Tutte Polynomials for Trees

Sharad Chaudhary
DEPARTMENT OF MATHEMATICS
DUKE UNIVERSITY
DURHAM, NORTH CAROLINA

Gary Gordon
DEPARTMENT OF MATHEMATICS
LAFAYETTE COLLEGE
EASTON, PENNSYLVANIA

ABSTRACT

We define two two-variable polynomials for rooted trees and one two-variable polynomial for unrooted trees, all of which are based on the corank-nullity formulation of the Tutte polynomial of a graph or matroid. For the rooted polynomials, we show that the polynomial completely determines the rooted tree, i.e., rooted trees T_1 and T_2 are isomorphic if and only if $f(T_1) = f(T_2)$. The corresponding question is open in the unrooted case, although we can reconstruct the degree sequence, number of subtrees of size k for all k, and the number of paths of length k for all k from the (unrooted) polynomial. The key difference between these three polynomials and the standard Tutte polynomial is the rank function used; we use pruning and branching ranks to define the polynomials. We also give a subtree expansion of the polynomials and a deletion-contraction recursion they satisfy.

1. INTRODUCTION

In this paper, we are concerned with defining and investigating two-variable polynomials for trees and rooted trees that are motivated by the Tutte polynomial. The Tutte polynomial, a two-variable polynomial defined for a graph or a matroid, has been studied extensively over the past few decades. The fundamental ideas are traceable to Veblen [9], Birkhoff [1], Whitney [10], and Tutte [8], who were concerned with colorings and flows in graphs. The chromatic polynomial of a graph is a special case of the Tutte polynomial. Brylawski [3] gives an indication of the diversity of application of the Tutte polynomial by listing some 19 distinct areas of combinatorics, from acyclic orientations of graphs to zonotopes, in which some evaluation of the Tutte

Journal of Graph Theory, Vol. 15, No. 3, 317–331 (1991)
 CCC 0364-9024/91/030317-15$04.00

polynomial plays a role. An extensive introduction to the Tutte polynomial that gives a very nice account of its application to graph theory and coding theory can be found in [4].

In this paper, we will concentrate on how we can modify the definition of the Tutte polynomial to get a meaningful invariant for trees and rooted trees. (Although we will not discuss *greedoids* here, much of what we discuss can be generalized to them.) Before modifying the definition, we give two equivalent definitions of the (ordinary) Tutte polynomial of a graph.

Let $G = (V, E)$ be a (finite) graph, with vertex set V and edge set E. We say $F \subseteq E$ is *acyclic* if it contains no cycles. Define the *rank* of $A \subseteq E$, denoted $r(A)$, by

1.1. $r(A) \equiv \max_{F \subseteq A}\{|F| : F \text{ is acyclic}\}$.

If $r(A) = |A|$, then we say A is an *independent* set of edges. Thus the rank of a set of edges A is simply the size of the largest independent subset of A. We now use this definition of rank to define the Tutte polynomial (or, more accurately, the *corank-nullity* polynomial):

1.2. $T(G; t, z) \equiv \sum_{A \subseteq E(G)} t^{r(E(G)) - r(A)} z^{|A| - r(A)}$.

We can reformulate the definition in a recursive way. We first remind the reader of two basic operations in graph theory. *Deleting* an edge e in a graph G means simply erasing the edge; *contracting* the edge e means erasing e and then identifying the two end points of e. We denote the resulting graphs by $G - e$ (for deletion) and G/e (for contraction). A *loop* in a graph is an edge that has only one end point; an *isthmus* (or *bridge*) is an edge whose deletion increases the number of connected components of G. Thus, for example, every edge in a tree is an isthmus. Using the notion of independence given above, an edge e is a loop if and only if it is in no independent set; e is an isthmus if and only if it is in every maximal independent set. We will not allow loops to be contracted or isthmuses to be deleted, although we remark that it is possible to carry out the rest of this formulation without this restriction. We can now give a recursive definition of the Tutte polynomial:

1.3(a). If G has no edges, then $T(G; x, y) = 1$.
(b). $T(G) = xT(G/e)$ if e is an isthmus.
(c). $T(G) = yT(G - e)$ if e is a loop.
(d). $T(G) = T(G/e) + T(G - e)$ if e is neither a loop nor an isthmus.

