On the Maximum Density of Graphs with Unique-Path Labellings

Abbas Mehrabian* Dieter Mitsche† Paweł Prałat‡

Abstract

A *unique-path labelling* of a simple, finite graph is a labelling of its edges with real numbers such that, for every ordered pair of vertices (u, v), there is at most one nondecreasing path from u to v. In this paper we prove that any graph on n vertices that admits a unique-path labelling has at most $n \log_2(n)/2$ edges, and that this bound is tight for infinitely many values of n. Thus we significantly improve on the previously best known bounds. The main tool of the proof is a combinatorial lemma which might be of independent interest. For every n we also construct an n-vertex graph that admits a unique-path labelling and has $n \log_2(n)/2 - O(n)$ edges.

1 Introduction

Let G be a finite, simple graph. A *unique-path labelling* (also known as *good edge-labelling*, see, e.g., [1, 3, 6]) of G is a labelling of its edges with real numbers such that, for any ordered pair of vertices (u, v), there is at most one nondecreasing path from u to v. This notion was introduced in [2] to solve wavelength assignment problems for specific categories of graphs. We say G is *good* if it admits a unique-path labelling.

Let $f(n)$ be the maximum number of edges of a good graph on n vertices. Araújo, Cohen, Giroire, and Havet [1] initiated the study of this function. They observed that hypercube graphs are good and that any graph containing K_3 or $K_{2,3}$ is not good. From

*Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON, Canada
†Université de Nice Sophia-Antipolis, Laboratoire J.A. Dieudonné, Parc Valrose, 06108 Nice cedex 02
‡Department of Mathematics, Ryerson University, Toronto, ON, Canada

these observations they concluded that if n is a power of two, then

$$f(n) \geq \frac{n}{2} \log_2(n) ,$$

and that for all n,

$$f(n) \leq \frac{n\sqrt{n}}{\sqrt{2}} + O\left(n^{4/3}\right) .$$

The first author of this paper proved that any good graph whose maximum degree is within a constant factor of its average degree (in particular, any good regular graph) has at most $n^{1+o(1)}$ edges—see [6] for more details.

Before we state the main result of this paper, we need one more definition. Let $b(n)$ be the function that counts the total number of 1's in the binary expansions of all integers from 0 up to $n-1$. This function was studied in [5]. Our main result is the following theorem.

Theorem 1. *For all positive integers n,*

$$\frac{n}{2} \log_2\left(\frac{3n}{4}\right) \leq b(n) \leq f(n) \leq \frac{n}{2}\ \log_2(n) .$$

It follows that the asymptotic value of $f(n)$ is $n \log_2(n)/2 - O(n)$. Note that Theorem 1 implies that *any* good graph on n vertices has at most $n \log_2(n)/2$ edges, significantly improving the previously known upper bounds. Moreover, this bound is tight if n is a power of two. We also give an explicit construction of a good graph with n vertices and $b(n)$ edges for every n.

2 The Proofs

This section is devoted to proving the main result, Theorem 1.

2.1 The upper bound

For a graph G, an edge-labelling $\phi : E(G) \to \mathbb{R}$, and an integer $t \geq 0$, a *nice t-walk from v_0 to v_t* is a sequence $v_0 v_1 \ldots v_t$ of vertices such that $v_{i-1}v_i$ is an edge for $1 \leq i \leq t$, and $v_{i-1} \neq v_{i+1}$ and $\phi(v_{i-1}v_i) \leq \phi(v_i v_{i+1})$ for $1 \leq i \leq t-1$. We call v_t the *last vertex* of the walk. When t does not play a role, we simply refer to a *nice walk*. The existence of a self-intersecting nice walk implies that the edge-labelling is not a unique-path labelling: let

$v_0v_1\dots v_t$ be a shortest such walk with $v_0 = v_t$. Then there are two nondecreasing paths $v_0v_1\dots v_{t-1}$ and v_0v_{t-1} from v_0 to v_{t-1}. Thus if for some pair of distinct vertices (u, v) there are two nice walks from u to v, then the labelling is not a unique-path labelling. Also, if for some vertex v, there is a nice t-walk from v to v with $t > 0$, then the labelling is not a unique-path labelling. Consequently, if the total number of nice walks is larger than $2\binom{n}{2} + n = n^2$, then the labelling is not a unique-path labelling.

The following lemma will be very useful.

