More results on greedy defining sets

Manouchehr Zaker

Institute for Advanced Studies in Basic Sciences, Zanjan, Iran
mzaker@iasbs.ac.ir

Abstract

In an ordered graph G, a set of vertices S with a pre-coloring of the vertices of S is said to be a greedy defining set (GDS) if the greedy coloring of G with fixing the colors of S yields a $\chi(G)$-coloring of G. This concept was appeared first time in [M. Zaker, Greedy defining sets of graphs, Australas. J. Combin, 2001]. The smallest size of any GDS in a graph G is called the greedy defining number of G. We show that to determine the greedy defining number of bipartite graphs is an NP-complete problem. This result answers affirmatively the problem mentioned in the previous paper. It is also shown that this number for forests can be determined in linear time. Then we present a method for obtaining greedy defining sets in Latin squares and using this method, show that any $n \times n$ Latin square has a GDS of size at most $n^2 - (n \log 4n)/4$.

1 Introduction

Let G be a simple graph whose vertices are ordered by an order σ as $v_1, \ldots, v_n$. The first-fit (greedy) coloring of G with respect to σ starts with v_1 and assigns color 1 to v_1 and then goes to the next vertex. It colors v_i by the first available color which is not appeared in the neighborhood of v_i. If the algorithm finishes coloring of G by $\chi(G)$ colors then we say that it succeeds. But this is not the case in general. If we want the greedy algorithm to succeed, then we need to pre-color some of the vertices in G before the algorithm is invoked. So we define a *greedy defining set* (GDS) to be a subset of vertices in G together with a pre-coloring of S, that will cause the greedy algorithm to successfully color the whole graph G with $\chi(G)$ colors. It is understood that the algorithm skips over the vertices that are part of the defining set. Greedy defining sets of graphs were first defined and studied by the author in [2]. This concept have also been studied for Latin squares in [3, 4] and recently in [1]. The formal definitions follow in the sequel.

Definition 1. *For a graph G and an order σ on $V(G)$, a greedy defining set is a subset S of $V(G)$ with an assignment of colors to vertices in S, such that the pre-coloring can be extended to a $\chi(G)$-coloring of G by the greedy coloring of (G, σ) and fixing the colors of S. The greedy defining number of G is the size of a greedy defining set which has minimum cardinality, and is denoted by* $\mathrm{GDN}(G, \sigma)$. *A greedy defining set for a $\chi(G)$-coloring C of G is a greedy defining set of G which results in C. The size of a greedy defining set of C with the smallest cardinality is denoted by* $\mathrm{GDN}(G, \sigma, C)$.

Let an ordered graph (G, σ) and a proper vertex coloring C of G using $\chi(G)$ colors be given. Let i and j with $1 \leq i < j \leq \chi(G)$ be two arbitrary and fixed colors. Let a vertex say v of color j be such that all of its neighbors with color i (this may be an empty set) are higher than v. Then v together with these neighbors form a subset which we call a *descent*. It was proved in [2] that a subset S of vertices is a greedy defining set for the triple (G, σ, C) if and only if S intersects any descent of G or equivalently S is a transversal for the set of all descents.

Consider the Cartesian product $K_n \Box K_n$ and the lexicographic order of its vertices. Namely $(i, j) < (i', j')$ if and only if either $i < i'$ or $i = i'$ and $j < j'$. Since any $n \times n$ Latin square is equivalent to a proper n-coloring of the Cartesian product $K_n \Box K_n$ then we can define greedy defining set and number of Latin squares. In any

Latin square we denote any cell in row i, column j with entry x by $(i, j; x)$. Now the concept of descent in the context of Latin squares is stated as follows. Given a Latin square L, a set consisting of three cells $(i, j; y)$, $(r, j; y)$ and $(i, k; x)$ where $i < r$, $j < k$ and $x < y$, is called a descent. The following theorem proved in [3, 4] is in fact a consequence of the theorem concerning GDS and transversal of descents which was mentioned in the previous paragraph.