Definitions 1.2 and 1.3 are equivalent under the substitution $x = t + 1$ and $y = z + 1$. It is also clear from either definition that if G is a tree, then $T(G) = (t + 1)^n = x^n$, where n is the number of edges of G. For rooted trees, the Tutte polynomial is not defined.

As mentioned above, our goal is to modify the original Tutte polynomial so that it will provide more information about trees. We can easily define a "Tutte-like" polynomial by using 1.2 with a different rank function. If the rank function is reasonably well behaved, then the resulting polynomial should have some interesting properties and reformulations. In Section 2, we do this for rooted trees, and in Section 3 we concentrate on (unrooted) trees. For the two polynomials we consider in Section 2, the polynomial completely determines the rooted tree. The same question for unrooted trees is still open, although there are many familiar properties of the tree that can be determined from the polynomial. We explore these properties in Section 3.

2. ROOTED TREES

We begin by defining two rank functions for rooted trees, each of which gives a polynomial via 1.2. Let $T_r = (V, E)$ be a rooted tree, i.e., a tree with a distinguished vertex $*$, and let $A \subseteq E$. The *branching rank of A*, denoted $b(A)$, is defined by

2.1. $b(A) \equiv \max_{S \subseteq A}\{|S| : S \text{ is a rooted subtree, rooted at } *\}$.

Dually, define the *pruning rank of A* by

2.2. $p(A) \equiv \max_{S \subseteq A}\{|S| : S$ is the *complement* of a rooted subtree, rooted at $*\}$.

Thus, $p(A) = |A|$ if and only if $b(E - A) = |E - A|$. The sets satisfying $p(A) = |A|$ are precisely the sets of edges that can be "pruned" from T and leave a rooted tree.

Some general comments are in order here. Both b and p satisfy the following three properties, where $A, B \subseteq E$, x and y are edges, and r represents the rank function b or p.

R1. $r(A) \leq |A|$,
R2. $A \subseteq B$ implies $r(A) \leq r(B)$,
R3. If $r(A) = r(A \cup \{x\}) = r(A \cup \{y\})$, then $r(A) = r(A \cup \{x, y\})$.

Any rank function satisfying R1, R2, and R3 defines a *greedoid.* The sets A that satisfy $r(A) = |A|$ are called the *feasible sets* of the greedoid. (Thus, feasible is to greedoid as independent is to graph or matroid.) In this context, we are considering a rooted tree as a greedoid in two different ways. The feasible sets in the branching case are just the rooted subtrees; the feasible sets in the pruning case are the complements of rooted subtrees. Although we will not require the generality of greedoids here, we will use the term feasible set as is necessary.

We can now use the two rank functions b and p with Definition 1.2 to create two polynomials for rooted trees. Explicitly, we have

2.3. $f_b(T_r; t, z) \equiv \sum_{A \subseteq E} t^{b(E)-b(A)} z^{|A|-b(A)}$.

2.4. $f_p(T_r; t, z) \equiv \sum_{A \subseteq E} t^{p(E)-p(A)} z^{|A|-p(A)}$.

(Note that $p(E) = b(E) = |E|$.)

The main result of this section is that each of these polynomials, separately, determine the rooted tree. We state this explicitly now.

Theorem 1. Suppose T_1 and T_2 are rooted trees.

(a) If $f_b(T_1) = f_b(T_2)$, then T_1 and T_2 are isomorphic rooted trees.
(b) If $f_p(T_1) = f_p(T_2)$, then T_1 and T_2 are isomorphic rooted trees.

The proof of part a appears in [6]. An outline of the proof of part b is given later in this section. Both proofs depend on the fact that $\mathbf{Z}[x, y]$ is a unique factorization domain.

By Theorem 1, it is possible to calculate one polynomial from the other; given $f_b(T_r)$ (say), we can recover T_r (by the theorem), and hence calculate $f_p(T_r)$. It is possible to provide a "direct" way to calculate one polynomial from the other, but we omit the details here (see [5] for the full story).

