On approximating the b-chromatic number

Sylvie Corteel\textsuperscript{a}, Mario Valencia-Pabon\textsuperscript{b}, Juan-Carlos Vera\textsuperscript{c}

\textsuperscript{a}CNRS, PRiSM, Université de Versailles, 45 Av. des Etats Unis, 78035 Versailles, France
\textsuperscript{b}Departamento de Matemáticas, Universidad de los Andes, Cra. 1 No. 18A - 70, Bogotá, Colombia
\textsuperscript{c}Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA 15213, USA

Received 9 July 2003; received in revised form 4 August 2004; accepted 21 September 2004

Abstract

We consider the problem of approximating the b-chromatic number of a graph. We show that there is no constant \(\varepsilon > 0\) for which this problem can be approximated within a factor of \(120/133 - \varepsilon\) in polynomial time, unless \(P = NP\). This is the first hardness result for approximating the b-chromatic number.

\(© 2004\) Elsevier B.V. All rights reserved.

Keywords: Combinatorial problems; Approximation algorithms; Graph coloring

1. Introduction

We consider finite undirected graphs without loops or multiple edges. A coloring (i.e. proper coloring) of a graph \(G = (V, E)\) is an assignment of colors to the vertices of \(G\), such that any two adjacent vertices have different colors. A coloring is called a b-coloring, if for each color \(i\) there exists a vertex \(x_i\) of color \(i\) such that for every color \(j \neq i\), there exists a vertex \(y_j\) of color \(j\) adjacent to \(x_i\) (such a vertex \(x_i\) is called a dominating vertex for the color class \(i\)). The b-chromatic number \(\varphi(G)\) of a graph \(G\) is the largest number \(k\) such that \(G\) has a b-coloring with \(k\) colors. The b-chromatic number of a graph was introduced by Irving and Manlove [3] when considering minimal proper colorings with respect to a partial order defined on the set of all partitions of the vertices of a graph. They proved that determining \(\varphi(G)\) is NP-hard for general graphs, but polynomial-time solvable for trees. Recently, Kratochvil et al. [6] have shown that determining \(\varphi(G)\) is NP-hard even for bipartite graphs. Some bounds for the b-chromatic number of a graph are given in [3,5].

In this paper we prove that there is no constant \(\varepsilon > 0\) for which this problem can be approximated within a factor of \(120/113 - \varepsilon\) in polynomial time, unless \(P = NP\). No hardness of approximation was previously known for this problem.

The organization of the paper is as follows. In Section 2 we give the preliminaries. In Section 3 we present the hardness of approximation result. We end in Section 4 with some concluding remarks.
2. Preliminaries

Let $P$ be a maximization problem and let $x \geq 1$. For an instance $x$ of $P$ let $OPT(x)$ be the optimal value. An $x$-approximation algorithm for $P$ is a polynomial-time algorithm $\mathcal{A}$ such that on each input instance $x$ of $P$ it outputs a number $\mathcal{A}(x)$ such that $OPT(x)/x \leq \mathcal{A}(x) \leq OPT(x)$.

To show the hardness of approximating the b-chromatic number we relate it to the hardness of approximating the optimization version of the $k$-ESAT problem. Let $k$ be an integer greater than 1.

**k-ESAT problem**

*Instance:* A set $X = \{x_1, x_2, \ldots, x_n\}$ of boolean variables, a collection $C = \{c_1, c_2, \ldots, c_p\}$ of disjunctive clauses with exactly $k$ different literals, where at least one literal is a variable or a negated variable in $X$.

*Question:* Does there exist a truth assignment for the variables in $X$ such that each clause in $C$ is satisfied?

The decision version of the $k$-ESAT problem is NP-complete for $k \geq 3$ [1]. Johnson showed in [4] the following result.

**Theorem 1** (Johnson [4, Theorem 3]). Let $(X, C)$ be an instance of the $k$-ESAT problem. Then, there is a deterministic polynomial-time algorithm that finds a truth assignment for variables in $X$ which satisfies at least $|C|(1 - 1/2^k)$ clauses in $C$.

