

---

# Graphs with Identical Chromatic Symmetric Functions

---

## A PROPOSAL FOR THE UNDERGRADUATE RESEARCH AWARD

*Author:*

Keeler M. Russell

*Mentor:*

Dr. Jeremy L. Martin



UNIVERSITY OF KANSAS

May 21, 2012

# Project Description

---

## Summary of Purpose

All graphs, sets of vertices joined by edges, have an associated chromatic symmetric function  $X_G$ . This project focuses on whether or not a tree, a particular type of graph, can be reconstructed uniquely from its chromatic symmetric function. This is an open, or unanswered, question in mathematics, defined more fully throughout the course of this proposal.

## Background Information

A *graph*  $G$ , denoted  $G(V, E)$ , is a set of vertices  $V$  connected by a set of edges  $E$ . The *order* of a graph is the number of vertices in  $V$ . A *tree*  $T$  is a graph such that any two vertices are connected by a unique *simple path*, a sequence of distinct vertices such that each vertex is connected to the next in the sequence by a single edge. An excellent entry-level introduction to graph theory is [1].

We say that two graphs are *isomorphic* if they are structurally identical. Also, an *invariant* is a property of a graph  $G$  that depends only on the structure of the  $G$ , and not on a particular representation or drawing of  $G$ . For example, the number of vertices in a graph or number of edges in a graph are simple invariants. Isomorphic graphs, since they are structurally identical, must have the same values for their invariants. However, note that the reverse is not true in general: two non-isomorphic graphs can have the same values for invariants. For example, consider Figure 1: both graphs have the same number of vertices and edges, yet they are clearly unique structurally. A *complete invariant* is one that is unique to an isomorphic graph; that is, only that graph and no others can have that value for the invariant, thus enabling reconstruction of that graph from its invariant.

A *proper coloring* is an assignment of colors to the vertices of a graph such that no two adjacent vertices (vertices which share an edge) share a color. To count the proper colorings of a graph  $G$ , we use the *chromatic polynomial*  $\chi_G(k)$ , which counts the number of proper colorings of  $G$  given at most  $k$  colors. This is another type of invariant. For any tree  $T$ ,  $\chi_T(k) = k(k-1)^{n-1}$ , where  $n$  is the number of vertices in the tree. Clearly the chromatic polynomial is useless in distinguishing trees because it gives the same information regardless of the structure of the tree. Building on the chromatic polynomial, Richard Stanley introduced yet another invariant, the *chromatic symmetric function*  $X_G$ , as a symmetric function generalization of the chromatic polynomial [5]. The chromatic symmetric function  $X_G$  of a graph  $G$  is an ordinary generating function which counts, for each integer partition  $\lambda \vdash n$ , the number of proper colorings that exist using the color  $i$   $\lambda_i$  times. For example, in Figure 1, the term  $4m_{2,2,1}$  in  $X_G$  indicates that there are 4 ways to color two vertices one color, two other vertices a second color, and one final vertex a third color. A standard reference for symmetric functions is [6]. For reference, we denote the chromatic symmetric function of a general graph  $G$  is denoted as  $X_G$ , that of a tree  $T$  as  $X_T$ , and occasionally abbreviate ‘chromatic symmetric function’ as ‘CSF’.

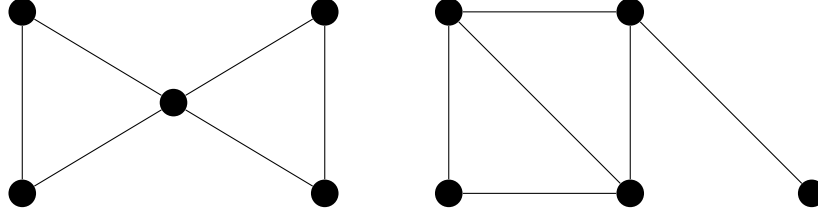


Figure 1: Two graphs  $G$  and  $H$  with the same chromatic symmetric function. In this case,  $X_G = X_H = 120m_{1,1,1,1,1} + 24m_{2,1,1,1} + 4m_{2,2,1}$ .

## Research Context

Here, we may state the overall research question concretely: is  $X_T$  a complete isomorphism invariant for trees? That is, does  $X_T$  determine a tree's structure up to isomorphism? Or, equivalently, is it possible for any two non-isomorphic trees to have the same  $X_T$ ?