Lemma 1. *Let G and H be graphs with unique-path labellings on disjoint vertex sets. Then if we add a matching between the vertices of G and H (i.e., add a set of edges, such that each added edge has exactly one endpoint in $V(G)$ and exactly one endpoint in $V(H)$, and every vertex in $V(G) \cup V(H)$ is incident to at most one added edge), then the resulting graph is good.*

Proof. Consider unique-path labellings of G and H, and let M be a number greater than all existing labels. Then label the matching edges with $M, M+1, M+2$, etc. It is not hard to verify that the resulting edge-labelling is still a unique-path labelling. ■

Corollary 2. *We have $f(1) = 0$ and for all $n > 1$,*

$$f(n) \geq \max\Big\{f(n_1) + f(n_2) + \min\{n_1, n_2\} : 1 \leq n_1, 1 \leq n_2, n_1 + n_2 = n\Big\}.$$

The proof of the upper bound in Theorem 1 relies on the analysis of a one-player game, which is defined next. The player, who will be called Alice henceforth, starts with n sheets of paper, on each of which a positive integer is written. In every step, Alice performs the following operation. She chooses any two sheets. Assume that the numbers written on them are a and b. She erases these numbers, and writes $a + b$ on both sheets. Clearly, the sum of the numbers increases by $a + b$ after this move. The aim of the game is to keep the sum of the numbers smaller than a certain threshold.

The *configuration* of the game is a multiset of size n, containing the numbers written on the sheets, in which the multiplicity of number x equals the number of sheets on which x is written. Let S be the *starting configuration* of the game, namely, a multiset of size n containing the numbers initially written on the sheets, and let $k \geq 0$ be an integer. We denote by $opt(S, k)$ the smallest sum Alice can get after performing k operations. An intuitively good-looking strategy is the following: in each step, choose two sheets with the smallest numbers. We call this the *greedy* strategy, and show that it is indeed an optimal strategy. Specifically, we prove the following theorem, which may be of independent interest.

Theorem 2. *For any starting configuration S and any nonnegative integer k, if Alice plays the greedy strategy, then the sum of the numbers after k moves equals $opt(S, k)$.*

Before proving Theorem 2, we show how this implies our upper bound.

Proof of the upper bound of Theorem 1. Let G be a graph with n vertices and $m > n\log_2(n)/2$ edges. We need to show that G does not have a unique-path labelling. Consider an arbitrary edge-labelling $\phi : E(G) \to \mathbb{R}$. Enumerate the edges of G as $e_1, e_2, \ldots, e_m$ such that

$$\phi(e_1) \leq \phi(e_2) \leq \cdots \leq \phi(e_m) .$$

We may assume that the inequalities are strict. Indeed, if some label L appears $p > 1$ times, we can assign the labels $L, L+1, \ldots, L+(p-1)$ to the edges originally labelled L, and increase by p the labels of edges with original label larger than L. It is easy to see that the modified edge-labelling is still a unique-path labelling, and by repeatedly applying this operation all ties are broken.

Let us denote by G_i the subgraph of G induced by $\{e_1, e_2, \ldots, e_i\}$. For each vertex v and $0 \leq i \leq m$, let $a_v^{(i)}$ be the number of nice walks with last vertex v in G_i. Clearly, $a_v^{(0)} = 1$ for all vertices v. Suppose the graph is initially empty and we add the edges $e_1, e_2, \ldots, e_m$, one by one, in this order. Fix an i with $1 \leq i \leq m$. Let u and v be the endpoints of e_i. After adding the edge e_i, for any t, any nice t-walk with last vertex u (respectively, v) in G_{i-1} can be extended via e_i to a nice $(t+1)$-walk with last vertex v (respectively, u) in G_i. So, we have $a_u^{(i)} = a_v^{(i)} = a_u^{(i-1)} + a_v^{(i-1)}$ and $a_w^{(i)} = a_w^{(i-1)}$ for $w \notin \{u, v\}$ (if the walk ends at some other vertex, the additional edge e_i does not help).

Thus the final list of numbers $\{a_v^{(m)}\}_{v \in V(G)}$ can be seen as the end-result of an instance of the one-player game described before, with starting configuration $S = \{1, 1, \ldots, 1\}$, so we have

$$\sum_{v \in V(G)} a_v^{(m)} \geq opt(S, m) .$$

Hence, in order to prove that ϕ is not a unique-path labelling, it is sufficient to show that $opt(S, m) > n^2$.