In contrast with this difficulty, a direct procedure for calculating $f_p(T_r; t, 0)$ from $f_b(T_r; t, 0)$ (or vice versa) is easy to give.

Proposition 2. $f_p(T_r; t, 0) = t^n f_b(T_r; t^{-1}, 0)$, where n is the number of edges of T_r.

Proof. The coefficient of t^i in $f_p(t, 0)$ is just the number of subsets of edges A with $p(A) = |A|$ and $n - p(A) = i$. Thus, this coefficient counts the number of pruning feasible subsets of cardinality $n - i$. Similarly, the coefficient of t^j in $f_b(t, 0)$ counts the number of branching feasible subsets of cardinality $n - j$. Since A is branching feasible if and only if $E - A$ is pruning feasible, the proposition follows.

Evidently, setting $z = 0$ in either polynomial reduces the information carried by the polynomial. To see this explicitly, consider the following example.

Example 3. Let S_r and T_r be the two rooted trees of Figure 1. We set $y = z + 1$ (see Proposition 11) for ease of notation and note that from Proposition 2, any counterexample for f_b will also be a counterexample for f_p.

$$f_b(S_r) = t^7y^4 + t^6(2y^4 + y^3) + 4t^5y^3 + t^4(2y^3 + 2y^2) + t^3(2y^2 + y) + 2t^2y + t + 1.$$

$$f_b(T_r) = t^7y^4 + t^6(2y^4 + y^3) + t^5(y^4 + 2y^3 + y^2) + t^4(y^3 + 3y^2) + t^3(2y^2 + y) + 2t^2y + t + 1.$$

$$f_p(S_r) = t^7y^6 + t^6y^4 + 2t^5y^3 + t^4(y^3 + y^2 + y) + t^3(y^2 + 2y + 1) + t^2(y + 3) + 3t + 1.$$

$$f_p(T_r) = t^7y^6 + t^6y^4 + t^5(y^3 + y^2) + t^4(2y^2 + y) + t^3(y^2 + 2y + 1) + t^2(y + 3) + 3t + 1.$$

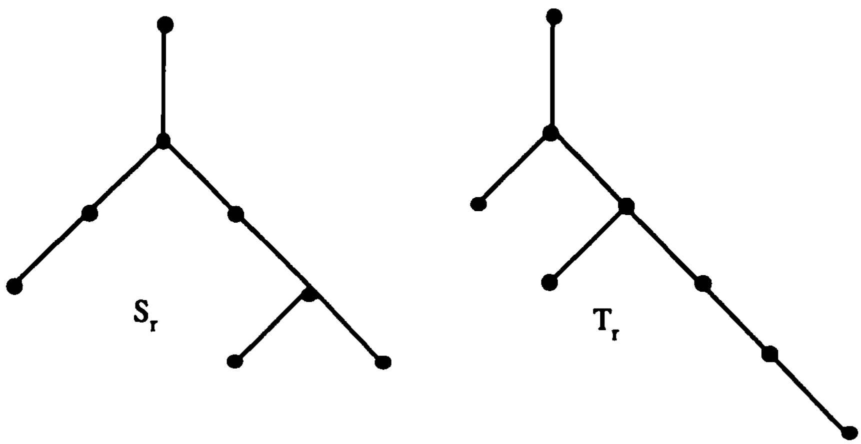

FIGURE 1

The ordinary Tutte polynomial derives much of its utility and interest from the recursive Definition 1.3. Both f_b and f_p also have recursive descriptions that are similar to 1.3.

Proposition 4. (a) Let e be any edge incident with the root $*$ of the rooted tree T_r. Then $f_b(T_r) = f_b(T_r/e) + t^{b(E)-b(E-e)}f_b(T_r - e)$

(b) Suppose e is the only edge incident with the root $*$ of the rooted tree T_r. Then $f_p(T_r) = t(z + 1)f_p(T_r/e) + 1 - tz$.

The proof of (a) is essentially the same as the proof of the equivalence of 1.2 and 1.3. Loosely, this involves identifying the subsets of $E - e$ with the subsets of E that do not contain e (and comparing the corresponding terms in 2.3 and 2.4) and identifying the subsets of E/e with the subsets of E that do contain e. We also remark that, strictly speaking, $T_r - e$ is not a rooted tree; it is a rooted tree together with some unreachable (from the root) edges. These edges are simply treated as loops, i.e., their rank is zero in $T_r - e$. A complete proof of (a) appears in [6].

