Applied Mathematics and Computation 320 (2018) 142–148

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

A calculus for measuring the elegance of abstract graphs

Abbe Mowshowitz[a,*], Matthias Dehmer[b,c]

[a] Department of Computer Science, The City College of New York (CUNY), 138th Street at Convent Avenue, New York, NY 10031, USA
[b] Department of Biomedical Computer Science and Mechatronics, UMIT, Hall in Tyrol, Austria
[c] College of Computer and Control Engineering, Nankai University, Tianjin 300350, PR China

ARTICLE INFO

Keywords:
Graph elegance
Graph aesthetics
Graph construction
Entropy
Information content of graphs
Hydrocarbon trees

ABSTRACT

This paper introduces a system for measuring the elegance of a graph based on the steps needed to build the graph and on its symmetry structure. The measure is designed to capture the essence of the notion of elegance in mathematics, namely, simplicity and clarity. The term "elegance" is used instead of "aesthetics" to distinguish the measure from those dependent on visual representation of a graph. Elegance is based solely on the abstract properties of a graph. A framework for measurement is defined and applied in a special case.

1. Introduction

The notion of elegance in mathematics, although well understood by practicing mathematicians, is rarely defined explicitly. This paper aims to extract the essence of the largely intuitive idea of elegance and apply it to an abstract mathematical object, namely, a graph. The aesthetics of graphs has been treated primarily from the perspective of visual representations, which makes use of the geometrical properties of drawings; but relatively little has been written about the aesthetics of a graph as a purely abstract object, specified solely in terms of vertices and edges [3,9,29,35]. We use the term "elegance" rather than "aesthetics" to avoid confusion, since the measure introduced here is independent of geometry. In keeping with the intuitive idea, elegant graphs are those which can be constructed easily from a set of graph primitives and exhibit a high degree of symmetry. The two principles underlying our definition of elegance are *least effort* and *simplicity*. A graph is relatively elegant if it requires minimal effort to construct and is simple in structure.

Clearly, this notion of elegance is dependent upon the choice of primitives. However, there are constraints so the choice is not arbitrary. If one wants to be able to build all graphs of a certain type, care must be taken to use a set of primitives that will capture all such graphs. For example, all trees could be constructed from an edge as primitive, but only some trees could be built from a path of length two or more. Of course, one might include include both an edge and a path of length two as primitives. The problem of selecting primitives is analogous to the situation for the propositional calculus, where there is a trade off between the size and simplicity of the set of primitives, on the one hand, and the length of proofs of validity, on the other hand. In the context of graph elegance, the number of construction steps can sometimes be reduced by enlarging the set of primitives. This fact underscores the relativity of measurement and points up the need to consider the requirements of a particular application.

The use of symmetry in the definition of graph elegance might be challenged on grounds that almost all graphs have the trivial automorphism group [14]. However, the term "almost all" is misleading in the present context. There are sufficiently

* Corresponding author.
E-mail addresses: abbe@cs.ccny.cuny.edu, abbemow@earthlink.net (A. Mowshowitz).

https://doi.org/10.1016/j.amc.2017.09.023
0096-3003/

many finite graphs arising in applications in chemistry, psychology and other fields that have non-trivial automorphism group to justify the use of symmetry as part of the definition of elegance.

The elements of a calculus for quantitative measurement of elegance are defined in Section 2. This is followed in Section 3 by an example of elegance measurement applied to the class of graphs representing hydrocarbons. Elegance measurement is compared to other ways of investigating graph structure in Section 4. Extensions of the hydrocarbon example are presented in Section 5 and the results are discussed in the conclusion.

