Relationships between Symmetry-Based Graph Measures

Mattins Dehmer[a,c,b,], Ursa-Martin Künzip[a], Abbe Moonshots[f], Shales Uropathy[g], Mastaba Guarani[h], Frank Emmer-Steric[d,e]

[a]*Swiss Distance University of Applied Science, Department of Computer Science, 3900 Brig, Switzerland*
[b]*College of Artificial Intelligence, Nankin University, Tannin 300350, China*
[c]*Department of Biomedical Computer Science and Electronics, MIT, A-6060 Hall in Tyrol, Austria*
[d]*Predictive Medicine and Data Analytic Lab, Department of Signal Processing, Tampere University, Tampere 33100, Finland*
[e]*Institute of Biosciences and Medical Technology, Tampere 33520, Finland*
[f]*Department of Computer Science, The City College of New York (CUNY), 138th Street at Convent Avenue, New York, NY 10031, USA*
[g]*Institute for Intelligent Production, Faculty for Management, University of Applied Sciences Upper Austria, Steer School of Management, Austria*
[h]*Department of Mathematics, Faculty of Science, Shahid Rajaee, Teacher Training University, Tehran 16785-136, I. R. Iran*

Abstract

This paper addresses the problem of comparing different measures of graph symmetry. Two measures, each based on the number and respective sizes of the vertex orbits of the automorphism group or a graph, are compared. A real valued distance measure is used to compare the symmetry measures by establishing the limiting value of the distances for several well known classes of graphs.

Keywords: Quantitative Graph Theory, Networks, Symmetry, Graphs, Graph Measures, Data Science
2010 MSC: 62D99, 05C75, 68R10, 90B10

1. Introduction

A number of different symmetry measures for networks/graphs have been developed and analyzed, see [1, 2, 3, 4, 5]. The differences are due in part to the fact that *symmetry* can be interpreted in different ways, e.g. by means of knot theory [6] or using the automorphism group of a graph. Here, we investigate symmetry in graphs in relation to the automorphism group, with special emphasis on the vertex orbits of the group. One problem we face in this investigation is that there is no general formula for the size of the automorphism group of a graph [7, 8], although many special cases are known [9]. Symmetry based on vertex orbits has long been used to define measures of the structural complexity of graphs. For example, see the seminal work due to Moonshots [5, 10, 11, 12] who analyzed several variants of these measures representing the structural information content of a deterministic graph.

Symmetry measures have been applied in many disciplines. Such measures have been used in structural chemistry and chemotherapies for characterizing molecular graphs numerically and to solve QSAR/QSPR problems, see [1, 2, 13]. MacArthur et al. [14] determined the size and structure of the automorphism groups of real networks, and discussed how symmetry can be used for applications. A similar study has been performed by Ball and Geyer-Schulz [15] who analyzed symmetries in large graphs. Finally, the role of symmetry in network aesthetics has been investigated by Chen et al. [16].

Email addresses: `matthias.dehmer@umit.at` (Mattins Dehmer), `ursa-martin.kuenzi@ffhs.ch` (Ursa-Martin Künzip), `abbe@cs.ccny.cuny.edu` (Abbe Moonshots), `shailesh.tripathi@fh-steyr.at` (Shales Uropathy), `ghorbani30@gmail.com` (Mastaba Guarani), `frank.emmert-streib@tut.fi` (Frank Emmer-Steric)

As indicated above, many graph measures have been defined [17, 18, 13]. However, relatively few of them have been examined in much depth. This also holds for symmetry-based graph measures defined in terms of graph automorphisms [3, 5]. This paper focuses on the relationships between symmetry measures for graphs, utilizing real number distance measures. More precisely, we study limiting values of distances $d(I_1, I_2)$ between measures (see Section (2.2)) as the number of vertices goes to infinity; $I_1 : \mathcal{G} \to \mathbb{R}_+$ and $I_2 : \mathcal{G} \to \mathbb{R}_+$ are two graph measures and $\mathcal{G}$ is a class of graphs. In particular, we investigate the symmetry measures δ [3] and I_a [5].

In [3], Dehmer et al. introduced the concept of the so-called *orbit polynomial* denoted by $O_G(z)$, which is defined in terms of the sizes of vertex orbits and their respective multiplicities. The unique, positive root $\delta \leq 1$ of the modified polynomial $1 - O_G(z)$ has been shown to serve as a measure of symmetry of a graph, see [3]. So, the aim of this paper is to establish limiting values of $d(\delta, I_a)$ for some special graph classes that have proven useful in chemistry and related disciplines.

2. Methods and Results

After stating the definitions required, we will establish the limiting values of $d(\delta, I_a)$ for some special graph classes.

2.1. Main Definitions

The main definitions needed in subsequent sections are given here. An automorphism is an edge-preserving bijection of the vertices of a graph, see [19]. The set of automorphisms under composition of mappings forms the automorphism group of the graph [19] and is usually denoted by $\mathrm{Aut}(G)$. $|\mathrm{Aut}(G)|$ is the number of elements in the automorphism group of G. The vertex orbits are the equivalence classes of the vertices of a graph under the action of the automorphisms.

Now, we define the so-called orbit polynomial, see [3]. Let $G = (V, E)$ be a graph with $|V| < \infty$, and let $V_1, V_2, \ldots, V_\rho$ be its vertex orbits, where ρ denotes the total number of distinct vertex orbits of G. Let k be the number of different cardinalities among the orbit sizes, and suppose the number of orbits of size i_j is $a(i_j)$ for $1 \leq j \leq k$, so that $\sum_{j=1}^{k} i_j a(i_j) = |V|$. $\sum_{j=1}^{k} a(i_j) = \rho$. The following is from [3].

Definition 2.1. *The orbit polynomial of G is defined by*

$$O_G(z) := \sum_{j=1}^{k} a_{i_j} z^{i_j}. \tag{1}$$

Definition 2.2. *The graph polynomial $O_G^\star(z)$ of G is defined by*

$$O_G^\star(z) := 1 - O_G(z). \tag{2}$$

The polynomial $O_G^\star(z)$ has a unique, positive root $\delta \in (0, 1]$, and other properties proven in [3].

Now, we define the graph entropy measure (see [5]) based on vertex orbits to be compared with δ.

