Non-isomorphic caterpillars with identical subtree data

David Eisenstat\textsuperscript{a}, Gary Gordon\textsuperscript{b}

\textsuperscript{a}University of Rochester, Rochester, NY 14627-6602, USA
\textsuperscript{b}Department of Mathematics, Lafayette College, Easton, PA 18042-1781, USA

Received 3 March 2005; received in revised form 20 January 2006; accepted 25 January 2006
Available online 4 April 2006

Abstract

The greedoid Tutte polynomial of a tree is equivalent to a generating function that encodes information about the number of subtrees with \( I \) internal (non-leaf) edges and \( L \) leaf edges, for all \( I \) and \( L \). We prove that this information does not uniquely determine the tree \( T \) by constructing an infinite family of pairs of non-isomorphic caterpillars, each pair having identical subtree data. This disproves conjectures of [S. Chaudhary, G. Gordon, Tutte polynomials for trees, J. Graph Theory 15 (1991) 317–331] and [G. Gordon, E. McDonnell, D. Orloff, N. Young, On the Tutte polynomial of a tree, Congr. Numer. 108 (1995) 141–151] and contrasts with the situation for rooted trees, where this data completely determines the rooted tree.

Keywords: Tree; Subtree; Greedoid Tutte polynomial

1. Introduction

When \( T \) is a \textit{rooted} tree, the greedoid Tutte polynomial \( f(T) \) uniquely determines \( T \) [7, Theorem 2.8]. In this note we show that this result does not extend to unrooted trees: we construct an infinite collection of pairs of non-isomorphic \textit{caterpillars} (trees in which all of the non-leaf vertices form a path), each pair having the same greedoid Tutte polynomial (Corollary 2.7). This extends a construction in [5], where caterpillars with the same degree sequence and path data are created using a generating function approach.

From a combinatorial perspective, this greedoid Tutte polynomial encodes data about the number of subtrees of the tree with \( I \) internal (non-leaf) edges and \( L \) leaf edges. In fact, the greedoid Tutte polynomial is equivalent to a two-variable generating function \( \sum S x^{I(S)} y^{L(S)} \), where the subtree \( S \) has \( I(S) \) non-leaf and \( L(S) \) leaf edges and the sum extends over all subtrees. Thus, our main result (Theorem 2.6) can be stated purely combinatorially:

\textit{Main result:} Let \( c(T; I, L) \) denote the number of subtrees of the tree \( T \) having exactly \( I \) internal edges and \( L \) leaf edges. Then there exist infinitely many pairs of non-isomorphic trees \( T_1 \) and \( T_2 \) such that \( c(T_1; I, L) = c(T_2; I, L) \) for all \( I \) and \( L \).

Attempts to reconstruct graphs or matroids from polynomials have been attempted before. See [3,4] for classes of matroids and graphs for which unique reconstruction is possible. Note that de Mier and Noy consider the standard Tutte polynomial; we use a \textit{greedoid} version of this invariant (see remarks following Definition 2.2).

\* This work was supported by NSF Grant DMS-0243763.
E-mail addresses: de001m@mail.rochester.edu (D. Eisenstat), gordong@lafayette.edu (G. Gordon).

0012-365X/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.disc.2006.01.022
2. The counterexamples

Let $T$ be a tree with edge set $E$, where $|E| = n$. We define the rank of a subset of edges as follows:

**Definition 2.1.** For $A \subseteq E$, the rank of $A$ is given by

$$r(A) = \max_{F \subseteq A} \{|F| : F \text{ is the complement of a subtree of } T\}.$$ 

This rank function is the pruning rank of the tree; for $A \subseteq E$, we have $r(A)$ is the largest number of edges in $A$ which can be pruned from $A$, where the pruning process removes leaves, one by one, until no more leaves remain. During this process, edges that are not leaves initially (and hence, cannot be pruned initially) may become available for pruning later.

Definition 2.1 gives the tree $T$ an antimatroid structure, but we will not need this generality here. However, we do point out that the antimatroid structure completely determines the tree; in particular, it is possible to uniquely reconstruct the tree from the rank function of the antimatroid [1, Corollary 3.5]. Thus, the counterexamples given in this section provide examples of non-isomorphic antimatroids sharing the same greedoid Tutte polynomial.

**Definition 2.2.** Let $T$ be a tree with rank function as in Definition 2.1. Then the greedoid Tutte polynomial is defined by

$$f(T; t, z) = \sum_{A \subseteq E} t^{n-r(A)} z^{|A| - r(A)}.$$ 

This definition gives the standard Tutte polynomial of a graph (more precisely, the Whitney corank-nullity polynomial) when we use the cycle-rank function (i.e., $r(A)$ is the size of the largest cycle-free subset of $A$). To avoid confusion, we refer to the polynomial considered here as the greedoid Tutte polynomial throughout this note. We point out that the standard Tutte polynomial of a tree is simply $(t + 1)^{|E|}$, and so is of (essentially) no value in this situation. More information about the connection between these invariants can be found in [7].

