(r+1)$-colored in $O(rn^2)$ steps.*

Proof. We begin by setting up a data structure. Then we give an algorithm based on the proof of Theorem 1 that uses a procedure derived from the proof of Lemma 3. As before, we may assume that $n=(r+1)s$.

3.1. Data structure

We order the vertices of G in an arbitrary way: $u_1,\dots,u_n$. Let G_i denote the spanning subgraph of G whose edges are those of G incident with at least one of $u_1,\dots,u_i$. In particular, G_0 has no edges and $G_{n-1}=G$.

We shall maintain global arrays L',F,C,H,N such that:

1. L' is an $n\times r$-array representing the current subgraph G_i in the same way that L is used to represent G.
2. F is an n-array, where $F(v)$ is the color of v in a coloring of G_i.
3. C is an $(r+1)$-array, where $C(X)$ is a list of vertices in the color class X of the coloring F; we can add vertices to $C(X)$ in $O(1)$ steps and remove vertices in $O(s)$ steps.
4. H is an $(r+1)\times(r+1)$-array, where $H(X,Y)$ is the number of witnesses of XY (so $XY\in E(\mathcal{H})$ iff $H(X,Y)>0$).
5. N is an $n\times(r+1)$-array N, where $N(v,X)$ is the number of neighbors of v in color class X.

Many times during the algorithm the current coloring will be locally modified by switching a vertex v from class X to class Y, which requires updating C,F,H, and N. This takes $O(r+s)$ steps: First set $F(v):=Y$, $H(X,Y):=H(X,Y)–1$ and $H(Y,X):=H(Y,X)+1$. Next, for each neighbor y of x, do: $N(y,X):=N(y,X)-1$, $N(y,Y):=N(y,Y)+1$; if $N(u,X)=0$ then $H(F(u),X):=H(F(u),X)+1$; if $N(u,Y)=1$ then $H(F(u),Y):=H(F(u),Y)-1$. Finally, add v to $C(Y)$ and remove v from $C(X)$.

3.2. Algorithm

The algorithm receives the graph G represented by the array L. First it builds the global data structure for the empty graph G_0. Then it iteratively refines the coloring and the data structure using a procedure $\mathcal{P}$ described in the next subsection which takes $O(r^2 s)$ steps. After $n-1$ iterations it returns an equitable $(r+1)$-coloring of G using in total $O(r^3 s^2)$ steps.

Initialization. As in the proof of Theorem 1, by possibly adding some artificial vertices, we may assume that $n=rs$. The algorithm starts with all entries in all arrays, except the input array L, set to zero. Thus L' represents the graph G_0 with no edges. Divide the vertices arbitrarily (there are no edges yet) into color classes of size s. Then update F,C,H,N to reflect this equitable coloring in $O(r^2 s)$ steps.

Iteration i ($i=1,\dots,n-1$). Add the edges in $E(u_i)$ to G_{i-1} to form G_i and update L', H, and N. This takes $O(r)$ steps. If u_i still has no neighbor in its class then go to Iteration $i+1$. Otherwise, switch u_i to a class in which it has no neighbors and update. This takes $O(r+s)$ steps and leaves us with a nearly equitable $(r+1)$-coloring of G_i and an updated data structure. Now run Procedure $\mathcal{P}$ on G_i. We will show that this yields an equitable coloring of G_i and an updated data structure in $O(r^2 s)$ steps. Go to Iteration $i+1$.

Then G will be colored equitably at the end of Iteration $n-1$.

3.3. Procedure $\mathcal{P}$

The procedure receives the global data structure for a nearly equitable $(r+1)$-coloring of a graph G_i. It recursively calculates an equitable $(r+1)$-coloring of G_i and updates the data structure in $O(r^2 s)$ steps.

Initialization. Use breadth-first search to calculate $\mathcal{A}$, $\mathcal{T}$ and R, where $\mathcal{T}$ is an in-branching of $\mathcal{A}$ in $\mathcal{H}$ rooted at V^- and $R(j)$ is the j-th class added to $\mathcal{A}$ by the search. Using the array H this takes $O(ar)$ steps. If $V^+ \in \mathcal{A}$ then we find an equitable $(r+1)$-coloring of G_i by shifting at most r witnesses along the V^+,V^--path in $\mathcal{H}$ determined by $\mathcal{T}$. Using C and N to find the witnesses, this takes $O(r(r+s)) \le O(r^2 s)$ steps. So suppose $V^+ \notin A$.

Decision. Determine which of Case 1 or Case 2 applies as follows. Let $W_1 := R(a)$ be the last class of $\mathcal{A}$. Search for a vertex $w \in W_1$ such that w is a witness for an edge $W_1 X$ in $\mathcal{H}[\mathcal{A}]$ and has a solo neighbor $y \in B$. Using the data structure this takes $O(rs)$ steps: Use $C(W_1)$ to find the s vertices v; check each of the r classes $U \neq W_1$ to see if (i) $U \in \mathcal{A}$ and $N(v,U)=0$ or

(ii) $U \in \mathcal{B}$ and $N(v,U)=1$. If some X passes test (i) and some X' passes test (ii) then accept $w:=v$. In this case, search $X' \cap N(w)$ to find a vertex y such that $N(y,W_1)=1$. If the search is successful then enter Case 1 with class $W=W_1$, witness w and its solo neighbor y. Otherwise the first terminal set has been found. After iterating this process at most b times, Case 1 is entered after at most a iterations or a b-subset $\mathcal{A}_0 \subseteq \mathcal{A}'$ of terminal sets is constructed after at most $b \le a$ iterations and Case 2 is entered. Thus Case 1 is entered after $O(ars)$ steps or Case 2 is entered after $O(brs)$ steps.