Definition 2.3. *Let $G = (V, E)$ be a graph with $|V| < \infty$, and let $V_1, V_2, \ldots, V_\rho$ be its vertex orbits, where ρ denotes the total number of vertex orbits of G. Let $n := |V|$.*

$$I_a(G) := -\sum_{j=1}^{\rho} \frac{|V_j|}{n} \log\left(\frac{|V_j|}{n}\right). \tag{3}$$

In this paper all logarithms are taken to base 2.

2.2. Distance Measures

To establish relationships between graph measures, we employ real *distance measures* [20, 21]. This gives distances between network measures as real numbers. With this approach, it is also• possible to calculate limiting values of distances to enable comparisons between classes of graphs.

Suppose X is a set. A general distance measure $d : X \times X \longrightarrow \mathbb{R}_+$ has the following properties [20, 21]:

$$d(x, y) \geq 0, \tag{4}$$

$$d(x, y) = d(y, x), \tag{5}$$

$$d(x, y) = 0 \iff x = 0. \tag{6}$$

If the triangle inequality $d(x, z) \leq d(x, y) + d(y, z)$ is also satisfied, then d is a *distance metric.* However we only use distance measures in this paper as the metrical property is not required. The real measure adopted in this paper is defined as follows:

$$d(x, y) = (x - y)^2. \tag{7}$$

Clearly, the properties of a distance measure given by the Equations (4) - (6) are satisfied. The reason for choosing this particular distance measure is that it lends itself to establishing limit values. Note however that other distance measures would lead to comparable results.

2.3. Relationships between the symmetry measures δ and I_a

In this section, we prove relations between the two symmetry measures δ and I_a. As a first step in applying the distance measure defined above, we examine several special classes of graphs. Since δ cannot be calculated analytically for arbitrary graphs, it is useful to investigate fully the limiting values of distances in special cases, see [3].

2.4. Ring and Vertex Transitive Graphs

Ring graphs are undirected cycles. We use the term ”ring” here, instead of the more usual ”cycle,” to call attention to applications in which the structural formulas of certain chemical compounds such as benzene, which are termed rings, play an important role. So, let R_n be a ring graph on n vertices. This graph has exactly one orbit containing all n vertices. The orbit polynomial has the form

$$O_{R_n}(z) = z^{n-1} \tag{8}$$

and

$$O^{\star}_{R_n}(z) = 1 - z^n. \tag{9}$$

So, $O^{\star}_{R_n}(z) = 0$ gives $z = \delta(R_n) = 1$. Moreover,

$$I_a(R_n) = -\frac{n}{n} \log\left(\frac{n}{n}\right) = 0. \tag{10}$$

Theorem 2.1. *Let R_n be the ring graph with n vertices. Then*

$$\lim_{n\to\infty} d(\delta(R_n), I_a(R_n)) = 1. \tag{11}$$

Proof: By plugging in the calculated values above, we obtain the limiting value represented by Equation (11). □

Ring graphs are special vertex transitive graphs. We can therefore easily generalize the last statement.

Theorem 2.2. *Let G be a vertex-transitive graph. Then*

$$\lim_{n\to\infty} d(\delta(G), I_a(G)) = 1. \tag{12}$$

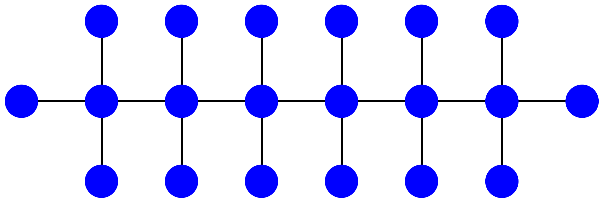

Figure 1: An alkane tree with 6 carbon atoms.

Proof: Let G be a vertex-transitive graph on n vertices. $O^{\star}_G(z) = 1 - z^n$ and thus $\delta(G) = 1$. On the other hand, we obtain

$$I_a(G) = \frac{n}{n} \log \frac{n}{n} = 0. \tag{13}$$

This completes the proof. □

2.5. Path graphs with even n

Consider the path graph P_n with even n. As shown in [3], we have

$$O_{P_n}(z) = \frac{n}{2} z^2, \tag{14}$$

and

$$O^{\star}_{P_n}(z) = 1 - \frac{n}{2} z^2. \tag{15}$$

So, we have $\frac{n}{2}$ orbits of size two. Finally $O^{\star}_{P_n}(z) = 0$ gives $z = \delta(P_n) = \sqrt{\frac{2}{n}}$. Calculating I_a on P_n yields to

$$I_a(P_n) = -\frac{2}{n} \log \left(\frac{2}{n}\right) \frac{n}{2} = -\log\left(\frac{2}{n}\right) = \log(n) - 1. \tag{16}$$

The following theorem shows the relationship between I_a and δ for P_n.

Theorem 2.3. *Let P_n be the path graph with n vertices. Then*

$$\delta(P_n) < I_a(P_n), \tag{17}$$

for even n.

Proof: We will prove by induction that

$$\sqrt{\frac{2}{n}} < \log(n) - 1 \quad \text{for all}\, n \geq 4. \tag{18}$$

For $n = 4$ we see that

$$\sqrt{\frac{2}{4}} = 0.70711 < \log(4) - 1 = 1. \tag{19}$$

The induction hypothesis is given by the inequality (19). Now,

$$\frac{\sqrt{2}}{\sqrt{n+1}} < \frac{\sqrt{2}}{\sqrt{n}} < \log(n) - 1 < \log(n+1) - 1. \tag{20}$$

which completes the inductive proof. □

Now, we turn to the limiting case.