We will need a combinatorial reformulation of the greedoid Tutte polynomial as a generating function.

**Definition 2.3.** Let $T$ be a tree. Then the subtree leaf–non-leaf generating function is defined by

$$g(T; x, y) = \sum_S x^{I(S)} y^{L(S)},$$ 

where the sum extends over all subtrees of $T$.

The connection between the greedoid Tutte polynomial and the generating function $g(T; x, y)$ is given in Proposition 13(b) of [2].

**Proposition 2.4.** Let $T$ be a tree with greedoid Tutte polynomial $f(T; t, z)$, and let $g(T; x, y) = \sum_S x^{I(S)} y^{L(S)}$, where the sum extends over all subtrees of $T$. Then

$$g(T; x, y) = f(T; y, xy^{-1} - 1),$$ 

$$f(T; t, z) = g(T; t(z + 1), t - 1).$$

We will use the following notation. Let $T$ be a caterpillar with spine $\{v_1, \ldots, v_r\}$, and let $d_i$ be the degree of vertex $v_i$. Fix positive integers $k$ and $m$ with $1 \leq k < m \leq r$ and define

$$e_i(k, m) = \begin{cases} 
  d_i - 2 & \text{if } k < i < m, \\
  d_i - 1 & \text{if } i = k \text{ or } i = m.
\end{cases}$$

Finally, let $s_{k,m} = \sum_{j=k}^{m} e_i(k, m)$. 

**Theorem 2.6.** We now choose \( \text{has} \).

**Lemma 2.5.** Let \( T \) be a caterpillar with spine \( \{v_1, \ldots, v_r\} \). Then the number of subtrees of \( T \) with \( L \) leaves and \( I \) non-leaves is

\[
\sum_{i=1}^{r-1} \left( \binom{S_{i,i+I}}{L} - \binom{S_{i,i+I-1}}{L} - \binom{S_{i+1,i+I}}{L} + \binom{S_{i+1,i+I-1}}{L} \right).
\]

**Proof.** Note that a subtree with exactly \( I \) internal edges must have non-leaf vertices \( \{v_i, \ldots, v_{i+I}\} \) for some \( 1 \leq i \leq r-I \). We now choose \( L \) vertices which are adjacent to these vertices to form \( L \) leaves, paying attention to two considerations:

1. \( v_i \) and \( v_{i+I} \) must each have at least one adjacent vertex chosen; otherwise \( v_i \) or \( v_{i+I} \) would be a leaf and \( S \) would not have \( I \) internal edges.

2. For the \( d_k \) vertices adjacent to \( v_k \), note that two vertices are already used \((v_{k-1} \text{ and } v_{k+1})\) when \( i < k < i+I \) and one vertex is already used at the endpoints \( v_i \) and \( v_{i+I} \).

The first consideration above is resolved easily: count all possible ways to select \( L \) vertices as leaves, then subtract those selections in which no vertices adjacent to \( v_i \) or \( v_{i+I} \) are chosen. Finally, add in those selections in which both \( v_i \) and \( v_{i+I} \) are excluded, since these have been removed twice.

For the second consideration, just count the number of vertices which legitimately can be chosen as leaves: each \( v_k \) has \( d_k - 2 \) possible choices (for \( i < k < i+I \)), while \( v_i \) and \( v_{i+I} \) have \( d_i - 1 \) and \( d_{i+1} - 1 \) choices, respectively. This coincides precisely with the definition of the \( e_i(k, m) \).

**Theorem 2.6.** Let \( T_1(\alpha, \beta) \) be a caterpillar with \((\text{non-leaf})\) degree sequence \( \{\alpha+1, \beta+1, \alpha+\beta+1, 1\} \) and let \( T_2(\alpha, \beta) \) be a caterpillar with \((\text{non-leaf})\) degree sequence \( \{\alpha+1, \alpha+\beta+1, 1, \alpha+1, \beta+1\} \), as in Fig. 1, where \( \alpha \) and \( \beta \) are positive integers. Then, \( g(T_1) = g(T_2) \).

**Proof.** We must show that \( T_1 \) and \( T_2 \) have the same number of subtrees with \( L \) leaves and \( I \) non-leaves for all values of \( L \) and \( I \). Let \( t_i(L, I) \) denote the number of such subtrees in \( T_i \), for \( i = 1, 2 \), and note that \( 0 \leq I \leq 4 \).

1. \( I = 0 \): Subtrees with no internal edges are stars, and the number of such subtrees is completely determined by the degree sequence. But \( T_1 \) and \( T_2 \) have the same degree sequences, so \( t_1(L, 0) = t_2(L, 0) \) for all \( L \geq 0 \).