We now prove part (b) of Proposition 4. Let $A \subseteq E$ with $e \notin A$. Then the rank $p(A)$ and the cardinality $|A|$ considered in T_r/e are the same as $p(A)$ and $|A|$ when considered in T_r. Furthermore, $p(E) = p(E - e) + 1$, where $p(E)$ is computed in T_r and $p(E - e)$ is computed in T_r/e. Thus

$$\sum_{e \notin A \subseteq E} t^{p(E)-p(A)} z^{|A|-p(A)} = tf_p(T_r/e).$$

Now if $e \in A$, we note that if $A \neq E$, then $p(A)$ in T_r is also the same as $p(A - e)$ in T_r/e, $|A| = |A - e| + 1$, and $p(E) = p(E - e) + 1$. Thus $p(E) - p(A)$ (in T_r) is one more than $p(E - e) - p(A - e)$ (in T_r/e) and $|A| - p(A)$ (in T_r) is also one more than $|A - e| - p(A - e)$ (in T_r/e). Thus

$$\sum_{e \in A \subseteq E} t^{p(E)-p(A)} z^{|A|-p(A)} = tzf_p(T_r/e) - tz + 1,$$

since we need to account for the case $A = E$. Adding these two disjoint collections together completes the proof.

It is easy to describe how either of the polynomials behaves under direct sum or one-point union of two rooted trees. The proof of the next proposition is a straightforward application of either the definitions 2.3 and 2.4, or of Proposition 4.

Proposition 5. Let T_r be the rooted tree formed by joining the two rooted trees S_r and U_r at their roots. Then, for $i = b$ or p, we have $f_i(T_r) = f_i(S_r)f_i(U_r)$.

Propositions 4 and 5 can be used in combination to give a complete recursive description of either polynomial. We give two applications of this technique.

Example 6. We compute $f_b(T_r)$ and $f_p(T_r)$ for the rooted tree of Figure 2.

We first compute $f_b(T_r)$. By Proposition 4(a), we have $f_b(T_r) = f_b(T_r/a) + t^5 f_b(T_r - a) = f_b(T_r/a) + t^5(z + 1)^4$. ($f_b(T_r - a) = (z + 1)^4$ since $T_r - a$ consists of 4 loops.) Now let L_r and R_r be the rooted left and right subtrees of T_r/a, i.e., L_r is the rooted tree consisting of the single edge b, and R_r is the rooted tree consisting of edges c, d, and e, with root at the "top" of c. By Proposition 5, we have $f_b(T_r/a) = f_b(L_r)f_b(R_r)$. By 4(a) again (or directly), we find $f_b(L_r) = 1 + t$ (since $f_b(L_r/b) = 1$) and $f_b(R_r) = f_b(R_r/c) + t^3(z + 1)^2$. Now $f_b(R_r/c) = (t + 1)^2$ from the above. Putting this all together yields $f_b(T_r) = (t^5 + t^4 + 2t^3 + 3t^2 + 3t + 1) + (4t^5 + 2t^4 + 2t^3)z + (6t^5 + t^4 + t^3)z^2 + 4t^5z^3 + t^5z^4$.

Calculating $f_p(T_r)$ involves singing the same tune with slightly different words. Specifically, we replace the terms of the form $t^k(z + 1)^{k-1}$ by $(1 - tz)$ and we multiply the terms arising from contraction by $t(z + 1)$. This time, we obtain $f_p(T_r) = (t^5 + 3t^4 + 3t^3 + 2t^2 + t + 1) + (2t^5 + 6t^4 + 5t^3 + 2t^2)z + (t^5 + 3t^4 + 2t^3)z^2$.

