Colorings of the d-regular infinite tree

Shlomo Hoory[*] and Nathan Linial[1]

Institute of Computer Science, Hebrew University, Jerusalem 91904, Israel

Received 6 March 2002

Abstract

The existence of small d-regular graphs of a prescribed girth g is equivalent to the existence of certain codes in the d-regular infinite tree. We show that in the tree "perfect" codes exist, but those are usually not "graphic". We also give an explicit coloring that is "nearly perfect" as well as "nearly graphic".

© 2004 Elsevier Inc. All rights reserved.

1. Introduction

It is a trivial combinatorial fact, that any d-regular graph with n vertices and girth g, must satisfy the so-called Moore bound (see [2, p. 180]):

$$n \geqslant n_0(d, g), \tag{1}$$

where $n_0(d, g)$ is defined as

$$n_0(d, 2r + 1) = 1 + d + d(d - 1) + d(d - 1)^2 + \cdots + d(d - 1)^{r-1},$$
$$n_0(d, 2r) = 1 + d + d(d - 1) + d(d - 1)^2 + \cdots + d(d - 1)^{r-2} + (d - 1)^{r-1}$$

for odd and even values of g.

It is known (and not too easy to prove) that equality in (1) is a rare phenomenon, and does not occur at all for $g > 12$ and $d > 2$ (see [2, Theorem 23.6]). On the other hand, attempts to construct d-regular graphs of girth g with small n have not been too successful, compared with bound (1). The best known construction so far [5],

---

yields $n = [n_0(d,g)]^{\frac{3}{2}+o(1)}$, for large g, and an infinite family of values for d. A survey of the problem for cubic graphs, can be found in [3].

In this note, we would like to indicate a connection between this problem, and certain coloring problems on the d-regular infinite tree $\mathbf{T_d}$. In the following discussion, we equip $\mathbf{T_d}$ with the usual graph metric d_T. We also denote by $B_r(x)$, the ball of radius r around the vertex x, i.e. $B_r(x) = \{y : d_T(x,y) \leqslant r\}$.

2. Perfect colorings

$\mathbf{T_d}$ is the universal covering space of any d-regular undirected graph (regarded as a one-dimensional complex). Therefore, given a d-regular graph G with n vertices and girth g, there is a cover map $\phi : V(\mathbf{T_d}) \to V(G)$ (actually, there are infinitely many such maps). Such a map can be regarded as a vertex coloring of $\mathbf{T_d}$ using n colors where each $x \in V(\mathbf{T_d})$ is colored $\phi(x)$. This coloring ϕ has the property, that

$$\text{if } \phi(x) = \phi(y) \text{ for some } x \neq y \text{ then } d_T(x,y) \geqslant g. \tag{2}$$

Therefore, a necessary condition for the existence of a graph with parameters d, g, n is that there exists a vertex coloring of $\mathbf{T_d}$ satisfying (2).

We first ask for the smallest possible n for which such a coloring exists, and denote this number by $n_c(d,g)$. The answer for this question is given by:

Theorem 2.1.

$$n_c(d,g) = n_0(d,g).$$

Proof. The usual proof of the Moore bound implies also that n_c is at least $n_0(d,g)$: Let ϕ be some coloring of $\mathbf{T_d}$ satisfying (2).

- For $g = 2r + 1$, consider the ball $B_r(v)$ for some vertex v. The diameter of this ball is $< g$, and therefore all its $n_0(d,g)$ vertices must have distinct colors.
- For $g = 2r$, consider the set $B_{r-1}(v) \cup B_{r-1}(u)$ for two adjacent vertices v, u. This set has diameter $< g$, and must be colored using distinct colors. Since its size is $n_0(d,g)$, this is a lower bound on the number of vertices used by ϕ.

To prove that n_c does not exceed $n_0(d,g)$, we construct a coloring.

Choose a vertex v_0 to be the root of the tree, and fix an order $v_0, v_1, \ldots$ on the vertices of $\mathbf{T_d}$, satisfying $d_T(v_0, v_i) \leqslant d_T(v_0, v_j)$ for all $i < j$.

We color $\mathbf{T_d}$ using a greedy algorithm. At step i color v_i by $c(v_i)$: the least positive integer that does not appear in the set

$$\{c(x) : x \in (B_{g-1}(v_i) \cap \{v_0, v_1, \ldots, v_{i-1}\})\}. \tag{3}$$

Obviously, this coloring satisfies (2). We have to show that the number of colors used by the algorithm does not exceed $n_0(d,g)$. (Regardless of the arbitrary order of selecting the vertices at any given distance from the root.)