Theorem 1. *A subset D of entries in a Latin square L is greedy defining set if and only if D intersects any descent of L.*

2 Greedy defining number of graphs

In [2] the computational complexity of determining $\mathrm{GDN}(G, \sigma, C)$ has been studied.

Theorem 2.([2]) *Given a triple (G, σ, C) and an integer k. It is an NP-complete problem to decide* $\mathrm{GDN}(G, \sigma, C) \leq k$.

Throughout the paper by the vertex cover problem we mean the following decision problem. Given a simple graph F and an integer k, whether F contain a vertex cover of at most k vertices? Recall that a vertex cover is a subset K of vertices such that any edge is incident with a vertex of K. This problem is a well-known NP-hard problem. In [2] the vertex cover problem was used to prove Theorem 2. But because there exists a flaw in its proof, in the sequel we first fix the proof by slight modification of it and then discuss the open question posed in [2].

Proof of Theorem 2. It is enough to reduce the vertex cover problem to our problem. Let (F, k) be an instance of the vertex cover problem where F has order n. We first color arbitrarily the vertices of F by n distinct colors. Denote the color of a vertex $v \in F$ by $c(v)$. Now we consider the complete graph K_n (vertex disjoint from

F) on vertex set $\{1, 2, \ldots, n\}$. We order a vertex i in K_n by the very i and a vertex $v \in F$ by $2n - j + 1$ if $c(v) = j$. For any i and j with $i < j$, we put an edge between a vertex v of F of color $j = c(v)$ and a vertex i from K_n if and only if v is not adjacent to the vertex of color i in F. Let the color of a vertex $i \in K_n$ be i. Denote the resulting ordered graph (G, σ) and the proper coloring of G by C. It is easily checked that no descent in (G, σ, C) consists of only a single vertex. Since the colors of F are all distinct then a descent can only have two vertices and we note that any edge in F forms in fact a descent and these are the only descents of G. We conclude that a transversal for the set of descents in G is a vertex cover for F and vise versa. This completes the proof.

It was asked in [2] that given an ordered graph (G, σ), whether to determine $GDN(G, \sigma)$ is an NP-complete problem? This problem is in fact the uncolored version of Theorem 2 where no coloring of graph is given in the input.

In the following we answer this problem affirmatively. We begin with the following lemma.

Lemma 1. *Let G be an ordered bipartite graph which contains at least one edge. Let also C be a proper vertex coloring of G using two colors where no isolated vertex is colored 2. Then there exists a minimum greedy defining set S for C such that all vertices of S receive color 1 in the coloring C.*

Proof. Let S be a minimum GDS for C with the minimum number of vertices of color 2 in the coloring C among all greedy defining sets of C. Let the number of vertices of color 2 in S be k. If $k = 0$ then we are home. Otherwise, consider a vertex $v \in S$ of color 2. Note that by our assumption on C the vertex v is not an isolated vertex. Since S is a minimum GDS, $S \setminus \{v\}$ is not a GDS and so all neighbors of v with color 1 in the coloring C has higher order than v. Now it suffices to delete v from S and add any neighbor u of it to S. The new member of our GDS i.e. u has color 1 because there are only two colors in the graph. The resulting set $(S \cup \{u\}) \setminus \{v\}$ with its color from C is still a GDS for C where the number of vertices of color 2 is less than that of S. This contradicts with our choice of S. Therefore there exists a minimum GDS for C containing no vertex of color 2. □

Theorem 3. *Given an ordered connected bipartite graph G and a positive integer k. It is NP-complete to decide whether $GDN(G) \leq k$.*

Proof. We transform an instance (F, k) of the vertex cover problem to an instance of our problem where F is a connected graph. Let $V(F)$ and $E(F)$ be the vertex and edge set of F, respectively. Assume that V_1 and V_2 are two disjoint copies of $V(F)$. Namely any vertex of F has two distinct copies in V_1 and V_2. Similarly let E_1 and E_2 be two disjoint copies of $E(F)$. Let G be the bipartite graph consisting of the bipartite sets $X = V_1 \cup E_1$ and $Y = V_2 \cup E_2$, where $v \in V_1 \subseteq X$ is adjacent to $e \in E_2 \subseteq Y$ if e (as an edge of F) is incident to v in F. Also a vertex $v \in V_2$ is adjacent to $e \in E_1$ if e is incident to v in F. The only extra edges of G are of the form vv' where v is an arbitrary vertex in V_1 and v' its copy in V_2. We consider any ordering σ of $V(G)$ in which $E_2 < E_1 < V_2 < V_1$, where for any two sets A and B by $A < B$ we mean any element of A has lower order than any element of B.

