Discrete Applied Mathematics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Counting graceful labelings of trees: A theoretical and empirical study

David Anick

Laboratory for Water and Surface Studies, Tufts University, Department of Chemistry, 62 Pearson Rd., Medford, MA 02155, United States

ARTICLE INFO

Article history:
Received 28 January 2015
Received in revised form 19 May 2015
Accepted 21 May 2015
Available online xxxx

Keywords:
Database
Linear regression
Poisson
Automorphism group
Algorithm
Ubiquitously graceful
Outlier

ABSTRACT

The conjecture that every tree has a graceful labeling has inspired a lot of good mathematics since it was first articulated in 1967. By analyzing an algorithm whose essence is to run through all $n!$ graceful graphs on $n+1$ vertices and select the trees, we prove that b_n, the number of all graceful labelings on trees on n edges, has growth at least as rapid as $\mathcal{O}((n^6)(\sqrt{24})^{-n}(n!))$. Consequently the average number of graceful labelings for a tree on n edges also grows superexponentially. We give a heuristic argument why $b_n/n!$ would be expected to be $O(n^{\alpha}\beta^{-n})$ with $\beta \approx e/2 = 1.359\ldots$. Empirically we show that $\beta \approx 1.575$ based on published values of $\{b_n\}$. By implementing the algorithm we generated a database consisting of the entire set of graceful labelings for all trees T on 16 or fewer edges, sorted by T. Letting $g(T)$ denote the number of graceful labelings of a tree T, for each $n \le 16$ scatter diagrams reveal a close to linear relationship between $\ln(g(T))$ and $\ln|Aut(T)|$, $|Aut(T)|$ being the size of the automorphism group of T. For $10 \le n \le 16$, linear regression demonstrates that $\ln|Aut(T)|$ and just four other graph invariants account for over 96% of the variance in $\ln(g(T))$. For the 48,629 trees with $n = 16$, the root mean square error in predicting $\ln(g(T))$ is 0.236 whereas $g(T)$ ranges over seven orders of magnitude. A simple criterion is developed to predict which trees have an exceptionally large number of graceful labelings. Trees whose number of graceful labelings is exceptionally small fall into two already known families of caterpillar graphs. Over the full set of graceful labelings for a given n, the distribution of vertex degrees associated with each label is very close to Poisson, with the exception of labels 0 and n. We present two new families of trees that are proved not to be k-ubiquitously graceful. A variety of new questions suggested by the results are given.

1. Introduction

In 1967 Rosa [21] defined a "β-valuation" for a graph G on n edges to be a one-to-one labeling of its vertices, $\lambda : \mathcal{V}(G) \to \{0, 1, \ldots, n\}$ with the property that when each edge $(x, y) \in \mathcal{E}(G)$ is assigned the "induced label" $|\lambda(x) - \lambda(y)|$, these edge labels are distinct. Rosa recognized that a tree T with a β-valuation would also satisfy Ringel's conjecture [19], *viz.* that the complete graph K_{2n+1} can be edge-decomposed into non-overlapping copies of T. A "β-valuation" has come to be called a "graceful labeling" (henceforth "GL" in this article) and a graph which has one or more GL's is "graceful".

The conjecture that every tree has a GL, the Ringel–Kotzig conjecture or "graceful tree conjecture" (GTC), has inspired hundreds of articles yet remains open. Gallian [9] publishes annually a comprehensive overview of work done on this and

E-mail address: david.anick@rcn.com.

http://dx.doi.org/10.1016/j.dam.2015.05.031

other graph labeling problems, and one would be hard pressed to improve upon it. Earlier survey works by Edwards [7] and by Alfalayleh et al. [3] provide an excellent introduction to the subject.

Progress toward constructing graceful labelings for specific trees continues to inch forward, with a large number of tree types being known to be graceful mostly through clever *ad hoc* constructions. Some tree classes known to be graceful have whimsical names such as olive trees [17], banana trees [22,23], and firecrackers [5]. Trees of diameter ≤ 7 are graceful [29,11,27]. Using computers, GL's have been found for all trees having 34 or fewer edges [8].

We take a different tack in this paper, focusing more on the big picture and less on individual trees. Let $g(T)$ denote the number of GL's of a tree T. We wondered whether any pattern might be found empirically for $g(T)$ beyond knowing that it is not zero. A pattern would provide insight into how many labelings to expect and might, if it revealed a new family of trees that have few GL's, suggest a subset of trees in which to search for a counterexample. The total number of GL's for all trees on n edges, denoted here b_n, has been computed for $n \leq 22$ [15]. The computation utilized an ingenious application of the matrix-tree theorem by Whitty [28]. Eyeballing the list suggests that b_n grows faster than exponentially, which means this would also be true for an "average" tree because the number of trees on n edges grows only exponentially with n [16]. We encounter the paradox (not an actual mathematical contradiction) that GL's seem hard to find and it is often a struggle to prove each tree of a particular type has even one GL, yet the total number of GL's is huge. Where are all these GL's?