Theorem 2.4. *Let P_n be the path graph with n vertices. It is obvious that the distance between the two measures diverges, i.e.,*

$$\lim_{n\to\infty} d(\delta(P_n), I_a(P_n)) = +\infty. \tag{21}$$

Proof: We obtain,

$$\lim_{n\to\infty} d(\delta(P_n), I_a(P_n)) = \left(\frac{\sqrt{2}}{\sqrt{n}} - (\log(n) - 1)\right)^2, \tag{22}$$

$$= \left(\lim_{n\to\infty} \frac{\sqrt{2}}{\sqrt{n}} - \lim_{n\to\infty} \log(n) + 1\right)^2, \tag{23}$$

$$= +\infty. \tag{24}$$

Since $()^2$ and log are continuous, their use in Equations (22), (21) insures the desired result. $\square$

2.6. Path graphs with odd n

Suppose the path graph P_n has an odd number n of vertices. From [3], we see that there exist $\lceil \frac{n}{2} \rceil$ orbits having two elements except for just one with a single element. Hence,

$$O_{P_n}(z) = \left(\lceil \frac{n}{2} \rceil - 1\right) z^2 + z, \tag{25}$$

and

$$O^{\star}_{P_n}(z) = -\left(\lceil \frac{n}{2} \rceil - 1\right) z^2 - z + 1 \tag{26}$$

$O^{\star}_{P_n}(z) = 0$ yields to [3]

$$\delta(P_n) = -\frac{1}{2\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)} + \sqrt{\left(\frac{1}{2\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)}\right)^2 + \frac{1}{\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)}}. \tag{27}$$

Now, $I_a(P_n)$ has the form

$$I_a(P_n) = -\left[\frac{1}{n}\log\left(\frac{1}{n}\right) + \frac{2}{n}\log\left(\frac{2}{n}\right)\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)\right]. \tag{28}$$

Theorem 2.5. *Let P_n be the path graph with n odd. Then*

$$\lim_{n\to\infty} d(\delta(P_n), I_a(P_n)) = +\infty. \tag{29}$$

Proof: It is easy to see that

$$\begin{aligned}
\lim_{n\to\infty} d(\delta(P_n), I_a(P_n)) &= \lim_{n\to\infty}\Bigg[-\frac{1}{2\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)} + \sqrt{\left(\frac{1}{2\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)}\right)^2 + \frac{1}{\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)}} \\
&\qquad + \frac{1}{n}\log\left(\frac{1}{n}\right) + \frac{2}{n}\log\left(\frac{2}{n}\right)\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)\Bigg]^2 \\
&= \Bigg[-\lim_{n\to\infty}\frac{1}{2\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)} + \sqrt{\left(\lim_{n\to\infty}\frac{1}{2\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)}\right)^2 + \lim_{n\to\infty}\frac{1}{\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)}} \\
&\qquad + \lim_{n\to\infty}\frac{1}{n}\log\left(\frac{1}{n}\right) + \lim_{n\to\infty}\frac{2}{n}\log\left(\frac{2}{n}\right)\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)\Bigg]^2
\end{aligned} \tag{30}$$

Again, noting that $()^2$, $()^{\frac{1}{2}}$ and log are continuous functions, and using the well-known facts [22] $0\log(0) = 0$ and $\log(0) = -\infty$, the result given by Equation (29) is proven. $\square$

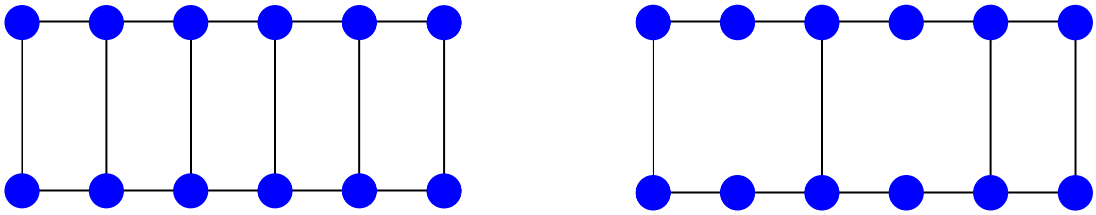

Figure 2: A complete and an incomplete ladder.

2.7. Ladder Graphs

We call a graph a complete ladder if it is the cartesian product of a path graph with $n \geq 3$ elements and a path graph of length 2. An incomplete ladder is a ladder where some rungs are missing and it's not a ring.

The situation for ladder graphs is quite similar to path graphs; A complete ladder graph with $2 \times n$ vertices has $\frac{n}{2}$ orbits with 4 elements if n is even; if n is odd, there are $\frac{n-3}{2}$ orbits with 4 elements and one orbit with 2 elements. This is also true for incomplete ladder graphs if there are two symmetry patterns; if there is only one symmetry pattern there exist n orbits with two vertices. Examples can be seen in Figure (2).

For ladder graphs with two symmetry patterns and even n, the related orbit polynomial yields to

$$O^*(z) = 1 - \frac{n}{2} z^4. \tag{31}$$

This polynomial possesses the unique positive root

$$\delta = \sqrt[4]{\frac{2}{n}}. \tag{32}$$

For odd n, we obtain

$$O^*(z) = 1 - z^2 - \frac{n-1}{2} z^4, \tag{33}$$

and

$$\delta = \sqrt{\frac{\sqrt{2n-1} - 1}{n-1}}. \tag{34}$$

For incomplete ladders with only one symmetry pattern, we get the related orbit polynomial

$$O^*(z) = 1 - nz^2 \tag{35}$$

with

$$\delta = \frac{1}{\sqrt{n}}. \tag{36}$$

The entropy for ladder graphs with two symmetry patterns and even n yields to

$$I_a = -\frac{n}{2} \frac{4}{2n} \log\left(\frac{2}{2n}\right) = \log(n). \tag{37}$$

For ladder graphs with two symmetries and odd n, the entropy becomes to

$$I_a = -\frac{n-1}{2} \frac{4}{2n} \log\left(\frac{4}{2n}\right) - \frac{2}{2n} \log\left(\frac{2}{2n}\right) = \log(n)\frac{m-1}{n}. \tag{38}$$

Finally, for incomplete ladders with only one symmetry pattern, the entropy equals

$$I_a(z) = -n\frac{2}{2n} \log\left(\frac{2}{2n}\right) = \log(n). \tag{39}$$

These computations lead to the following analogous version of the Theorems (2.4), (2.5).