Firstly, it *is* possible for two unique graphs to have the same  $X_G$ , so it is not a complete isomorphism invariant for graphs in general (see Figure 1). However, it is unknown whether  $X_T$  is a complete isomorphism invariant for trees, which is the question this research aims to tackle. For trees with 23 or less vertices, this question has been answered affirmatively by Li-Yang Tan [8]. Through exhaustive computation, Tan showed that no trees of 23 vertices or less have the same  $X_T$ .

Additionally, Dr. Martin, et. al. have shown that certain important graph invariants can be recovered from  $X_G$  of graphs in general and that for certain classes of trees,  $X_T$  distinguishes them up to isomorphism [4]. This is the state of the art for the chromatic symmetric function.

## Significance

This proposal primarily focuses on pure math, an area of research that can be difficult to justify given its lack of direct, real-world significance at conception; the application to the real world often comes later. However, graphs (and especially trees), are utilized to incredible effect in the domain of computer science. Many efficient algorithms and data structures rely heavily on trees as their backbone. I wish to push the frontier of what is known about the chromatic symmetric function of a tree beyond the current state of the art, potentially discerning an answer to Stanley's open question of whether  $X_T$  is a complete isomorphism invariant for trees.

## Aim of Research

The first goal is computationally determine whether trees of greater than 23 vertices have the same  $X_T$ . If this is accomplished, a pair of trees forming a counterexample for the open question given above may be found, and this pair can be analyzed to determine exactly *why* they each have different  $X_T$ . If such a counterexample is not found, then at the very least the boundary of 23 vertices will have been breached, and the code that I will write for these computations can be made freely available online for future research.

The second goal is purely mathematical: to construct an infinite family of pairs of graphs  $(G_0, H_0)$ ,  $(G_1, H_1)$ ,  $(G_2, H_2)$ , ... for which  $X(G_i)$  is identical to  $X(H_i)$  for all

natural numbers  $i$ . Stanley, in his original paper introducing  $X_G$ , presented a pair of graphs which have the same  $X_G$ , one which looks like a bowtie, the other something like a kite (see Figure 1). By ‘gluing’ copies of these graphs together in a clever way, a method could be provided for constructing an infinite number of graphs with the same chromatic symmetric function. This resembles Gordon and Eisenstat’s construction of an infinite family of non-isomorphic ‘caterpillar’ trees with identical subtree data [3]. Although it would not answer the open question for trees, it would represent an opposing line of inquiry which examines *why* this infinite family has pairs of graphs with the same  $X_G$ . This could provide insight into what sort of qualities make one graph’s  $X_G$  differ from another’s.

The third, most abstract, and most difficult goal is to treat the chromatic symmetric function as a linear transformation on the vector space of graphs of order  $n$  to the vector space of symmetric functions homogeneous of degree  $n$ . This transformation must have a *kernel* or *nullspace*, a set that contains any of the transformations that result in 0, since the dimension of the vector space of graphs of order  $n$  is much larger than the dimension of the vector space of symmetric functions of degree  $n$ . The next question to ask would be, “What kind of things exist in this kernel?” Given two graphs  $G$  and  $H$  of the same order, if  $X_G = X_H$  then the formal linear combination  $G - H$  is in the kernel. By understanding the kernel of this transformation, it may be possible to gain new insight into the open question of the chromatic symmetric functions of trees. Note that if a tree can be found in the kernel, then two trees share the same CSF, thus answering the question.

## Methods

The first step is to develop a program that generates all non-isomorphic trees on  $n$  vertices, then check to make sure that none of them have the same  $X_T$ . My primary tool in this investigation will be *Sage*, an open-source mathematics software package [7]. For graphs in general, the isomorphism problem becomes infeasible: the number of graphs on  $n$  vertices grows exponentially, and a fast algorithm for this process is not known. However, for trees, there exist efficient algorithms for enumerating all trees on  $n$  vertices [9] and for checking for isomorphism between two trees [2]. To compute  $X_T$  for a given tree  $T$ , Stanley gives a formula in [5] which can serve as the basis for the code.

Constructing an infinite family of trees with identical  $X_T$  involves more intuition than programming. Dr. Martin has confirmed this goal’s feasibility. I can investigate Stanley’s original paper on the chromatic symmetric function further for insight into this problem. Additionally, Dr. Martin’s own work serves as an excellent reference for the combinatorial interpretation of the coefficients of  $X_G$ . With some effort and fact-checking using *Sage*, I should be able to construct such a family.