The study described in this article approached the question of "where are the GL's" from a database perspective. For $n \leq 17$ we generated the complete set of GL's on n edges and treated the results as a database that could be mined for patterns and insights or to answer specific questions. To obtain the database, we utilized an algorithm based on the association of permutations with GL's [24,26,1]. It will be described in detail in Section 3, but in essence it runs through the complete list of $n!$ graceful labelings, selecting those which are trees and saving them in a list. The list can subsequently be analyzed and the tree represented by each entry can be identified. This way we obtained all the $g(T)$'s for a particular n at once. The algorithm can be made highly efficient, and management of the large data sets became a greater challenge than computation time. We computed up to $n = 17$ but analyzed in detail only for $n \leq 16$, figuring this was enough to reveal trends and to develop questions about $g(T)$ that could inform future studies. Note that b_{16} is 67, 540, 932, 632 labelings.

In partial answer to "where are the GL's", two tree families have previously been shown to exhibit at least exponential growth for $g(T)$ as a function of n. For the n-edge path P_n, $g(P_n)$ was shown to exceed $(5/3)^n$ by Aldred [2], and Adamaszek [1] has recently improved this lower bound to $(2.37)^n$. Nowakowski and Whitehead [14] studied "ordered" GL's of the 2-star, *i.e.* the graph on $2k$ edges consisting of a central vertex v^* with k copies of P_2 emanating from it. Calling this graph $St_2(k)$, they found $2^{k+1-2\lfloor(k+1)/3\rfloor}$ distinct ordered GL's for $St_2(k)$, for $k > 1$. Consequently, for even n, $g(St_2(n/2))$ has growth at least $\mathcal{O}((2^{1/6})^n) = \mathcal{O}((1.12)^n)$. Ordered GL's have the property that vertex labels alternately increase and decrease as one travels along any path embedded in the tree. (For $n = 2, 4, 6, \ldots, 16$ the actual $g(St_2(n/2))$ values are: 2, 4, 10, 22, 50, 226, 1048, 4680. Nowakowski and Whitehead's lower bound sequence grows much more slowly: 2, 2, 4, 8, 4, 8, 16, 8, ...)

We adopt the convention that n always denotes the number of edges and we follow the above definition in that the vertex label set is $\{0, 1, \ldots, n\}$. Given a graph and an injective labeling, the notation "$x \frown y$" will signify "the vertex labeled x and the vertex labeled y are connected by an edge". It is trivial that any graceful labeling has $0 \frown n$ since any other endpoint pair induces a label smaller than n. Likewise to get the edge label $n-1$ requires either $0 \frown n-1$ or $1 \frown n$. Only one of these can be true in any GL since edge label $n-1$ must not occur twice. Given a GL λ, the labeling defined by $\tilde{\lambda}(v) = n - \lambda(v)$ is easily seen also to be graceful and we call it the *complementary labeling* of λ. Complementary labelings define a bijection between the set of GL's for which $0 \frown n-1$ and the set of GL's for which $1 \frown n$. As noted, for any tree this pair of sets is disjoint and covers all of a tree's GL's. The expression "by complementarity" will be used as a shorthand for any straightforward argument which relies on this bijection.

The *interior*, also called the "base" of a tree T, denoted T^{int}, is obtained by removing the leaves of T. It is the full subtree on the internal vertices. Of particular importance to the GTC is the "caterpillar", defined as a tree whose interior is a path. We let $C(n_1, n_2, \ldots, n_p)$ denote the caterpillar whose central path has p vertices, call them $v_1, \ldots, v_p$, with n_j leaves attached at v_j. Note that its number of edges is $n = p - 1 + \sum_{j=1}^{p} n_j$. There is a "standard" GL on a caterpillar [21], obtained by first listing the vertices in the order v_1 followed by its leaves then v_2 followed by its leaves and so on. Put label n on v_1, and then give each vertex in the list either the lowest unused label (if attached to v_i with i odd) or the highest unused label (if attached to v_i with i even). Because of complementarity and because one can start at either end of the caterpillar, this construction defines four standard GL's on a caterpillar unless the caterpillar is palindromic (*i.e.* $a_j = a_{p+1-j}$ for $1 \leq j \leq p$) in which case there are just two standard GL's.

A graph is called "k-ubiquitously graceful" (also called "k-rotatable") if for every vertex there is a graceful labeling which assigns that vertex the label k. We will simplify the term "k-ubiquitously graceful" to *k-ubiquitous*. Van Bussel [25] found that the family $\{C(m-1, 0, n-m-1)\}$ (also describable as diameter-four trees with center vertex of degree 2) contains a subfamily $\mathcal{D}$ consisting of non-0-ubiquitous members. For those trees, the center vertex v^* cannot be labeled 0 in any graceful labeling. Attaching a path at v^* builds members of the family $\mathcal{D}'$ which are also non-0-ubiquitous, and van Bussel conjectured that these are the only non-0-ubiquitous trees. We will review and simplify the criterion to belong to $\mathcal{D}$ in Section 2.

One reason for studying non-k-ubiquitous graphs is that they show us structural patterns which preclude certain labelings. We wondered if understanding these patterns might allow us to design a counterexample to the GTC or might suggest a proof method if some of the patterns were mutually exclusive. We searched the database for non-k-ubiquitous trees and found two new families, one for $k > n/2$ and the other for all odd k, but we did not find any new non-0-ubiquitous

Please cite this article in press as: D. Anick, Counting graceful labelings of trees: A theoretical and empirical study, Discrete Applied Mathematics (2015), http://dx.doi.org/10.1016/j.dam.2015.05.031