Theorem 2.6. *Let L_n be a ladder graph with $2n$ vertices. Then*

$$\lim_{n\to\infty} d(\delta(L_n), I_a(L_n)) = +\infty. \tag{40}$$

2.8. Star graphs

In this section, we investigate the limiting value of our distance measure for star graphs S_n with n vertices. The star graph has one orbit with $n-1$ vertices and $n-1$ orbits with just one vertex. So,

$$O_{S_n}(z) = z^{n-1} + z, \tag{41}$$

and

$$O^{\star}_{S_n}(z) = -z^{n-1} - z + 1 \tag{42}$$

Now, assume that $\delta < 1$ [3] satisfies the equation

$$-\delta^{n-1} - \delta + 1 = 0. \tag{43}$$

To obtain the limiting value of δ, consider the algebraic equation

$$\lim_{n\to\infty} \left(-\delta^{n-1} - \delta + 1\right) = 0. \tag{44}$$

Clearly, $-\delta + 1 = 0$ and, hence, $\delta = 1$. For $I_a(S_n)$, we obtain

$$I_a(S_n) = -\left[\frac{1}{n}\log\left(\frac{1}{n}\right) + \frac{n-1}{n}\log\left(\frac{n-1}{n}\right)\right]. \tag{45}$$

Theorem 2.7. *Let S_n be the star graph with n vertices. Then*

$$\lim_{n\to\infty} d(\delta(S_n), I_a(S_n)) = 1. \tag{46}$$

Proof: We have shown above that $\lim_{n\to\infty}\delta_n(S_n) = 1$. So,

$$\lim_{n\to\infty} d(\delta(S_n), I_a(S_n)) = \lim_{n\to\infty}\left[\delta_n(S_n) - \frac{1}{n}\log\left(\frac{1}{n}\right) - \frac{n-1}{n}\log\left(\frac{n-1}{n}\right)\right]^2 \tag{47}$$

$$= \left[\lim_{n\to\infty}\delta_n(S_n) - \lim_{n\to\infty}\frac{1}{n}\log\left(\frac{1}{n}\right) - \lim_{n\to\infty}\frac{n-1}{n}\log\left(\frac{n-1}{n}\right)\right]^2 \tag{48}$$

Noting the continuity of $()^2$, log and $0\log(0)$ and $1\log(1) = 0$, the desired result, Equation (46), follows. □

2.9. Regular Rooted Trees

Rooted trees often occur in computer science related areas [23]. In this section, we study a special case thereof namely regular rooted trees. A complete regular rooted tree of degree d and height h is a rooted tree where every inner vertex has d children and every leaf has distance h from the root vertex. Such a tree $T_{d,h}$ has $n = \frac{d^{h+1}}{d-1}$ vertices. The orbits of this tree represent the vertices on the same level, so there are $h+1$ many orbits, with $1, d^1, d^2, \ldots, d^h$ elements.

We prove the following theorem.

Theorem 2.8. *The entropy of a regular rooted tree with degree d and height h equals*

$$I_a(T_{d,h}) = \log(d^{h+1} - 1) - \log(d-1) - \log(d)\cdot\frac{hd^{h+2} - (h+1)d^{h+1} + d}{(d^{h+1}-1)(d-1)}. \tag{49}$$

For fixed degree d, the limit $\lim_{h\to\infty} I_a(T_{d,h})$ yields to

$$\lim_{h\to\infty} I_a(T_{d,h}) = \frac{d}{d-1}\log(d) - \log(d-1). \tag{50}$$

For fixed height h, the limiting value equals

$$\lim_{d\to\infty} I_a(T_{d,h}) = 0. \tag{51}$$

Proof:

$$I_a(T_{d,h}) = -\sum_{k=0}^{h} \frac{d^k}{n} \log\left(\frac{d^k}{n}\right), \tag{52}$$

$$= \frac{1}{n}\sum_{k=0}^{h} d^k \cdot (\log(n) - k\log(d)),$$

$$= \frac{\log(n)}{n}\sum_{k=0}^{h} d^k - \frac{\log(d)}{n}\sum_{k=0}^{h} kd^k,$$

$$= \log(n) - \frac{\log(d)}{n}\sum_{k=0}^{h} kd^k. \tag{53}$$

The expression $\sum_{k=0}^{h} kd^k$ is an arithmetic-geometric sequence

$$\sum_{k=0}^{h} kd^k = \frac{hd^{h+2} - (h+1)d^{h+1} + d}{(d-1)^2}. \tag{54}$$

Hence, we get

$$I_a(T_{d,h}) = \log(n) - \frac{\log(d)}{n} \cdot \frac{hd^{h+2} - (h+1)d^{h+1} + d}{(d-1)^2},$$

$$= \log\left(\frac{d^{h+1}-1}{d-1}\right) - \frac{\log(d)\cdot(d-1)}{d^{h+1}-1} \cdot \frac{hd^{h+2} - (h+1)d^{h+1} + d}{(d-1)^2},$$

$$= \log(d^{h+1}-1) - \log(d-1) - \log(d) \cdot \frac{hd^{h+2} - (h+1)d^{h+1} + d}{(d^{h+1}-1)(d-1)}. \tag{55}$$

Now we consider the limiting value of the entropy for a fixed degree d and for $h \to \infty$. In order to compute these limits, we reformulate Equation (55):

$$I_a(T_{d,h}) = \log\left(\frac{d^{h+1}-1}{d^{h+1}} \cdot d^{h+1}\right) - \log(d-1) - \log(d) \cdot \frac{h\cdot(d-1) - 1 + d^{-h}}{(1-d^{-h-1})\cdot(d-1)},$$

$$= \log\left(\frac{d^{h+1}-1}{d^{h+1}}\right) + (h+1)\cdot\log(d) - \log(d-1) - \log(d)\cdot\left(h - \frac{1-d^{-h}}{d-1}\right)\cdot\frac{1}{1-d^{-h-1}}. \tag{56}$$

Now we fix the degree d and calculate the limit $h \to \infty$; then the first term in Equation (56) converges to 0 an the last factor converges to 1. Thus

$$\lim_{h\to\infty} I_a(T_{d,h}) = \lim_{h\to\infty}\left((h+1)\cdot\log(d) - \log(d-1) - \log(d)\cdot\left(h - \frac{1}{d-1}\right)\right) \tag{57}$$

$$= \log(d) - \log(d-1) + \frac{\log(d)}{d-1} \tag{58}$$

$$= \frac{d}{d-1}\log(d) - \log(d-1) \tag{59}$$

On the other hand, if we fix the height h an consider the limit $d \to \infty$, the first summand in Equation (56) converges to 0 an the last factor converges to 1 too, hence

$$\lim_{d\to\infty} I_a(T_{d,h}) = \lim_{d\to\infty}\left((h+1)\cdot\log(d) - \log(d-1) - \log(d)\cdot\left(h - \frac{1}{d-1}\right)\right), \tag{60}$$

$$= \lim_{d\to\infty}\left((\log(d) - \log(d-1)) + \frac{\log(d)}{d-1}\right), \tag{61}$$

$$= \lim_{d\to\infty}\log\left(\frac{d}{d-1}\right) + \lim_{d\to\infty}\frac{\log(d)}{d-1} = 0. \tag{62}$$