The MAX $k$-ESAT problem is the optimization version of the $k$-ESAT problem in which, given an instance of $k$-ESAT, the goal consists of finding the maximum number of clauses that can be satisfied simultaneously by any truth assignment of the boolean variables. The MAX $k$-ESAT problem is NP-hard [1].

Note that in the case $k = 3$, Theorem 1 gives an $8/7$-approximation algorithm for the MAX 3-ESAT problem. Moreover, Hästad showed in [2] the following inapproximability result for the MAX 3-ESAT problem.

**Theorem 2** (Hästad [2, Theorem 6.1]). The MAX 3-ESAT problem is not approximable within $8/7 - \varepsilon$ for any $\varepsilon > 0$, unless $P = NP$.

In the following section, we use Theorem 2 restricted to a special kind of instances in order to obtain an inapproximability result for the b-chromatic number problem of a graph.

**Definition 1.** We say that an instance $(X, C)$ of MAX 3-ESAT is non-trivial if $|C| > 4$, and for all $x \in X$

- There is no $c \in C$ such that $x, \overline{x} \in c$.
- There are $c, d \in C$ such that $x \in c$ and $\overline{x} \in d$.

We now show that Theorem 2 holds when restricted to non-trivial instances of MAX 3-ESAT.

**Corollary 1.** The MAX 3-ESAT problem is not approximable within $8/7 - \varepsilon$ for any $\varepsilon > 0$, even when restricted to non-trivial instances.

**Proof.** We present a proof by contradiction. Assume that there is an $(8/7 - \varepsilon)$-approximation algorithm running in polynomial time $p(|X| + |C|)$ for non-trivial instances $(X, C)$ of the MAX 3-ESAT problem, for some $0 < \varepsilon \leq 1/7$. We prove that there is an $(8/7 - \varepsilon)$-approximation algorithm for the MAX 3-ESAT problem. This contradicts Theorem 2.

We prove this by induction on $|X| + |C|$. The base case is trivial. Now, let $k > 1$ and assume that the statement holds for all instances $(X, C)$ such that $|X| + |C| < k$, and let $(X, C)$ be an instance of MAX 3-ESAT such that $|X| + |C| = k$. If the instance is non-trivial, the statement follows from our initial assumption. If not we have three possible cases:

- There is $x \in X$ such that there is $c \in C$ with $x, \overline{x} \in c$. Let $C' = C \setminus \{c\}$. By induction hypothesis applied to $(X, C')$, we can get, in polynomial time, a truth assignment for the variables in $X$ that satisfies at least $|C'|/(8/7 - \varepsilon)$ clauses in $C'$. This assignment also satisfies $c$ and therefore satisfies at least
  
  $$\frac{|C'|}{8/7 - \varepsilon} + 1 \geq \frac{|C|}{8/7 - \varepsilon}$$

  clauses of $C$.

- There is $x \in X$ such that no clause $c \in C$ contains $x$. Let $X' = X \setminus \{x\}$ and $C' = C \setminus \{c \in C : x \in c\}$. By induction hypothesis we can get, in polynomial time, a truth assignment for the variables in $X'$ that satisfies at least $|C'|/(8/7 - \varepsilon)$ clauses in $C'$. Now
Fig. 1. Partial construction of $G$ from $(X, C)$, where the clause $c_i \in C$ contains the literals $x_1$ and $x_n$.

we assign the value True to $x$, and all clauses in $C$ containing it are satisfied. Therefore we have a truth assignment satisfying at least
\[
\frac{|C'|}{8/7 - \varepsilon} + |C \setminus C'| \geq \frac{|C|}{8/7 - \varepsilon}
\]
classes.

- There is $x \in X$ such that no clause $c \in C$ contains $x$. This case is analogous to the previous one.

Therefore, there is a $(8/7 - \varepsilon)$-approximation algorithm for the MAX 3-ESAT problem running in polynomial time $O(k^2) p(k)$, where the $O(k^2)$ term represents the time needed to find the desired $x$ and construct $X'$ and $C'$ and is certainly not the best possible. □

3. Hardness of approximation

In this section we prove the hardness result for approximating the b-chromatic number problem of a graph.