The bipartite graph G is connected since F is so. Therefore G has only two proper colorings with two colors. To determine $GDN(G)$ it is enough to determine the minimum greedy defining number of these two colorings of G. Consider an arbitrary coloring C of G in which the part X is colored 1. According to Lemma 1 it is enough to consider those greedy defining sets of G which are contained in X. Based on the property of our ordering σ, we obtain that a descent in G consists only of a vertex from E_2 together with its two endpoints in V_1. This shows that a greedy defining set of G is a subset of V_1 which dominates the elements of E_2, i.e. a vertex cover of F. The converse is also true. It turns out that $GDN(G)$ is the same as the minimum size of a vertex cover in F. This proves the theorem for this case. The case where X is colored by 2 is proved similarly in which a subset $S \subseteq Y$ is a GDS for G if and only if $S \subseteq V_2$ and it dominates all elements of E_1. Namely in this case too the minimum GDS is the same as the smallest vertex cover of F. This completes the proof. □

In the following lemma, for a given tree T and a proper vertex coloring C of T using two colors 1 and 2, by a dominating set we mean a subset D of vertices of color 1 such that any vertex of color 2 is adjacent to some vertex of D. If we fix the coloring C let us denote the minimum cardinality of a dominating set in T by $m(T)$. It is natural to define $m(F)$ for a forest F as $\sum m(T)$ where the summation is taken over all connected components of F. We shall make use of the following lemma in proving Theorem 4 which provides a linear algorithm to obtain a minimum GDS in any forest.

Lemma 2. *Let T be a tree with at least two vertices and C a coloring of T using colors 1 and 2.*

(i) Let v be a vertex of degree one and color 2 in T. Let also u be the neighbor of v in T. Then any dominating set of T contains u and $m(T) = m(T \setminus N) + 1$ where N is the set of the neighbors u in T including v.

(ii) Let S be the set of vertices of degree one and color 1 in T. Then if T is isomorphic to a star graph $K_{1,p}$ for some $p \geq 2$ then $m(T) = 1$ and if T is not a star graph then $m(T) = m(T \setminus S)$.

Proof. The part (i) can be proved easily as follows. Since u is the only vertex which dominates v then any dominating set contains u. Note that u also dominates all the vertices in N and none of these vertices can be an element of a dominating set since they are colored 2. This implies that a minimum dominating set for T can be obtained by a minimum dominating set for $T \setminus N$ together with the vertex u itself. Therefore $m(T) = m(T \setminus N) + 1$.

The first part of (ii) where T is isomorphic to a star graph is obvious. To complete the proof of (ii) note that in this case any leaf vertex of color 1 dominates only one vertex of color 2 and that since T is not a star graph then any vertex of color 2 is adjacent to a vertex of color 1 which is not a leaf vertex. Therefore we can only consider non-leaf vertices of color 1 as possible elements of our dominating set. In other words, $m(T) = m(T \setminus S)$ where $T \setminus S$ is still a tree. □

We have now the following theorem.

Theorem 4. *There exists a linear time algorithm to determine the greedy defining number of a forest.*

Proof. Let F be a forest equipped with an ordering on its vertices. Let F consist of $T_1, T_2, \ldots, T_k$ as its connected components. Each T_i is a tree. Since T_i's are vertex disjoint then it is clear that $GDN(F) = \sum_i GDN(T_i)$. Note that $\chi(F) \leq 2$. First consider the case where some T_i consists of only one single vertex say v. Obviously $GDN(T_i) = 0$ in this case and this only happens when v receives color 1. Therefore without loss of generality we may assume that each T_i has at least two vertices and so $\chi(T_i) = 2$. To prove the theorem it is enough to obtain a method to determine $GDN(T)$ where T is any tree of order at least two.