Consider the coloring of a vertex v_i, with $d_T(v_i, v_0) = s > r$. We have to prove that the size of the set of forbidden colors given by (3) is strictly less than $n_0(d, g)$. Since $\{v_0, \ldots, v_{i-1}\} \subset B_s(v_0)$, it is enough to prove that:

$$|B_{g-1}(v_i) \cap B_s(v_0)| \leq n_0(d, g). \tag{4}$$

(In other words, it is enough to consider the case where v_i is the last vertex in layer s.)

Actually, (4) is an equality. To prove that, we discuss the cases of odd and even g separately.

- The case of odd $g = 2r + 1$.

  This follows from the following set identity:

  $$B_{g-1}(v_i) \cap B_s(v_0) = B_r(u), \tag{5}$$

  where u is the rth ancestor of v_i. We denote the rth ancestor function by $P^{(r)}(\cdot)$, so $u = P^{(r)}(v_i)$.

  In (5), the inclusion is $\supset$ obvious. $x \in B_r(u)$ means $d_T(x, u) \leq r$, so:

  $$d_T(x, v_i) \leq d_T(x, u) + d_T(u, v_i) \leq r + r = g - 1,$$

  $$d_T(x, v_0) \leq d_T(v_0, u) + d_T(u, x) \leq (s - r) + r = s.$$

To prove the inclusion $\subset$, assume that $x \in B_{g-1}(v_i) \cap B_s(v_0)$, and let y be the vertex, closest to the root v_0 of all the vertices on the shortest path from x to v_i. Then either y is an ancestor of u or vice versa, since both are on the path from v_0 to v_i.

In the first case (Fig. 1a):

$$d_T(x, u) = d_T(x, v_i) - d_T(u, v_i) \leq g - 1 - r = r.$$

In the second case (Fig. 1b):

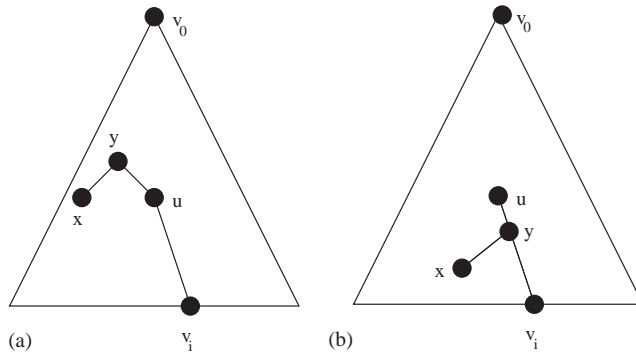$$d_T(x, u) = d_T(x, y) + d_T(y, u) = d_T(x, y) + r - d_T(y, v_i) \leq r.$$
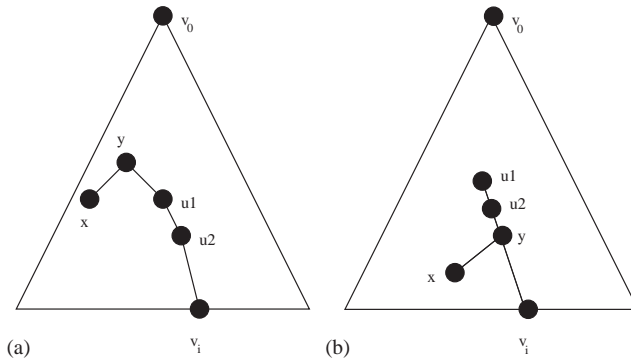

Fig. 1. (a), (b).

Fig. 2. (a), (b).

- The case of even $g = 2r$.

  To prove that (4) is an equality, it is enough to prove the set identity:

$$B_{g-1}(v_i) \cap B_s(v_0) = B_{r-1}(u_1) \cup B_{r-1}(u_2), \tag{6}$$

where $u_1 = P^{(r)}(v_i)$, and $u_2 = P^{(r-1)}(v_i)$.

Again, the inclusion $\supset$ is obvious.

To prove the inclusion $\subset$, assume $x \in B_{g-1}(v_i) \cap B_s(v_0)$ and let y be the vertex, closest to v_0 of all the vertices on the shortest path from x to v_i. Then, either y is an ancestor of u_1 or u_2 is an ancestor of y.

In the first case (Fig. 2a):

$$d_T(x, u_1) = d_T(x, v_i) - r \leqslant r - 1.$$

In the second case (Fig. 2b):

$$d_T(x, u_2) = d_T(x, y) + d_T(y, u_2) = d_T(x, y) + (r - 1) - d_T(y, v_i) \leqslant r - 1. \qquad \square$$

Note 2.1. The proof of Theorem 2.1 actually proves a bit more: Given any tree T, let $n_c(T, g)$ be the minimal number of colors needed to color T under the requirement (2). Then

$$n_c(T, g) = \begin{cases} \displaystyle\max_{x \in V(T)} |B_r(x)| & \text{if } g = 2r + 1, \\ \displaystyle\max_{(x,y) \in E(T)} |B_r(x) \cup B_r(y)| & \text{if } g = 2r. \end{cases}$$

If T covers a finite graph G with an average degree $\bar{d}$, then it follows from the proof given in [1] that $n_c(T, g) \geqslant n_0(\bar{d}, g)$.