Let $(X, C)$ be an instance of the 3-ESAT problem. We define $G(X, C) = (V, E)$ to be the graph constructed as follows:

Let $X = \{x_1, x_2, \ldots, x_n\}$ be the set of boolean variables, and let $C = \{c_1, c_2, \ldots, c_p\}$ be the collection of disjunctive clauses, with $c_i = \{l_{i,1}, l_{i,2}, l_{i,3}\}$ for $i = 1, 2, \ldots, p$, where $l_{i,j} = x_k$ or $l_{i,j} = \overline{x_k}$ for some $1 \leq k \leq n$.

Let
\[
V = \{v\} \cup \{z_i : 1 \leq i \leq p - 1\} \cup \{w_j : 1 \leq j \leq 2p\} \\
\cup \{y_i : 1 \leq i \leq p\} \cup \{x_{i,j}, \overline{x_{i,j}} : 1 \leq i \leq n, 1 \leq j \leq p\},
\]
and let
\[
E = \{(z_i, w_j) : 1 \leq i \leq p - 1, 1 \leq j \neq i \leq 2p\} \\
\cup \{(v, z_i) : 1 \leq i \leq p - 1\} \cup \{(v, y_i) : 1 \leq i \leq p\} \\
\cup \{(y_i, y_{i'}) : 1 \leq i < j \leq p\} \\
\cup \{(x_{i,j}, \overline{x_{i,j}}) : 1 \leq i \leq n, 1 \leq j, k \leq p\} \\
\cup \{(y_i, x_{j,k}) : 1 \leq i \leq p, 1 \leq j \leq n, 1 \leq k \leq p, x_j \in c_i\} \\
\cup \{(y_i, \overline{x_{j,k}}) : 1 \leq i \leq p, 1 \leq j \leq n, 1 \leq k \leq p, \overline{x_j} \in c_i\}.
\]

Notice that $|V| = 2np + 4p$.

The resulting graph $G(X, C) = (V, E)$ is shown in Fig. 1.

**Theorem 3.** Let $(X, C)$ be a non-trivial instance of the 3-ESAT problem, where $|X| = n$ and $|C| = p$. Then, $\varphi(G(X, C)) = p + t$ where $t$ is the maximum number of clauses that can be satisfied in $C$. 
The proof of Theorem 3 requires Propositions 1 and 2 below.

**Proposition 1.** Let \( (X, C) \) be a non-trivial instance of the 3-ESAT problem, where \(|X| = n\) and \(|C| = p\). Let \( t \) be the maximum number of clauses that can be satisfied in \( C \). Then there is a b-coloring of \( G(X, C) \) with \( p + t \) colors.

**Proof.** Fix a truth assignment of the variables that satisfies exactly \( t \) clauses. W.l.o.g. assume that the clauses satisfied in \( C \) are \( c_1, c_2, \ldots, c_t \).

Color the vertices of \( G(X, C) \) with \( p + t \) colors as follows:

- for \( 1 \leq i \leq p - 1 \), assign color \( i \) to vertex \( z_i \).
- assign color \( p \) to vertex \( v \).
- for \( 1 \leq i \leq t \), assign color \( p + i \) to vertex \( y_i \).

The previous vertices will be the dominating vertices of each one of the \( p + t \) color classes.

For \( 1 \leq j \leq p + t \), assign color \( j \) to vertex \( w_j \), and for \( 1 \leq j \leq p - t \), assign color \( p + t \) to vertex \( w_{p+t+j} \). In this way, the vertex \( z_j \) is dominating for the color class \( i \).

Vertex \( v \) is already a dominating vertex for the color class \( p \).

For \( t + 1 \leq i \leq n \), assign to vertex \( y_i \) the color \( i - t \).