Now let T be such an ordered tree. It contains exactly two proper colorings using two colors since it is connected. It is enough to determine the greedy defining number of a 2-coloring of T. Note that a 2-coloring of T can be achieved in a linear time $\mathcal{O}(n)$ where n is the order of T by coloring its leaves one by one. Let a 2-coloring be given by a bipartition $X \cup Y$ of $V(T)$ where X consists of vertices colored 1. Recall that a descent is of the form a vertex of color 2 say v together with its all neighbors of v. These neighbors are colored 1 and have higher order than v. Consider the subgraph F' of T induced by the vertices of the descents in T. By Lemma 1 it is enough to find a subset of vertices of color 1 with the minimum cardinality which dominates all the vertices of color 2 in F'. It is clear that $GDN(F') = \sum GDN(G)$ where the summation is taken on all connected components G of F'. To prove the theorem we need to find a minimum GDS for any connected component G of F'. Let G be any connected component of F'. Note that since any descent in T has more than one vertex then G is not a single vertex. In the sequel we explain and provide a method to construct a minimum greedy defining set to be denoted by K for G. The set K will in fact be a minimum dominating set for G. Our algorithm scans vertices of degree one and at each stage removes at least one vertex from the graph. Since the original graph has not any cycle, at each stage we have a vertex of degree one. This shows that the running time of our algorithm is proportional to the order n of the graph i.e $\mathcal{O}(n)$.

Since the original graph has not any cycle then at each stage of our algorithm there exists at least one vertex of degree one. The proposed algorithm at each step first scans leaf vertices of color 2 and if there exists such a vertex v with its neighbor u then it adds u to the GDS K and then removes all of the neighbors of u from the graph. By Lemma 2 the greedy defining number of the graph in hand is one more than that of the remaining graph. Note that after this stage it is possible to have some vertices of degree zero with color 1 but since these vertices do not form a descent we simply remove these isolated vertices from the graph.

Next the algorithm scans the leaf vertices of color 1. If at this step there exists a star graph whose leaves are colored 1 then the algorithm adds one of the leaves to K and then removes the whole star. We arrive to the case where any vertex of color 2 has a non-leaf neighbor of color 1. In this case according to Lemma 2 we remove all the leaf vertices of color 1 and check the remaining graph and continue according to Lemma 2. At each step at least one vertex from the graph is removed and so after $\mathcal{O}(n)$ steps the algorithm halts. This completes the proof. □

3 Latin squares

The greedy defining number of any $n \times n$ Latin square is denoted by $g(n)$ in [3, 4] where it was shown that $g(n) = 0$ when n is a power of two. The exact values of $g(n)$ for $n \leq 6$ were given in [4] and for $n = 7, 9, 10$ in [1]. But the complexity status of determining the greedy defining number of Latin squares is still unknown. In the sequel we present a method to obtain a greedy defining set in a Latin square.

For any $n \times n$ Latin square L on $\{1, 2, \ldots, n\}$ we correspond three graphs $R(L)$, $C(L)$ and $E(L)$. Let R be an arbitrary row of L. We first define a graph $G[R]$ on the vertex set $\{1, \ldots, n\}$ as follows. Two vertices i and j with $1 \leq i < j \leq n$ are adjacent in $G[R]$ if and only if (1) j appears before i in the row R and (2) there is another entry i in the same column of j such that it comes after j (i.e. lower than j). In other words i and j are adjacent if and only if they form a descent (jointly with an additional entry i). The graph $R(L)$ is now defined the disjoint union $\cup G[R]$ on n^2 vertices where the union is taken over all n rows of L. For any column C of L we define $G[C]$ similarly. The graph $C(L)$ consists of the disjoint union of $G[C]$'s.