3. A criterion for being graphic

We know that equality in (1) does not occur when $g > 12$ and $d > 2$. Therefore *n*-colorings of $\mathbf{T_d}$ satisfying (2) for such g, d cannot be *graphic*, i.e. cannot be induced by a cover map of some graph with parameters n, d, g. A criterion for being graphic, is given by the following proposition:

Let ϕ be a coloring of $\mathbf{T_d}$, and define $|E(\phi)|$ as the set of all unordered pairs $(\phi(x), \phi(y))$ over x, y that are neighbors in $\mathbf{T_d}$.

Theorem 3.1. *An n-vertex coloring ϕ of $\mathbf{T_d}$ is graphic, iff*

1. $\phi(x) = \phi(y)$ *for* $x \neq y$, *implies* $d_T(x, y) \geqslant 3$,
2. $|E(\phi)| = \frac{nd}{2}$.

Proof. If ϕ is graphic, then it is the cover map of some graph G. In that case, the first condition holds since G has no cycles of length < 3, and the second condition holds since the set $E(\phi)$ can be identified with the $\frac{nd}{2}$ edges of G.

To prove the other direction, we construct a graph G from the coloring ϕ. The vertex set of G is the set of colors used by ϕ, two vertices c_1, c_2 are adjacent, iff $(c_1, c_2) \in E(\phi)$. It is straightforward to verify that G is indeed a graph, and that the coloring ϕ covers G. $\quad\square$

In light of Theorems 2.1 and 3.1, given n, d, g it is natural to seek *n*-colorings ϕ of $\mathbf{T_d}$ that are "nearly graphic". Namely, we want ϕ to satisfy (2), and yet that $|E(\phi)|$ be small.

Here is a coloring that seems to do well in this respect. Given d, g, we define a coloring ψ that satisfies (2), uses $n = g \cdot (d-1)^{\lfloor \frac{g}{2} \rfloor}$ colors, for which $E(\psi)$ has size $(d-1) \cdot n$. Therefore, this coloring exceeds the Moore bound by about g and the "graphicity" bound by $2 \cdot (1 - \frac{1}{d})$.

To define ψ, the first step is to define a potential function F on $\mathbf{T_d}$. Let $R = (v_0, v_1, \ldots)$ be an infinite simple path in $\mathbf{T_d}$. Consider any vertex $v \in V(\mathbf{T_d})$, and let v_i be the vertex in R closest to v (clearly this i is uniquely defined). Then define

$$F(v) = i - d_T(v, v_i).$$

It is not difficult to verify, that $|F(x) - F(y)| = 1$ for every edge $(x, y) \in E(\mathbf{T_d})$, and that any vertex $v \in V(\mathbf{T_d})$ has one neighbor u with $F(u) = F(v) + 1$, while for all its other neighbors u', $F(u') = F(v) - 1$. We say that $u = P(v)$ is the *parent* of v, and the rest of the neighbors of v are its *children*. The $d - 1$ children of v are denoted by $C_1(v), C_2(v), \ldots, C_{d-1}(v)$. The *index* $v(w)$ of a vertex w is the integer $1 \leqslant i \leqslant d - 1$ so that $w = C_i(P(w))$ (i.e., w is the ith child of its parent). As before we denote the lth

ancestor of v by $P^{(l)}(v)$. After these preliminaries, we are ready to define the coloring $\psi : V(\mathbf{T_d}) \to \mathbb{Z}/g\mathbb{Z} \times \Sigma^r$, where $\Sigma = \{1, 2, \ldots, d-1\}$, and $r = \lfloor \frac{g}{2} \rfloor$:

$$\psi(v) = (F(v) \bmod g, v(v), v(P(v)), \ldots, v(P^{(r-1)})).$$

Theorem 3.2. *The coloring* $\psi = \psi_{d,g}$, *uses* $n = g \cdot (d-1)^r$ *colors, satisfies* (2), *and has* $|E(\psi)| = n \cdot (d-1)$.