$\square$

In what follows, we investigate δ for trees $T_{d,h}$. The polynomial $1 - O_{T_{d,h}}(z)$ equals

$$O^*_{T_{d,h}}(z) = 1 - \sum_{k_0}^{h} z^{d^k}. \tag{63}$$

These polynomials have degree > 4 (except for trivial cases) and, therefore, there is no analytical formula for their roots by radicals, see [24]. Instead of exact formulas, we give upper and lower bounds for δ.

Theorem 2.9. *Let $T_{d,h}$ be a complete regular rooted tree. It holds,*

$$\delta(T_{d,h}) < z_1^{(d)} := \frac{d}{d+1}. \tag{64}$$

A tighter upper bound is

$$\delta(T_{d,h}) < z_2^{(d)} := \frac{(d-1) \cdot d^d + (d+1)^d}{(d+1) \cdot d^d + (d+1)^d}. \tag{65}$$

Proof: For a fixed degree d, $(O^*_{T_{d,h}}(z))_{h\in\mathbb{N}}$ is a sequence of polynomials with $O^*_{T_{d,h}}(0) = 1$ for all h and for every $z > 0$ the sequence $(O^*_{T_{d,h}}(z))_{h\in\mathbb{N}})$ is decreasing. Therefore the sequence $(\delta(T_{d,h}))_{h\in\mathbb{N}}$ is decreasing too.

Hence, to obtain an upper bound of $(\delta(T_{d,h}))$ it suffices to find upper bounds for $\delta(T_{d,1})$.

$\delta(T_{d,1})$ is the positive root of the polynomial $O^*_{T_{d,1}}(z) = 1 - z - z^d$. We determine two approximations of the Newton-Iteration method [25] for the positive root of $O^*_{T_{d,1}}$. We obtain

$$z_{i+1} := z_i - \frac{O^*_{T_{d,1}}(z_i)}{{O^*_{T_{d,1}}}'(z_i)} = z_i - \frac{z^d + z - 1}{dz^{d-1} + 1}, \tag{66}$$

if starting with $z_0 = 1$. As ${O^*_{T_{d,1}}}'(z) = -1 - dz^{d-1} < 0$ and ${O^*_{T_{d,1}}}''(z) = -d(d-1)z^{d-2} < 0$ for all $z > 0$, $O^*_{T_{d,1}}(z)$ (restricted to the domain $\mathbb{R}^+$) is concave downwards. Every tangent of this function at $z > 0$ lies above the curve. If we choose z_i such that $O^*_{T_{d,1}}(z_i) < 0$, the zero of the tangent z_{i+1} is greater than $\delta(T_{d,1})$. So, if we start the Newton-Iteration method with $z_0 = 1$, we obtain the upper bounds z_1 and z_2 for $\delta(T_{d,1})$:

$$z_1 = 1 - \frac{1^d + 1 - 1}{1^{d-1} + 1} = \frac{d}{d+1}, \tag{67}$$

$$\begin{aligned} z_2 &= \frac{d}{d+1} - \frac{\left(\frac{d}{d+1}\right)^d + \frac{d}{d+1} - 1}{d\left(\frac{d}{d+1}\right)^{d-1} + 1}, \\ &= \frac{d}{d+1} - \frac{d^d + (d+1)^{d-1} \cdot d - (d+1)^d}{(d+1) \cdot d^d + (d+1)^d}, \\ &= \frac{(d-1) \cdot d^d + (d+1)^d}{(d+1) \cdot d^d + (d+1)^d}. \end{aligned} \tag{68}$$

These upper approximations z_1 and z_2 are the bounds of the Inequalities (64), (65). $\square$

We mention by determining further approximations z_i of $\delta(T_{d,1})$, we obtain even tighter but more complicated bounds.

In order to find lower bounds for $\delta(T_{d,h})$, we compare the orbit polynomial with the geometric sequences sharing the first two summands. For each orbit polynomial $O_{T_{d,h}}(z)$ and each z with $0 < z < 1$ holds

$$O_{T_{d,h}}(z) = \sum_{k=1}^{h} z^{d^k} \leq \sum_{\ell=0}^{\infty} z^{1+\ell\cdot(d-1)} := q(z) = \frac{z}{1-z^{d-1}} \tag{69}$$

The solution of $q(z) = 1$ is a lower bound of $\delta(T_{d,h})$; this solution is the positive root of

$$z^{d-1} + z - 1 = 0. \tag{70}$$

For $d > 2$ this corresponds to the value of δ of the star graph S_d. The value range of δ of the star graph has already been investigated in [3]. If we take a closer look at the just obtained bound, we see that the value of the bound represents another root of the Equation (70). When investigating bounds for the moduli of the complex zeros of complex polynomials, Dehmer et al. gave a classification of zero bounds namely *implicit* and *explicit* bounds, see [26, 27, 28]. By definition, bounds have been called implicit if the bound value is a positive zero of a concomitant polynomial obtained by direct calculation [27, 28]. In contrast, explicit bounds can be calculated by just plugging in certain parameter; note that explicit bounds are functions of the polynomial coefficients.

For $d = 2$, the zero of the polynomial represented by Equation (70) equals $z = 0.5$ (we do not mean $\delta(S_2)$ because the star graph S_2 is a degenerated star). As a consequence, we get the following theorem.