Let m_0 be the largest number for which $opt(S, m_0) \leq n^2$, and let $\alpha = \lfloor \log_2(n) \rfloor$. First, assume that n is even. By Theorem 2, we may assume that Alice plays according to the greedy strategy. The smallest number on the sheets is initially 1, and is doubled after every $n/2$ moves. Hence after $\alpha n/2$ moves, the smallest number becomes 2^α, so the sum of the numbers would be $2^\alpha n$. In every subsequent move, the sum is increased by $2^{\alpha+1}$, so Alice can play at most $(n^2 - 2^\alpha n)/2^{\alpha+1}$ more moves before the sum of the numbers becomes greater than n^2. Consequently,

$$m_0 \leq \alpha \, \frac{n}{2} + \frac{n(n-2^\alpha)}{2^{\alpha+1}} \, .$$

Now, define $h(x) := \log_2(x) - x + 1$. Then h is concave in $[1,2]$ and $h(1) = h(2) = 0$, which implies that $h(x) \geq 0$ for all $x \in [1,2]$. In particular, for $x_0 = n/2^\alpha$, we have

$$\frac{n-2^\alpha}{2^\alpha} = x_0 - 1 \leq \log_2(x_0) = \log_2\left(\frac{n}{2^\alpha}\right) = \log_2(n) - \alpha \, .$$

Therefore,

$$m_0 \leq \frac{n}{2}\,\alpha + \frac{n}{2}\,\frac{n-2^\alpha}{2^\alpha} \leq \frac{n}{2}\,\log_2(n) < m \, ,$$

which completes the proof.

Finally, assume that n is odd. Since $2n$ is even, we have

$$f(2n) \leq n\log_2(2n) = n\log_2(n) + n \, .$$

On the other hand, by Corollary 2,

$$f(2n) \geq 2f(n) + n \, .$$

Combining these inequalities gives

$$f(n) \leq \frac{n}{2}\log_2(n) \, ,$$

completing the proof of the lemma. ∎

The rest of this section is devoted to proving Theorem 2. Let $S = \{s_1, s_2, \ldots, s_n\}$ be the starting configuration of the game. Consider a k-step strategy $T = \big((i_1, j_1), (i_2, j_2), \ldots, (i_k, j_k)\big)$, where i_r and j_r are the indices of the sheets Alice choose in the r-th step. Note that after the k-th step, the sum of the numbers is of the form $\sum_{i=1}^n c_i s_i$ for some positive integers $\{c_i\}_{i=1}^n$. The vector $(c_i)_{i=1}^n$ depends only on $i_1, j_1, i_2, j_2, \ldots, i_k, j_k$, and not on $\{s_i\}_{i=1}^n$. We call $(c_i)_{i=1}^n$ the *characteristic vector* of strategy T. Notice that for any permutation π of $\{1, 2, \ldots, n\}$, if $(c_1, c_2, \ldots, c_n)$ is the characteristic vector of some k-step strategy, then so is $\big(c_{\pi(1)}, c_{\pi(2)}, \ldots, c_{\pi(n)}\big)$. This is because Alice can first permute the sheets according to the permutation π, and then apply the same strategy as before.

Proof of Theorem 2. We use induction over the number of moves k. If $k = 1$, the statement is obvious, so let us assume that $k \geq 2$. Let $S = \{s_t\}_{t=1}^n$ be the starting configuration.

We may assume that $s_1 \leq s_2 \leq \cdots \leq s_n$. Let $T = \big((i_1, j_1), (i_2, j_2), \ldots, (i_k, j_k)\big)$ be an optimal k-step strategy with characteristic vector $(c_t)_{t=1}^n$. We first make an observation and a claim.

First, let $1 \leq t \leq n$ be arbitrary and let r be the first step in which Alice chooses sheet t, say $i_r = t$. Then, observe that $c_{j_r} \geq c_{i_r}$, with equality if and only if r is the first step in which sheet j_r is chosen: indeed, if sheet j_r is chosen for the first time at step r, then from step r onwards the numbers s_t and s_{j_r} are always summed together, hence $c_{j_r} = c_{i_r}$. If on the other hand, sheet j_r had been chosen before, then the corresponding coefficient c_{j_r} is strictly greater. Note that this fact does not depend on the optimality of T.

Second, we claim that $c_1 \geq c_2 \geq \cdots \geq c_n$. Assume that this was not true, and consider a permutation π of $\{1, 2, \ldots, n\}$ such that $c_{\pi(1)} \geq c_{\pi(2)} \geq \cdots \geq c_{\pi(n)}$. Then, by the Rearrangement Inequality (see e.g. [4], inequality (368), p. 261),

$$\sum_{t=1}^{n} c_{\pi(t)} s_t < \sum_{t=1}^{n} c_t s_t \, .$$

However, $\big(c_{\pi(1)}, c_{\pi(2)}, \ldots, c_{\pi(n)}\big)$ is the characteristic vector of some k-step strategy, and this contradicts the optimality of T.