2. \( I = 1 \): Such a subtree \( T_1 \) or \( T_2 \) uses exactly one of the non-leaf edges in \( T_1 \) and \( T_2 \). Then, there is a bijection between the four non-leaf edges of \( T_1 \) and those of \( T_2 \) so that the number of subtrees having \( L \) leaves using an edge in \( T_1 \) is the same as the number using the corresponding edge in \( T_2 \). One bijection is: \( a \leftrightarrow c', b \leftrightarrow d', c \leftrightarrow a', d \leftrightarrow b' \). Thus, \( t_1(L, 1) = t_2(L, 1) \) for all \( L \geq 0 \).

3. \( I = 2 \): We apply Lemma 2.5. After simplifying, we have (for \( i = 1, 2 \))

\[
t_1(L, 2) = 2 \left( \binom{2\alpha + 2\beta - 1}{L} - \binom{\alpha - 1}{L} + \binom{\beta - 1}{L} \right) - \left( \binom{2\alpha - \beta - 1}{L} + \binom{\alpha + 2\beta - 1}{L} - \binom{\alpha + \beta - 1}{L} \right).
\]

4. \( I = 3 \): We use the lemma again. This time, we have (for \( i = 1, 2 \))

\[
t_1(L, 3) = 3 \left( \binom{3\alpha + 2\beta - 2}{L} + \binom{2\alpha + 3\beta - 2}{L} - \binom{\alpha + \beta - 2}{L} \right).
\]
Table 1

\[
\begin{array}{ll}
D(T_1(\alpha, \beta)) & D(T_2(\alpha, \beta)) \\
(\alpha + \beta x)(1 + x^2 + x^3) & (\alpha + \beta x)(1 + x + x^3) \\
(\alpha + \beta x)(1 + x^2 + x^3 + x^4) & (\alpha + \beta x)(1 + x + x^2 + x^4) \\
(\alpha + \beta x)(1 + x^2 + x^3 + x^5) & (\alpha + \beta x)(1 + x + x^3 + x^5) \\
(\alpha + \beta x)(1 + x^2 + x^3 + x^4 + x^5) & (\alpha + \beta x)(1 + x + x^2 + x^4 + x^5) \\
(\alpha + \beta x)(1 + x + x^3 + x^4 + x^5) & \\
\end{array}
\]

(5) \(I = 4\): Note that all four internal edges of \(T_1\) and \(T_2\) are needed. Then, for \(i = 1, 2\)
\[
t_1(L, 4) = \left( \frac{3\alpha + 3\beta - 3}{L} \right) - \left( \frac{3\alpha + 2\beta - 3}{L} \right) \\
- \left( \frac{2\alpha + 3\beta - 3}{L} \right) + \left( \frac{2\alpha + 2\beta - 3}{L} \right).
\]
\(\square\)

Since \(T_1(\alpha, \beta)\) and \(T_2(\alpha, \beta)\) are not isomorphic for any positive integers \(\alpha \neq \beta\), we have the following:

**Corollary 2.7.** Let \(\alpha \neq \beta\) be positive integers. Then \(T_1(\alpha, \beta)\) and \(T_2(\alpha, \beta)\) are non-isomorphic trees with the same greedoid Tutte polynomial.

In [5], non-isomorphic caterpillars with the same degree sequence and the same number of paths of length \(k\) for all \(k\) are constructed. This amounts to creating two trees in which \(t_1(L, 0) = t_2(L, 0)\) for all \(L \geq 0\) and \(t_1(2, I) = t_2(2, I)\) for all \(I \geq 0\). Generating functions play a central role in generating those examples: if \(T\) is a caterpillar with spine vertices \(v_1, \ldots, v_I\), let \(D(T) = \sum_{i=1}^{I} e_i\), where \(e_i + 1 = \deg(v_i)\). Then, the polynomial \(x^I D(T; x)D(T; x^{-1})\) encodes the degree sequence and the number of paths of length \(k\) for any \(k\) [5, Lemma 2].

For our example, we find \(D(T_1(\alpha, \beta)) = (\alpha + \beta x)(1 + x^2 + x^3)\), and \(D(T_2(\alpha, \beta)) = (\alpha + \beta x)(1 + x + x^3)\). Thus, \(x^I D(T_1; x)D(T_1; x^{-1}) = x^I D(T_2; x)D(T_2; x^{-1})\), so \(T_1\) and \(T_2\) have the same degree sequence and the same number of paths of any length.

Further, we could create additional counterexamples by modifying one of the factors in this expression. The reader can check that the generating polynomials in Table 1 also produce non-isomorphic caterpillars with the same greedoid Tutte polynomial.

In general, let \(p(x)\) be a polynomial with coefficients drawn from \(\{0, 1\}\) whose coefficient list does not have two consecutive 0’s. We conclude with a conjecture that such polynomials will always generate caterpillars with identical subtree data.

**Conjecture 2.8.** Let \(T_1\) and \(T_2\) be caterpillars with \(D(T_1) = (\alpha + \beta x) p(x)\) and \(D(T_2) = (\beta + \alpha x) p(x)\). Then \(T_1\) and \(T_2\) have the same greedoid Tutte polynomial.

**Acknowledgements**

We thank Jennifer Feder, Greg Francos, and Amanda Redlich for useful discussions.

**References**