2. The calculus framework

Two components make up our measure of *graph elegance*: (1) the number of steps needed to construct the graph from a given set of (primitive) graphs using a specified set of operations, and (2) the structural information content of the graph. Let $P_1, \ldots, P_r$ be a set of graphs (primitives), and $\alpha_1, \ldots, \alpha_s$ a set of binary graph operations, and U the construction universe. U consists of the primitives and all graphs constructed using the primitives and already constructed graphs. A sequence of graphs $[G_0^n]$ is a *defining sequence* for a graph G if

$$G_0 = P_k, 1 \leq k \leq r \tag{1}$$

$$G_t = G_{t-1}\alpha_i P_j, t > 0, 1 \leq i \leq s, 1 \leq j \leq r, \quad \text{or} \tag{2}$$

$$G_t = G_{t-1}\alpha_i H, t > 0, 1 \leq i \leq r, H \in U, \tag{3}$$

and $G = G_n, n \geq 0$. The number n of operations α in the defining sequence is the *build number* of G denoted by $b(G)$. In addition to the build number, the structural information content of G is needed to compute the elegance index $e(G)$. Following [24–27] the information content of G is $I_{\text{Aut}}(G) = -\sum_{i=1}^{m} p_i \log(p_i)$ where p_i is the size of the ith orbit of the automorphism group of G divided by the number of vertices in G (see [27]). Finally, the elegance of G is obtained by adding its information content to the minimum build number over all possible defining sequences, i.e., $e(G) = I_{\text{Aut}}(G) + \min[G_0^n]b(G)$.

Finding appropriate primitives and operations is a nontrivial exercise, dependent upon the class of graphs one wishes to construct as well as the aims of measurement. Clearly, there are many possible sets of primitives and operations that could be used, making a general system applicable to arbitrary graphs difficult (but not impossible) to justify. Such a system must await results concerning the optimality of construction over different systems. Note that a graph may have several defining sequences, and non-isomorphic graphs may have the same build number and the same elegance value. Since elegance equates with ease of construction and simplicity, the smaller the quantitative value, the more elegant the graph. The following section illustrates the measurement system for a class of graphs representing simple hydrocarbons.

3. Application to hydrocarbon trees

Hydrocarbons can be modeled as trees in which every vertex is of degree 4 or 1. Vertices of degree 1 are called endpoints. Vertices of degree 4 correspond to carbon atoms; the endpoints correspond to hydrogen atoms. If k denotes the number of vertices of degree 4, the number of endpoints is $2k+2$, so the number n of vertices in the tree is given by $n = 3k+2$. To illustrate elegance measurement we adopt the following system for constructing the hydrocarbon trees (H-trees). Only one primitive is required, namely, the star S_5 with five vertices. Any H-tree can be constructed using this primitive in the following manner. If T is an H-tree with n vertices, join S_5 or a tree $S \in U$ to T by identifying an edge of T with an edge of the other tree, each of the two edges being incident to an endpoint. The resulting tree is again an H-tree. Note that the operation defined for H-trees depends on two edges, one in each of the operands as shown in Fig. 1.

The H-trees in Fig. 2 represent butane and isobutane. The H-tree for butane has build number 2, whereas that for isobutane has build number 3. The former can be constructed in two steps: joining S_5 with itself, and then joining the result with itself; the latter requires three steps: joining S_5 with itself, and then joining the result with S_5. Butane has four orbits containing 6,4,2,2 vertices, respectively, and its structural information content is 1.841; the orbit sizes of isobutane are 9,3,1,1, giving it a structural information content of 1.609. So, butane has an elegance value of 3.841, whereas isobutane's value is 4.609. Thus the butane H-tree can be said to be more elegant than the H-tree of isobutane. This difference agrees with the fact that isobutane is more expensive to produce than butane. This example illustrates the need to take account of two factors in measuring elegance, since the H-tree for butane is more complex than that for isobutane, but the latter requires an additional step to build, giving it a higher elegance value. In general, we might conjecture that more elegant molecules (i.e., those having relatively low elegance values) are less expensive to synthesize. In the H-tree model, the increase in effort required to build a graph with high elegance value comes from having to locate particular edges to identify among the choices available in the operand graphs. Building the simple chain of the butane tree requires little exploration among the graphs to be joined compared with the search effort needed to build the more symmetric tree representing isobutane.

Elegance values for trees corresponding to acyclic saturated hydrocarbons (i.e., simple chains of carbon atoms) are given in the following theorem.