For every \( 1 \leq i \leq n \), do the following. If \( x_j \) is true, choose \( 1 \leq s \leq p \) such that \( x_j \in c_s \). Notice that \( c_s \) is satisfied and therefore \( s \leq t \). Assign to each \( x_{i,j} \) color \( j \) and to \( x_{i,j} \) color \( p + s \), for \( 1 \leq j \leq p \). If \( x_j \) is false then \( \overline{x_j} \) is true, and proceed in the analogous way.

Now, we just need to check that the coloring is proper and that for \( 1 \leq i \leq t \), \( y_i \) is a dominating vertex for its color class.

The coloring is not proper only if there are \( 1 \leq i \leq p, 1 \leq j \leq n \) and \( 1 \leq k \leq p \) such that there is an edge between \( y_i \) and \( l_{j,k} \), where \( l_{j,k} = x_{j,k} \) or \( l_{j,k} = \overline{x_{j,k}} \), with \( y_i \) and \( l_{j,k} \) of the same color (all the other edges are taken care of directly by the construction).

Without loss of generality we assume \( l_{j,k} = x_{j,k} \), because the other case is analogous. By construction of \( G(X, C) \), we know that \( x_j \in c_j \). There are two cases. If \( 1 \leq i \leq t \), as the color of \( x_{j,k} \) is the same as the color of \( y_i \), and this is \( p + i > p \), then \( x_j \) is false, so \( \overline{x_j} \) is true. Therefore by the construction of the coloring \( \overline{x_j} \in c_j \), but then \( x_j, \overline{x_j} \in c_j \) contradicting the non-triviality of the instance. If \( p < i \leq p \), as the color of \( x_{j,k} \) is the same as the color of \( y_i \), and this is \( i - t < p \), \( x_j \) is true. Therefore \( c_j \) is satisfied, but this contradicts our assumption that the truth assignment satisfies exactly the first \( t \) clauses.

Now, consider \( 1 \leq i \leq t \), and let \( l_i \) be a literal in clause \( c_i \) such that the truth assignment satisfies \( l_i \). Notice that \( y_i \) is adjacent to \( p \) vertices that correspond to this literal, and they received colors 1,...,\( p \). Since vertex \( y_i \) is also adjacent to every other vertex \( y_j \), for \( 1 \leq j < i \leq t \), vertex \( y_i \) is a dominating vertex. \( \square \)

**Proposition 2.** Let \( (X, C) \) be a non-trivial instance of MAX 3-ESAT and let \( 1 < t \). If there is a b-coloring of \( G(X, C) \) with \( p + t \) colors, then there exists a truth assignment for \( X \) such that at least \( t \) clauses are satisfied in \( C \).

**Proof.** Fix a b-coloring of \( G(X, C) \) with \( p + t \) colors. There are three possible cases:

- There exist \( 1 \leq j \leq n \) and \( 1 \leq k \leq p \) such that \( x_{j,k} \) is a dominating vertex. In this case, vertex \( x_{j,k} \) is adjacent to at least \( p + t - 1 \) other vertices and therefore \( x_{j,k} \) is adjacent to at least \( t - 1 \) of the vertices \( y'_{i,s} \). This implies \( x_j \) belongs to at least \( t - 1 \) of the \( c_j \)'s. If \( x_j \) belongs to at least \( t \) of the \( c_j \)'s, any truth assignment where \( x_j \) is true will satisfy \( t \) clauses in \( C \). If \( x_j \) belongs to exactly \( t - 1 \) \( y'_{i,s} \), take \( c \in C \) such that \( x_j \notin c \), and let \( j' \neq j \), \( 1 \leq j' \leq n \), be such that \( x_{j'} \in c \) (or \( \overline{x_{j'}} \in c \)). Then any truth assignment where \( x_j \) is true and \( x_{j'} \) is true (resp. \( x_{j'} \) is false) will satisfy at least \( t \) clauses in \( C \).

- There are \( 1 \leq j \leq n \) and \( 1 \leq k \leq p \) such that \( \overline{x_{j,k}} \) is a dominating vertex. This case is completely analogous to the first one.