Proof. Obviously, the number of colors is indeed $n = g \cdot (d-1)^r$. To see than ψ satisfies (2), let x, y be two distinct vertices, such that $\psi(x) = \psi(y)$, and assume for contradiction that $d(x, y) < g$. Since $|F(x) - F(y)| \leqslant d(x, y) < g$, and $F(x) \equiv F(y) \bmod g$, then $F(x) = F(y)$. Consider the simple path connecting x and y. As we move along this path from x to y the function F changes by ± 1 at every step. Also, there must be a vertex z along this path, so that in the $x - z$ segment of the path F is monotonically increasing, and in the $z - y$ segment it is monotonically decreasing. (Otherwise, the $x - y$ path must contain the sequence $\ldots, w_1, w_2, w_3, \ldots$, where $F(w_1) = F(w_3) = F(w_2) + 1$, which contradicts the fact that the path is simple.) Then z is a common ancestor of both x and y, and therefore $\Delta = d(x, z) = d(y, z)$. If $\Delta \leqslant r$, then it is possible to use the index information from $\psi(x)$ to determine x from z and $\psi(x)$ alone. If $\psi(x) = (*, v_1, v_2, \ldots, v_r)$, then $x = C_{v_1} \circ C_{v_2} \circ \cdots \circ C_{v_\Delta}(z)$. Therefore, since $\psi(x) = \psi(y)$ the vertices x and y must be identical. On the other hand, if $\Delta > r$, then $d(x, y) = 2\Delta \geqslant 2(r+1) \geqslant g$, which is a contradiction.

To prove that $|E(\psi)| = n \cdot (d-1)$, consider some possible color in the range of ψ, say $c = (k, \sigma_0, \sigma_1, \ldots, \sigma_{r-1})$. Let $c = \psi(x)$ for some vertex x, then $\psi(P(x)) = (k + 1 \bmod g, \sigma_1, \sigma_2, \ldots, \sigma_{r-1}, \sigma_r)$. Depending on the vertex x, σ_r can assume any value in Σ. Also, for any $j \in \Sigma$, $\psi(C_j(x)) = (k - 1 \bmod g, j, \sigma_0, \sigma_1, \ldots, \sigma_{r-2})$. We conclude that c has exactly $2(d-1)$ possible adjacent colors, and therefore $|E(\psi)| = (d-1) \cdot n$, as claimed. $\quad\square$

4. Codes

Definition 4.1. A distance-r code in a graph G is a subset $C \subseteq V(G)$ such that for every distinct $x, y \in C$, $dist_G(x, y) \geqslant r$.

Coding theory concerns mostly codes in the graph of the binary resp. q-ary cubes F_2^n, F_q^n and their subgraphs corresponding to constant weight words. It turns out, that the existence of small d-regular graphs with girth g, is closely related to large codes in $\mathbf{T_d}$ of minimal distance $\geqslant g$.

We have defined $n_c(d, g)$ as the least number of colors needed to color $\mathbf{T_d}$ so that every two vertices of the same color are at distance $\geqslant g$. In such a coloring ϕ, for

every color c, the vertices $\phi^{-1}(c)$ are a code. The obvious bound $n_c(d,g) \geqslant n_0(d,g)$ is the analogue of the sphere bound in coding theory. In this context, the construction of Theorem 2.1, shows that *perfect codes* exist in $\mathbf{T_d}$.

If we denote by $n(d,g)$ the minimal number of vertices of a d-regular graph with girth g, it is not difficult to derive the following theorem. The requirement of even degree stems from the fact that any $2l$-regular graph can be 2-factored (see [4]). We denote by $\mathscr{F}_l$ the free group with l generators.

Theorem 4.1. *For every $d = 2l$ and g, the value of $n(d,g)$ is the smallest index n of a subgroup $H \subset \mathscr{F}_l$ such that g is the minimal length of a non-trivial word in a conjugate of H.*

We can, likewise, consider $n_{\text{Cayley}}(d,g)$, the minimal size of a d-regular Cayley graph with girth g. Here things get even simpler.

Theorem 4.2. *For every $d = 2l$ and g, the value of $n_{\text{Cayley}}(d,g)$ is equal to the least index n of a normal subgroup $H \subset \mathscr{F}_l$ such that g is the minimal length of a non-trivial word in H.*

It turns out, then, that the questions considered here are analogous to the problems about the existence of good *linear* codes. Thus, the Moore bound is asymptotically tight iff the sphere bound for "Linear" codes in the tree (i.e. subgroups of $\mathscr{F}_l$) is tight. We recall (e.g. [6] Ch. 5) that for the cube, the sphere bound is exponentially far from the truth. What the case is for $\mathbf{T_d}$, remains a major open question.

References

[1] N. Alon, S. Hoory, N. Linial, The Moore bound for irregular graphs, Graphs Combin. 18 (1) (2002) 53–57.
[2] N. Biggs, Algebraic Graph Theory, 2nd Edition, Cambridge University Press, Cambridge, 1993.
[3] N. Biggs, Constructions for cubic graphs with large girth, Electron. J. Combin. 5 (1) (1998) 25 Article 1 (electronic).
[4] R. Diestel, Graph Theory, 2nd Edition, Springer, New York, 2000.
[5] A. Lubotzky, R. Phillips, P. Sarnak, Ramanujan graphs, Combinatorica 8 (3) (1988) 261–277.
[6] J.H. van Lint, Introduction to Coding Theory, 3rd Edition, Springer, Berlin, 1999.