Case 1. Finish by shifting at most a vertices in $O(a(r+s))$ steps, for a total of $O(ars)$ steps. Now recursively apply the procedure to $G_i[B-y]$. By induction this takes $O(brs)$ steps. Thus the total number of steps for the procedure is $O(r^2s)$.

Case 2. First calculate $\mathcal{B}'$ using breadth-first search in $O(br)$ steps. Next begin constructing the maximal independent set I using First-Fit. Each time an element z is added to I, mark its solo neighbors in $A_0 := \bigcup \mathcal{A}_0$ with z. When a vertex $w \in W \in A_0$ has been marked twice, say by $z_1 \in Z$ and z_2, end the construction of I. This requires $O(brs)$ steps. Next shift vertices along a W,V^--path and a V^+,Z-path and move z_1 into W. Move w into some class in $\mathcal{B}$ that does not contain its neighbors. This takes $O(r^2s)$ steps, for a total of $O(r^2s)$ steps. Finally, recursively apply the procedure to $G_i[B \cup W^*]$, where $W^* = W - w' - w + z_1$. This takes $O(brs)$ steps, which is satisfactory since $b \le \frac{1}{2}r$. ∎

4. Open question

We have not been able to answer the following conjecture concerning an algorithmic version of Theorem 2.

Conjecture 5. There exists a polynomial time algorithm that will equitably color any graph G with $r+1$ colors, provided that $d(x)+d(y) \le 2r+1$ for all $xy \in E(G)$.

Acknowledgment. We thank Goran Konjevod and Douglas West for reading the text and making useful comments.

References

[1] N. Alon and Z. Füredi: Spanning subgraphs of random graphs, *Graphs and Combinatorics* **8** (1992), 91–94.

[2] J. Blazewicz, K. Ecker, E. Pesch, G. Schmidt and J. Weglarz: *Scheduling computer and manufacturing processes, 2nd ed.*, Springer, Berlin, 485 p. (2001).

[3] A. Hajnal and E. Szemerédi: Proof of a conjecture of P. Erdős, in: *Combinatorial Theory and its Application (P. Erdős, A. Rényi and V. T. Sós, eds.)*, pp. 601–623, North-Holland, London, 1970.
[4] S. Janson and A. Ruciński: The infamous upper tail, *Random Structures and Algorithms* **20** (2002), 317–342.
[5] H. A. Kierstead and A. V. Kostochka: A short proof of the Hajnal–Szemerédi Theorem on equitable coloring, *Combinatorics, Probability and Computing* **17** (2008), 265–270.
[6] H. A. Kierstead and A. V. Kostochka: An Ore-type theorem on equitable coloring, *J. Combinatorial Theory Series B* **98** (2008), 226–234.
[7] H. A. Kierstead and A. V. Kostochka: Ore-type versions of Brooks' theorem, *J. Combinatorial Theory Series B* **99(2)** (2009), 298–305.
[8] A. V. Kostochka and G. Yu: Ore-type graph packing problems, *Combinatorics, Probability and Computing* **16** (2007), 167–169.
[9] J. Komlós, G. Sárközy and E. Szemerédi: Blow-Up Lemma, *Combinatorica* **17(1)** (1997), 109–123.
[10] M. Mydlarz and E. Szemerédi: Algorithmic Brooks' Theorem, manuscript.
[11] V. Rödl and A. Ruciński: Perfect matchings in ϵ-regular graphs and the Blow-Up Lemma, *Combinatorica* **19(3)** (1999), 437–452.
[12] B. F. Smith, P. E. Bjorstad and W. D. Gropp: *Domain decomposition; Parallel multilevel methods for elliptic partial differential equations*; Cambridge University Press, Cambridge, 224 p. (1996).
[13] A. Tucker: Perfect graphs and an application to optimizing municipal services, *SIAM Review* **15** (1973), 585–590.

Henry A. Kierstead

School of Mathematical and Statistical Sciences
Arizona State University
P.O. Box 871804
Tempe, AZ 85287
USA
kierstead@asu.edu

Alexandr V. Kostochka

Department of Mathematics
University of Illinois
1409 W. Green St.
Urbana, IL 61801
USA
and
Sobolev Institute of Mathematics
Novosibirsk, 630090
Russia
kostochk@uiuc.edu

Marcelo Mydlarz

Yahoo! Research Latin America
Blanco Encalada 2120, 4h floor
Santiago
Chile
marcem@cs.rutgers.edu

Endre Szemerédi

Department of Computer Science
Rutgers University
110 Frelinghuysen Rd.
Piscataway, NJ 08854-8019
USA
szemered@cs.rutgers.edu