Theorem 2.10. *For a tree $T_{d,h}$, we obtain the following lower bounds for the δ:*

$$\delta(T_{d,h}) > \delta(S_d) \text{ if } d > 2 \text{ and } \delta(T_{2,h}) > 0.5. \tag{71}$$

Table (1) shows the bounds of δ for some trees.

d	$\delta(S_d)$	$\lim_{h\to\infty} \delta(T_{d,h})$	$\delta(T_{d,1})$	$z_2^{(d)}$	$z_1^{(d)}$
2	0.50000000	0.56612379	0.61803399	0.61904762	0.66666667
3	0.61803399	0.67074015	0.68232780	0.68604651	0.75000000
4	0.68232780	0.72230845	0.72449196	0.73123360	0.80000000
5	0.72449196	0.75454389	0.75487767	0.76438212	0.83333333
6	0.75487767	0.77804564	0.77808960	0.78995185	0.85714286
7	0.77808960	0.79653918	0.79654435	0.81036362	0.87500000
8	0.79654435	0.81165177	0.81165232	0.82707615	0.88888889
9	0.81165232	0.82430051	0.82430056	0.84103233	0.90000000
10	0.82430056	0.83507904	0.83507904	0.85287348	0.90909091
11	0.83507904	0.84439753	0.84439753	0.86305309	0.91666667
12	0.84439753	0.85255071	0.85255071	0.87190191	0.92307692
13	0.85255071	0.85975667	0.85975667	0.87966741	0.92857143
14	0.85975667	0.86618067	0.86618067	0.88653867	0.93333333
15	0.86618067	0.87195054	0.87195054	0.89266286	0.93750000
16	0.87195054	0.87716687	0.87716687	0.89815626	0.94117647
17	0.87716687	0.88191005	0.88191005	0.90311211	0.94444444
18	0.88191005	0.88624517	0.88624517	0.90760599	0.94736842
19	0.88624517	0.89022557	0.89022557	0.91169991	0.95000000
20	0.89022557	0.89389541	0.89389541	0.91544519	0.95238095

Table 1: Bounds for $\lim_{h\to\infty} \delta(T_{d,h})$. $\delta(S_d)$ is a lower bound. Upper bounds for $\delta(T_{d,h})$ are $z_2^{(d)}$ and $z_1^{(d)}$, see Equation (65) and Equation (64).

2.10. Alkane trees

In this section, we consider single-bond hydrocarbons (alkane) represented as trees with k carbon atoms, see [9]. This tree possesses $2k+2$ hydrogen atoms and has $n = 3k+2$ vertices [9]. Figure (1) shows an alkane tree with 5 carbon atoms represented by vertices of degree 4. We denote this graph by T_k. For even k, T_k has one orbit of size 6, $\frac{k-2}{2}$ of size 4, and $\frac{k}{2}$ of size 2, see [9]. We will discuss the case for k even with $k > 2$; the odd case can be analyzed in a similar way. The orbit polynomial of T_k has the form

$$O_{T_k}(z) = z^6 + \frac{k-2}{2} z^4 + \frac{k}{2} z^2 \tag{72}$$

and

$$O^{\star}_{T_k}(z) = -z^6 - \frac{k-2}{2} z^4 - \frac{k}{2} z^2 + 1 \tag{73}$$

From [3], we see that $O^{\star}_{T_k}(z)$ has a unique, positive root $\delta < 1$. To simplify the proof we first state and prove a lemma.

Lemma 2.11. *All zeros of the polynomial*

$$P(z) = 1 - (z^6 + \alpha_1 z^4 + \alpha_2 z^2) = -z^6 - \alpha_1 z^4 - \alpha_2 z^2 + 1 = 0, \quad \alpha_1, \alpha_2 \in \mathbb{R}, \tag{74}$$

lie in $|z| \geq \frac{1}{\sqrt{3M}}$ *where* $M := \max(|\alpha_1|, |\alpha_2|)$.

Proof: From the triangle inequality, it follows that

$$\begin{aligned} P(z)| &= |1 - (z^6 + \alpha_1 z^4 + \alpha_2 z^2)|, \\ &\geq 1 - |(z^6 + \alpha_1 z^4 + \alpha_2 z^2)|, \\ &\geq 1 - \{|z|^6 + |\alpha_1||z|^4 + |\alpha_2||z|^2\}, \\ &\geq 1 - M\{|z|^6 + |z|^4 + |z|^2\}, \\ &\geq 1 - M\{3|z|^2\} > 0. \end{aligned} \tag{75}$$

In addition, $|P(z)| > 0$ if $|z| < \frac{1}{\sqrt{3M}}$. Hence, all zeros of $P(z)$ lie in $|z| \geq \frac{1}{\sqrt{3M}}$. □

Applying Lemma (2.11) to $O^{\star}_{T_k}(z)$ (see Equation (73)), it is clear that all zeros of this polynomial lie in $|z| \geq \sqrt{\frac{2}{3}} \frac{1}{\sqrt{k}}$. Thus, $\sqrt{\frac{2}{3}} \frac{1}{\sqrt{k}} \to 0$ as $k \to \infty$.

As an example, take $k = 16$. This gives

$$O^{\star}_{T_{16}}(z) = -z^6 - 7z^4 - 8z^2 + 1, \tag{76}$$

and $\delta(T_{16}) = 0.3369$. The zero bound by Lemma (2.11) gives the region $|z| \geq \sqrt{\frac{2}{3}} \cdot \frac{1}{4} = 0.2041$.

In order to prove a result for the limiting case of T_k, we first show that the unique, positive zero of Equation (76) converges to zero as k increases without bound. We have computed the zero using the well-known Newton-Iteration method [25]. The following Table (2) illustrates the results.

From Table (2), we observe that $\delta_i \to 0$ as $k \to \infty$. The same result is shown in Figure (3).

Theorem 2.12. *Let* T_k *be the alkane tree with* k *even. Then*

$$\lim_{n \to \infty} d(\delta(T_k), I_a(T_k)) = +\infty. \tag{77}$$

Step i	k	$O^*_{T_k}(z)$	$(\delta(T_k))_i$
1	1	$-z^6 - 0.5z^4 - 0.5z^2 + 1$	1.0000000000
2	10	$-z^6 - 4z^4 - 5z^2 + 1$	0.4178030760
3	1000	$-z^6 - 499z^4 - 500z^2 + 1$	0.0446768826
4	100000	$-z^6 - 49999z^4 - 50000z^2 + 1$	0.0044720912
5	100000000	$-z^6 - 49999999z^4 - 5 \cdot 10^7 z^2 + 1$	0.0001414214
6	10000000000	$-z^6 - 4999999999z^4 - 5 \cdot 10^9 z^2 + 1$	0.0000141421

Table 2: The values of δ_i for different k.