The third and final goal is certainly nebulous. At the present time, I believe it is beyond my skill level, and Dr. Martin is not sure whether the approach will turn up anything useful, but that is exactly what makes the research so exciting! In the process of completing the first and second goals of this project, I am confident that I will pick up enough intuition about the chromatic symmetric function to make a reasonable stab at this problem. It is a more generalized version of the second goal, and so the same methods as above will apply.

## Significance to Me

---

First and foremost, this project is exciting for me. I have already begun to delve in without the grant, and will continue to delve further if not awarded the grant. I plan on continuing my education by acquiring a masters degree, with Computer Science and Mathematics as my two primary choices. As such, I feel that this project is an excellent preparation for graduate study, allowing me to try pushing against the limits of human knowledge in these fields simultaneously before even reaching graduate school.

I would love to see my research published, as I am sure any undergraduate would. More likely, though, I would present at a conference. Dr Martin informed me that many mathematics conferences love to see undergraduate efforts presented. In fact, he mentioned one in particular, the MAA (Mathematical Association of America) conference in San Diego, CA during January 2013. If I planned to present at this conference, that would allow me approximately a year to conduct my investigation.

## My Qualifications

---

This proposal requires knowledge of computing techniques to develop a correct and efficient algorithm for generating trees and checking their chromatic symmetric functions against one another. A brute force algorithm is out of the question, as such an algorithm would be entirely unfeasible given the sheer number of trees it would have to grind through. I feel that I am up to the task for several reasons: for one, I am a Computer Engineering student, and have formal training in algorithm development and analysis. Secondly, I make a living as a programmer working for Garmin on embedded GPS device code, meaning that I have programmed nearly every day for several years. Finally, I was awarded the Eta Kappa Nu Underclassman Achievement Award for being the top student in Programming II.

Additionally, during the summer of 2010, just after my freshman semester, I participated in the Honors Research Development Program, drafting a thesis with Dr. Martin. My current proposal is based on the work I completed during that summer, including reading research papers, learning elementary graph theory, and drafting academic papers using  $\text{\LaTeX}$ . Dr. Martin, however, became busy with classes and several graduate students soon afterward, and I had just started working for Garmin and researching briefly with the Bioinformatics and Computational Life Sciences Laboratory on West Campus. Consequently, the project fell by the wayside. During the interim since then, I have completed more mathematics courses, making me feel much more prepared to tackle a research project like this. With Dr. Martin on sabbatical this semester supporting my efforts and my greater understanding of the tasks at hand, I feel strongly qualified for this project and this research award.

## References

- [1] Miklos Bona. *A Walk Through Combinatorics: An Introduction to Enumeration and Graph Theory*. World Scientific Publishing Company, 2002.
- [2] Samuel Buss. Alogtime algorithms for tree isomorphism, comparison, and canonization. In *In Computational Logic and Proof Theory, 5th Kurt Godel Colloquium'97, Lecture Notes in Computer Science No. 1289*, pages 18–33. Springer-Verlag, 1997.
- [3] David Eisenstat and Gary Gordon. Non-isomorphic caterpillars with identical subtree data. *Discrete Math*, 306:827–830, 2006.
- [4] Jeremy L. Martin, Matthew Morin, and Jennifer D. Wagner. On distinguishing trees by their chromatic symmetric functions. *Journal of Combinatorial Theory Series A*, 115(2):237–253, 2008.
- [5] Richard P. Stanley. A symmetric function generalization of the chromatic polynomial of a graph. *Advances in Mathematics*, 111(1):166–194, 1995.
- [6] Richard P. Stanley. *Enumerative Combinatorics, Volume 2*. Cambridge University Press, 1999.
- [7] William Stein. Sage open-source mathematics software. <http://www.sagemath.org>.
- [8] Li-Yang Tan. Computational verification that the chromatic symmetric function determines up to isomorphism trees of order at most 23. <http://www.cs.wustl.edu/~lt1/csf.html>, 2007.
- [9] Robert Alan Wright, L. Bruce Richmond, Andrew M. Odlyzko, and Brendan D. McKay. Constant time generation of free trees. *SIAM J. Comput.*, pages 540–548, 1986.