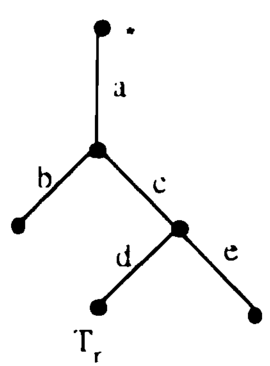

FIGURE 2

The second application of Propositions 4 and 5 yields algebraic characterizations of all possible polynomials that can occur as $f_b(T_r)$ or $f_p(T_r)$. Before giving this application, we give some easy evaluations of f_b and f_p (Proposition 7) and a result about irreducibility (Proposition 8).

The ordinary Tutte polynomial has several evaluations and properties that are easy to interpret in terms of the given graph. The same is true of the polynomials f_b and f_p. The proof of Proposition 7 is entirely straightforward and we omit it.

Proposition 7. For $i = b$ or p and T_r a rooted tree with n edges

(a) $f_i(T_r; 1, 0) =$ the number of rooted subtrees of T_r;
(b) $f_i(T_r; t, t^{-1}) = (t + 1)^n$;
(c) Special cases: $f_i(T_r; 1, 1) = 2^n$, $f_i(T_r; -1, -1) = 0$, and $f_i(T_r; 0, 1) = f_i(T_r; 0, 0) = 1$.

The proof of Proposition 8 for $f_b(T_r)$ appears in [6]. We give the proof for $f_p(T_r)$.

Proposition 8. Suppose T_r is a rooted tree whose root has degree one. Then, for $i = b$ or p, $f_i(T_r)$ is irreducible over $\mathbf{Z}[t, z]$.

Proof. Suppose the root $*$ of a rooted tree T_r has degree one, and $f_p(T_r; t, z) = g(t, z)h(t, z)$, where g and $h \in \mathbf{Z}[t, z]$. We write

(A) $g(t, z) = p_m(z)t^m + p_{m-1}(z)t^{m-1} + \cdots + p_1(z)t + 1$ and
(B) $h(t, z) = q_k(z)t^k + q_{k-1}(z)t^{k-1} + \cdots + q_1(z)t + 1$,

where $m + k = n$ is the number of edges of T_r. (We can assume both constant terms are 1 since $f_p(T_r; 0, 0) = 1$.) From definition 2.4, it is a straightforward calculation to show that the coefficient of t^n in $f_p(T_r; t, z)$ is $(z + 1)^c$, where c is the number of leaves in T_r. Thus we get $p_m(z) = (z + 1)^a$ and $q_k(z) = (z + 1)^b$, where a and b are nonnegative integers with $a + b = c$.

Next, we consider $f_p(T_r; t, -1)$. By Proposition 4(b), we know $f_p(T_r) = t(z + 1)f_p(T_r/e) + 1 - tz$. Plugging in $z = -1$ gives $f_p(T_r; t, -1) = (t + 1)$. Thus, we may assume $g(t, -1) = 1$ and $h(t, -1) = (t + 1)$. Combining this with (A) above yields

$$g(t, -1) = p_m(-1)t^m + p_{m-1}(-1)t^{m-1} + \cdots + p_1(-1)t + 1 = 1.$$

Thus, $p_i(-1) = 0$ for all $1 \le i \le m$, or $(z + 1)$ divides $p_i(z)$ for all $1 \le i \le m$. Thus, $p_i(1)$ is even for all i. Now since $f_p(T_r; 1, 1) = 2^n$, we know $g(1, 1) = \pm 2^d$ for some $d \ge 0$. But $g(1, 1) = p_m(1) + p_{m-1}(1) + \cdots + p_1(1) + 1 = \pm 2^d$ forces $g(1, 1) = \pm 1$.

Finally, by Proposition 7, we know $f_p(T_r; t, t^{-1}) = (t + 1)^n$. Thus, $g(t, t^{-1}) = \pm(t + 1)^e$ for some $e \ge 0$. Plugging in $t = 1$, we get

$g(1,1) = \pm 2^e$, so $e = 0$ by the above. Hence

$$g(t,t^{-1}) = p_m(t^{-1})t^m + p_{m-1}(t^{-1})t^{m-1} + \cdots + p_1(t^{-1})t + 1 = \pm 1.$$

But $p_m(t^{-1})t^m = (t^{-1} + 1)^a t^m$ has a term equal to t^m that cannot be canceled by any other term in $g(t, t^{-1})$. This contradiction proves the proposition.