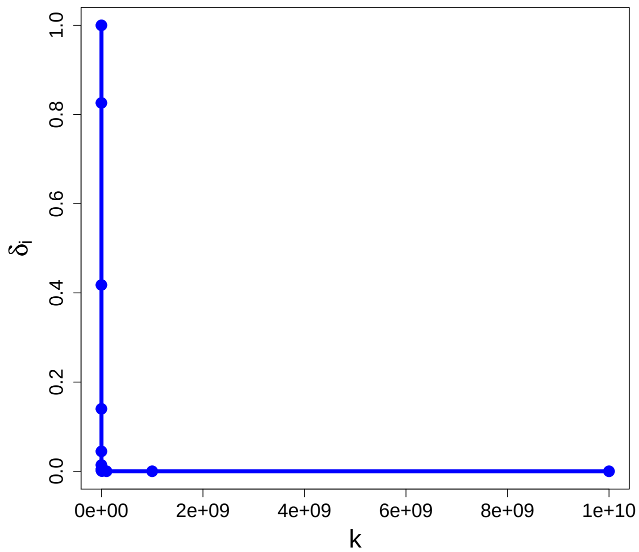

Figure 3: δ_i for different k.

Proof: First we compute $I(T_k)$. Appealing to the orbit structure of T_k above, we obtain

$$\begin{aligned} I(T_k) &= -\left[\frac{6}{3k+1}\log\left(\frac{6}{3k+1}\right) + \frac{k-2}{2}\cdot\frac{4}{3k+1}\log\left(\frac{4}{3k+1}\right)\right. \\ &\quad \left. + \frac{k}{2}\cdot\frac{2}{3k+1}\log\left(\frac{2}{3k+1}\right)\right], \qquad (78) \\ &= -\left[\frac{6}{3k+1}\log\left(\frac{6}{3k+1}\right) + \frac{2k-4}{3k+1}\log\left(\frac{4}{3k+1}\right)\right. \\ &\quad \left. + \frac{k}{3k+1}\log\left(\frac{2}{3k+1}\right)\right]. \qquad (79) \end{aligned}$$

Taking $\delta(T_k)$ as the unique, positive zero of $O^\star_{T_k}(z) = 0$, we make use of the observation $\lim_{k\to\infty} \delta(T_k) =$

0. Thus,

$$\lim_{k\to\infty} d(\delta(T_k), I_a(T_k)) = \lim_{k\to\infty}\Bigg[\delta(T_k) - \frac{6}{3k+1}\log\left(\frac{6}{3k+1}\right) + \frac{2k-4}{3k+1}\log\left(\frac{4}{3k+1}\right)$$

$$+\frac{k}{3k+1}\log\left(\frac{2}{3k+1}\right)\Bigg]^2, \tag{80}$$

$$= \Bigg[\lim_{k\to\infty}\delta(T_k) - \lim_{k\to\infty}\frac{6}{3k+1}\log\left(\frac{6}{3k+1}\right) - \lim_{k\to\infty}\frac{2k-4}{3k+1}\log\left(\frac{4}{3k+1}\right)$$

$$-\lim_{k\to\infty}\frac{k}{3k+1}\log\left(\frac{2}{3k+1}\right)\Bigg]^2. \tag{81}$$

In conclusion, we again use the fact that the functions $()^2$ and log are continuous. Noting that $\lim_{k\to\infty}\frac{2k-4}{3k+1} = \frac{2}{3}$ and $\lim_{k\to\infty}\frac{k}{3k+1} = \frac{1}{3}$ and, moreover, $0\cdot\log(0) = 0$ and $\log(0) = -\infty$, we finally obtain from Equation (81) that the limiting values given by Equation (77) hold. □

It is evident that the statement of Theorem (2.12) is also valid for odd k and can be proven analogously.

2.11. Complete Bipartite Graphs

In this section, we consider complete bipartite graphs K_{n_1,n_2} with two orbits of order n_1 and n_2. We assume that $n_1 \neq n_2$, otherwise bipartite graphs are transitive and these graphs have already been tackled in Section (2.4). A special case of complete bipartite graphs are star graphs that we already analyzed in Section (2.8). We start by discussing the graph measure δ of these graphs. We state the following lemma.

Lemma 2.13. *For every $\varepsilon > 0$ there only exist a finite number of pairs $(n_1, n_2) \in \dot{\mathbb{N}}^2$ with $\delta(K_{n_1,n_2}) < 1-\varepsilon$.*

Proof: Assume that $\delta := \delta(K_{n_1,n_2}) < 1-\varepsilon$. By definition of the related orbit polynomial, we obtain $\delta^{n_1} + \delta^{n_2} = 1$. Without loss of generality we assume $n_1 < n_2$. $\delta < 1$ implies $\delta^{n_1} > \delta^{n_2}$. So, $\frac{1}{2} < \delta^{n_1} < (1-\varepsilon)^{n_1}$ and $n_1 \leq \frac{\log(1/2)}{\log(1-\varepsilon)}$.

Moreover, as $\delta < 1-\varepsilon$, we also conclude $\delta^{n_1} < 1-\varepsilon$. Therefore, $\varepsilon \leq \delta^{n_2} \leq (1-\varepsilon)^{n_2}$, and finally $n_2 \leq \frac{\log(\varepsilon)}{\log(1-\varepsilon)}$. □

The computation of the entropy of a complete bipartite graph K_{n_1,n_2} is straightforward.

$$I_a(K_{n_1,n_2}) = p_1\log(p_1) + p_2\log(p_2), \tag{82}$$
$$\text{where } p_1 = \frac{n_1}{n_1+n_2} \text{ and } p_2 = \frac{n_2}{n_1+n_2}.$$

As a consequence of Lemma (2.13) and Equation (82), we obtain the next statement.

Theorem 2.14.