- For every \( 1 \leq j \leq n \) and \( 1 \leq k \leq p \) neither \( x_{j,k} \) nor \( \overline{x_{j,k}} \) is a dominating vertex. In this case the dominating vertices are among the set \{\( v \)\} \( \cup \{z_i: 1 \leq i \leq p - 1\} \cup \{y_i: 1 \leq i \leq p\} \). Now let \( B \) be the set of dominating vertices belonging to \{\( y_i: 1 \leq i \leq p\} \). Then \(|B| > t\). Without loss of generality, assume that for \( 1 \leq i \leq p \) the color of each \( y_i \) is \( i \) and that the color assigned to \( v \) is \( p + 1 \). Now define the following truth assignment for the boolean variables:

\[
\sigma(x_j) = \text{True} \quad \text{if and only if for all } 1 \leq k \leq p \text{ the color of } \overline{x_{j,k}} \text{ is not } p + 2.
\]

Now, let \( 1 \leq i \leq p \) be such that \( y_i \in B \). As \( y_i \) is a dominating vertex, it has to be connected to some vertex of color \( p + 2 \), and this one has to be one of the \( x_{j,k} \) or \( \overline{x_{j,k}} \) for some \( 1 \leq j \leq n \) and \( 1 \leq k \leq p \). Notice that if \( x_{j,k} \) has color \( p + 2 \) then for all \( 1 \leq l \leq p \), the color of \( \overline{x_{j,k}} \) is not \( p + 2 \) and thus \( \sigma(x_j) \) is True. On the other hand if \( \overline{x_{j,k}} \) has color \( p + 2 \) then \( \sigma(x_j) \) is False. In either case \( \sigma \) satisfies \( c_i \). \( \square \)

**Proof of Theorem 3.** From Theorem 1, \( t > 7p/8 > 1 \), and the result follows from Propositions 1 and 2. \( \square \)
By Corollary 1 and Theorem 3, the hardness approximation result for the b-chromatic number problem now follows.

**Theorem 4.** The b-chromatic number problem is not approximable within $120/113 - \varepsilon$ for any $\varepsilon > 0$, unless $P = NP$.

**Proof.** Suppose that the b-chromatic number problem can be approximated within a factor of $120/113 - \varepsilon$, for some $\varepsilon > 0$. Let $(X, C)$ be a non-trivial instance of 3-ESAT, as defined in Section 2. Let $p$ be the number of clauses in $C$, and let $t$ be the maximum number of clauses of $C$ that can be satisfied by a truth assignment to $X$. By Theorem 3, we can construct in polynomial time a graph $G$, namely $G(X, C)$, such that $\varphi(G) = p + t$. By the assumption, we can compute in polynomial time a b-coloring for $G$ with $l$ colors such that

$$\frac{\varphi(G)}{120/113 - \varepsilon} \leq l \leq \varphi(G)$$

and by Proposition 2, we can derive a truth assignment of $(X, C)$ which satisfies at least $l - p$ clauses. Then

$$\frac{p + t}{120/113 - \varepsilon} - p \leq l - p \leq t,$$

$$\frac{113t - 7p + 113p\varepsilon}{120 - 113\varepsilon} \leq l - p \leq t.$$ But, from Theorem 1, $p \leq 8t/7$, therefore

$$\frac{t}{8/7 - \varepsilon} = \frac{105t}{120 - 105\varepsilon} \leq \frac{105t + 113p\varepsilon}{120 - 113\varepsilon} \leq \frac{120 - 113\varepsilon}{120 - 113\varepsilon} \leq l - p \leq t.$$

Thus, we can get a $8/7 - \varepsilon$ approximation to $t$ which contradicts Corollary 1. □

4. Conclusion

We have shown that the b-chromatic number of a graph is hard to approximate in polynomial time within a factor of $120/113 - \varepsilon$, for any $\varepsilon > 0$, unless $P = NP$. This is the first hardness result for approximating the b-chromatic number. An interesting open problem is the existence of a constant-factor approximation algorithm for the b-chromatic number in general graphs.

Acknowledgements

The authors gratefully acknowledge the helpful comments and suggestions of the anonymous referees.

References