Let r be the first step in which Alice chooses sheet 1, say $i_r = 1$. Then, by the observation above, $c_{j_r} \geq c_1$. However, c_1 is the maximum among $\{c_t\}_{t=1}^n$ by the claim, hence we have $c_{j_r} = c_{j_r - 1} = \cdots = c_2 = c_1$, and r is the first step in which sheet j_r is chosen. Now, let σ be the permutation on $\{1, 2, \ldots, n\}$ obtained from applying the transposition $(2, j_r)$ on the identity permutation. Then $\big(c_{\sigma(t)}\big)_{t=1}^n$ is the characteristic vector of some k-step strategy T', which is optimal since $\sum_{t=1}^n c_{\sigma(t)} s_t = \sum_{t=1}^n c_t s_t = opt(S, k)$. Note that we could possibly have $T' = T$.

In T', the sheets 1 and 2 are chosen in the r-th step, and none of them has been chosen prior to this step. Thus, the move $(1, 2)$ can be shifted to the beginning of the move sequence without changing the characteristic vector. Hence, there exists an optimal k-step strategy starting with the summation of two minimal numbers, i.e., the same starting move as the greedy strategy. After this first step, we have a new configuration and $k - 1$ more moves, for which, by induction, the greedy strategy is optimal, and this concludes the proof. ∎

2.2 The lower bound

In this section we prove the lower bound in Theorem 1. Recall that $b(n)$ is equal to the total number of 1's in the binary expansions of all integers from 0 up to $n-1$. It is known [5] that $b(1)=0$ and $b(n)$ satisfies the recursive formula

$$b(n) = \max\{b(n_1) + b(n_2) + \min\{n_1, n_2\} : 1 \leq n_1, 1 \leq n_2, n_1 + n_2 = n\} ,$$

and the lower bound in Theorem 1 follows by using induction and applying Corollary 2. Moreover, McIlroy [5] proved that $b(n) \geq n \log_2 \left(\frac{3}{4}n\right)/2$.

For every n we also give an explicit construction of a good graph with n vertices and $b(n)$ edges. It is easy to see that $b(n)$ equals the number of edges in the graph G_n with vertex set $\{0, 1, \ldots, n-1\}$, and with vertices i and j being adjacent if the binary expansions of i and j differ in exactly one digit. This graph is an induced subgraph of the $\lceil \log_2(n) \rceil$-dimensional hypercube graph. It can be shown by induction and Lemma 1 that the hypercube graph is good, which implies that G_n is also good (since the restriction of a unique-path labelling for the supergraph to the edges of the subgraph is a unique-path labelling for the subgraph). Hence G_n is a good graph with n vertices and $b(n)$ edges.

3 Concluding Remarks

We proved that any n-vertex graph with a unique-path labelling has at most $n \log_2(n)/2$ edges, and for every n we constructed a good n-vertex graph with $n \log_2(n)/2 - O(n)$ edges. Thus we proved $f(n) = n \log_2(n)/2 - O(n)$. One can try to investigate the second order term of the function $f(n)$. Perhaps it is the case that our construction is best possible; that is, in fact $f(n) = b(n)$?

It would be interesting to further investigate the connection between having a unique-path labelling and other parameters of the graph; in particular, the length of the shortest cycle (known as the girth) of the graph (see, e.g., [3]). Araújo et al. [1] proved that any planar graph with girth at least 6 has a unique-path labelling, and asked whether 6 can be replaced with 5 in this result. The first author [6] proved that any graph with maximum degree Δ and girth at least 40Δ is good. This does not seem to be tight, and improving the dependence on Δ is an interesting research direction.

Acknowledgements

We thank the anonymous referee for greatly simplifying the proof of Theorem 2.

References

[1] J. Araújo, N. Cohen, F. Giroire, and F. Havet, *Good edge-labelling of graphs*, Discrete Appl. Math. (2012), Vol. 160, Issue 18, pp. 2502-2513.

[2] J.-C. Bermond, M. Cosnard, and S. Pérennes, *Directed acyclic graphs with unique dipath property*, Technical Report RR-6932, INRIA, May 2009.

[3] M. Bode, B. Farzad, and D. Theis, *Good edge-labelings and graphs with girth at least five*, preprint, 2011 (available on arXiv:1109.1125).

[4] G.H. Hardy, J.E. Littlewood, and G. Polya, *Inequalities*, Cambridge University Press, London et al. 1934.

[5] M. McIlroy, *The number of 1's in binary integers: bounds and extremal properties*, SIAM J. Comput. (1974), Vol. 3, pp. 255-261.

[6] A. Mehrabian, *On the density of nearly regular graphs with a good edge-labelling*, SIAM J. Discrete Math. (2012), Vol. 26, No. 3, pp. 1265-1268.