1. *For a fixed number n_1, the limit of the distances of the entropy and δ on bipartite graphs is*

$$\lim_{n_2\to\infty} d(\delta(K_{n_1,n_2}), I_a(K_{n_1,n_2})) = 1. \tag{83}$$

2. *If $(n_1^{(i)}, n_2^{(i)})$ is a sequence of pairs with $\lim_{i\to\infty} n_1^{(i)} = \lim_{i\to\infty} n_2^{(i)} = \infty$ and if $p_1 := \lim_{i\to\infty}\frac{n_1^{(i)}}{n_1^i+n_2^i}$ and $p_2 := \lim_{i\to\infty}\frac{n_2^{(i)}}{n_2^i+n_2^i}$ exist, then*

$$\lim_{n_2\to\infty} d(\delta(K_{n_1,n_2}), I_a(K_{n_1,n_2})) = (p_1\log(p_1) + p_2\log(p_2) - 1)^2. \tag{84}$$

3. Summary and Conclusion

This paper has initiated a comparative analysis of alternative quantitative measures of graph symmetry. Two measures based on the automorphism group of a graph were considered. For purposes of comparison a real valued distance measure was introduced, and the limiting values of the respective distances between the two symmetry measures on each of several classes of graphs were established. Although both symmetry measures compared are based on the same variables, namely the number and sizes of the orbits of the automorphism group, there can be considerable, and even extreme differences between them. For example, the distance between the two symmetry measures on the path graph diverges to infinity. An appropriate analogy is the case of two functions on the same set of variables having very different functional forms. The current research is viewed as a first step in a systematic analysis of symmetry measures. Distances between the two symmetry measures on additional classes of graphs, as well as comparative analysis of other measures, will be undertaken in future work.

4. Acknowledgments

Mattins Dehmer thanks the Austrian Science Funds for supporting this work (project P30031).

References

[1] D. Bonchev, Information Theoretic Indices for Characterization of Chemical Structures, Research Studies Press, Chichester, 1983.

[2] S. C. Basak, A. T. Balaban, G. D. Grunwald, B. D. Gute, Topological Indices: Their Nature and Mutual Relatedness, J. Chem. Inf. Comput. Sci. 40 (2000) 891–898.

[3] M. Dehmer, Z. Chen, F. Emmert-Streib, A. Mowshowitz, K. Varmuza, L. Feng, H. Jodlbauer, Y. Shi, J. Tao, The Orbit-Polynomial: A Novel Measure of Symmetry in Networks, IEEE Access 8 (2020) 36100–36112, ISSN 2169-3536, doi:\bibinfo{doi}{10.1109/ACCESS.2020.2970059}.

[4] A. Garrido, Symmetry of Complex Networks, Advanced Modeling and Optimization 11 (4).

[5] A. Mowshowitz, Entropy and the complexity of the graphs I: An index of the relative complexity of a graph, Bull. Math. Biophys. 30 (1968) 175–204.

[6] S.-J. Chen, K. A. Dill, Symmetries in proteins: A knot theory approach, J. Chem. Phys. 104 (1996) 5964–5973.

[7] C. J. Colbourn, K. S. Booth, Linear Time Automorphism Algorithms for Trees, Interval Graphs, and Planar Graphs, SIAM Journal on Computing 10 (1981) 203–225.

[8] M. Razinger, K. Balasubramanian, M. E. Munk, Graph automorphism perception algorithms in computer-enhanced structure elucidation, J. Chem. Inf. Comput. Sci. 33 (1993) 197–201.

[9] A. Mowshowitz, M. Dehmer, A Symmetry Index for Graphs, Symmetry: Culture and Science 21 (4) (2010) 321–327.

[10] A. Mowshowitz, Entropy and the complexity of graphs II: The information content of digraphs and infinite graphs, Bull. Math. Biophys. 30 (1968) 225–240.

[11] A. Mowshowitz, Entropy and the complexity of graphs III: Graphs with prescribed information content, Bull. Math. Biophys. 30 (1968) 387–414.

[12] A. Mowshowitz, Entropy and the complexity of graphs IV: Entropy measures and graphical structure, Bull. Math. Biophys. 30 (1968) 533–546.

[13] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors, Wiley-VCH, Weinheim, Germany, 2002.

[14] B. D. MacArthur, R. J. Sánchez-García, J. W. Anderson, Symmetry in complex networks, Discrete Applied Mathematics 156 (18) (2008) 3525–3531.

[15] F. Ball, A. Geyer-Schulz, How Symmetric Are Real-World Graphs? A Large-Scale Study, Symmetry 10 (1).

[16] Z. Chen, M. Dehmer, F. Emmert-Streib, A. Mowshowitz, Y. Shi, Toward Measuring Network Aesthetics Based on Symmetry, Axioms 6.

[17] F. Emmert-Streib, M. Dehmer, Networks for Systems Biology: Conceptual Connection of Data and Function, IET Systems Biology 5 (2011) 185–207.

[18] M. Dehmer, A. Mowshowitz, A History of Graph Entropy Measures, Information Sciences 1 (2011) 57–78.

[19] C. Godsil, G. Royle, Algebraic Graph Theory, Graduate Texts in Mathematics, Academic Press, 2001.

[20] H. H. Bock, Automatische Klassifikation. Theoretische und praktische Methoden zur Gruppierung und Strukturierung von Daten, Studia Mathematica, Vandenhoeck & Ruprecht, Göttingen, 1974.

[21] M. M. Deza, E. Deza, Encyclopedia of Distances, Springer; 2nd ed., 2012.

[22] R. G. Bartle, D. R. Sherbert, Introduction to Real Analysis, Wiley Publishing, 1999.

[23] T. H. Cormen, C. E. Leiserson, R. L. Rivest, Introduction to Algorithms, MIT Press, 1990.

[24] J. P. Tignol, Galois' Theory of Algebraic Equations, World Scientific Publishing Company, 2016.

[25] A. S. Householder, The Numerical Treatment of a single Nonlinear Equation, McGraw-Hill, new York, NY, USA, 1970.

[26] M. Dehmer, On the location of zeros of complex polynomials, Journal of Inequalities in Pure and Applied Mathematics .

[27] M. Dehmer, J. Kilian, On Bounds For The Zeros of Univariate Polynomials, International Journal of Applied Mathematics and Commuter Science 4 (2) (2007) 118–123.

[28] M. Dehmer, A. Mowshowitz, Bounds on the Moduli of Polynomial Zeros, Applied Mathematics and Computation 218 (2011) 4128–4137.