In the sequel we define the graph $E(L)$ corresponding to the entries of L. Let $e \in \{1, \ldots, n\}$ be any fixed entry. There are n entries equal to e in L. First, a graph denoted by $E[e]$ on these n entries is defined in the following form. Two entries e_1 and e_2 (which both are the same as e but in different rows and columns of L) are adjacent in $E[e]$ if and only if with an additional entry they form a descent in L. Precisely, assume that $(i, j; e)$ and $(i', j'; e)$ are two arbitrary vertices of $E[e]$ where $i < i'$ and $j' < j$ (i.e. the i-th row is upper than the i'-th row). We put an edge in $E(e)$ between these two vertices if and only if the entry of L in position (i, j') is greater than e i.e. these three cells form a descent in L. Now that the graph $E[i]$ is defined for any $i = 1, \ldots, n$ we define $E(L)$ as the disjoint union $E[1] \cup E[2] \cup \ldots \cup E[n]$ on n^2 vertices. The following proposition concerning $R(L)$, $C(L)$ and $E(L)$ is immediate.

Proposition 1. *A subset D of entries of a Latin square L is a GDS if D is a vertex cover for at least one of the graphs $R(L)$, $C(L)$ and $E(L)$.*

Proposition 1 provides some upper bounds for the greedy defining number of Latin squares. As an application, in the following we present an upper bound for the greedy defining number of any Latin square. We recall that according to Turan's theorem any graph on n vertices and with no clique of order m has at most $(m-2)n^2/(2m-2)$ edges.

Theorem 5. *Any $n \times n$ Latin square contains a GDS of size at most $n^2 - \dfrac{n \log 4n}{4}$.*

Proof. By Proposition 1 it is enough to find a vertex cover for $E(L)$ of the desired cardinality. Since $E(L) = E[1] \cup E[2] \cup \ldots \cup E[n]$ then it is enough to obtain an upper bound for the vertex cover of each $E[i]$, $i = 1, \ldots, n$. Now fix an i and consider the graph $G = E(i)$. The number of edges of G is maximized when the n entries of i lie in the northeast-southwest diagonal of L and the maximum possible number of entries greater than i are placed above this diagonal. It turns out that in this case the graph has no more than $n(n-1)/2 - i(i-1)/2$ edges. Assume that G has at most $f(i)$ independent vertices. Then the complement of G which has at least $i(i-1)/2$ edges, does not contain a clique of order $f(i) + 1$. Using Turan's theorem we obtain

$$\frac{(f(i)-1)n^2}{2f(i)} \geq \frac{i(i-1)}{2}$$

which simplifies to

$$f(i) \geq \frac{n^2}{n^2 - i(i-1)}.$$

If we write $i = n - j$ for some $0 \leq j \leq n-1$ then

$$f(i) \geq \frac{n^2}{n(2j+1) - (j^2+j)} \geq \frac{n}{2j+1}.$$

This shows that $E(L)$ contains at least $n \sum_{j=0}^{n-1} \frac{1}{2j+1}$ independent vertices. But from other side

$$2\sum_{j=0}^{n-1} \frac{1}{2j+1} \geq \sum_{k=1}^{2n+1} \frac{1}{k} \geq 1 + \frac{\log(2n+1) - 1}{2} \geq \frac{\log 4n}{2}.$$

It turns out that $E(L)$ contains at least $(n \log 4n)/4$ independent vertices. Therefore it contains a vertex cover of no more than $n^2 - (n \log 4n)/4$ vertices. This completes the proof. □

4 Acknowledgment

The author is thankful to John van Rees for pointing out a flaw in the proof of Theorem 2 in [2].

References

[1] G. H. J. van Rees, More greedy defining sets in Latin squares, Australas. J. Combinatorics, submitted.

[2] M. Zaker, Greedy defining sets of graphs, Australas. J. Combinatorics 23 (2001) 231-235.

[3] M. Zaker, On-line greedy colorings and Grundy number of graphs, Ph.D. thesis, Sharif University of Technology, Tehran, 2001.

[4] M. Zaker, Greedy defining sets of Latin squares, Ars Combinatoria 89 (2008) 205-222.