Now let $I_b \subseteq \mathbf{Z}[t, z]$ be the set of all possible branching polynomials, i.e., $f \in I_b$ if and only if there is some rooted tree T_r such that $f_b(T_r) = f$. Define $I_p \subseteq \mathbf{Z}[t, z]$ in a similar way. We can now characterize I_b and I_p as follows.

Proposition 9. The following properties completely characterize I_b:

(B1) $1 \in I_b$.
(B2) If $f, g \in I_b$, then $fg \in I_b$.
(B3) If $f \in I_b$ and n is the t-degree of f, then $\{f + t^{n+1}(z + 1)^n\} \in I_b$.

Proposition 10. The following properties completely characterize I_p:

(P1) $1 \in I_p$.
(P2) If $f, g \in I_p$, then $fg \in I_p$.
(P3) If $f \in I_p$, then $\{t(z + 1)f + 1 - tz\} \in I_b$.

Proof. To prove these two propositions, we first show the six statements (B1)–(P3) are valid, i.e., they produce polynomials that are in I_b or I_p. Now (B1) and (P1) are trivial, (B2) and (P2) correspond to joining two disjoint rooted trees together at the root, and (B3) and (P3) follow from adding a single edge to a rooted tree at the root and then moving the root to the other endpoint of the new edge. (This last process is exactly the reverse of contracting the only edge incident with the root in a tree that has a root of degree one.) By Propositions 4 and 5, the polynomials produced by (B1)–(P3) are valid polynomials.

To show that any polynomial in I_b or I_p can be obtained via (B1)–(B3) or (P1)–(P3), respectively, we can implement the algorithm given below. (We give the algorithm for I_p; the algorithm for I_b is similar.) This algorithm depends on Proposition 8 and the fact that $\mathbf{Z}[t, z]$ is a unique factorization domain. Essentially, it allows us to reverse the process of forming a polynomial via (P2) and (P3).

Input. $f_p(T_r)$.
Step 1. Factor $f_p(T_r)$ into irreducibles over $\mathbf{Z}[t, z]$.
Step 2. For each factor p_i, set $p_{i\text{-new}} = (p_i - 1 + tz)(t(z + 1))^{-1}$.
Step 3. Go to Step 1, with $p_{i\text{-new}}$ replacing $f_p(T_r)$.

Each factor produced in Step 1 corresponds to a rooted subtree in which the root $*$ has degree one. The operation in Step 2 simply contracts that single edge incident with $*$.

Example 6 essentially shows how to use this algorithm to compute $f_p(T_r)$ for the rooted tree of Figure 1. It should also be clear how the algorithm provides a proof of Theorem 1, since we can uniquely reconstruct the rooted tree from the polynomial.

The final topic in this section concerns a description of the polynomials in terms of subtrees, rather than all subsets of edges. We omit the proof of (a) below. It is similar to the proof of (b), which we include. We denote by $L(S)$ the collection of leaves of a rooted subtree S, and by $M(S)$ the set of edges e that can be adjoined to S. Thus $M(S) = \{e \in E - S : S \cup \{e\}$ is a rooted subtree$\}$. We stipulate that if the root $*$ has degree one in a subtree, then the sole edge incident to $*$ is not counted as a leaf (unless the subtree has only one edge).

Proposition 11. Let $R(T_r)$ be the collection of all rooted subtrees of T_r, all rooted at $*$.

(a) $f_b(T_r; t, z) = \sum_{S \in R(T_r)} t^{|T_r| - |S|}(z + 1)^{|T_r| - |S| - |M(S)|}$.

(b) $f_p(T_r; t, z) = \sum_{S \in R(T_r)} t^{|S|}(z + 1)^{|S| - |L(S)|}$.

Proof of (b). We first note that the union of pruning feasible sets is feasible, i.e., if S_1 and S_2 are both the complements of rooted subtrees of T_r, then so is $S_1 \cup S_2$. Thus, if $p(A) = k$ for some $A \subseteq E$, then there is a *unique* feasible set B with $|B| = k$ and $B \subseteq A$. In this case, we say that A *uses* B, and denote by $U(B)$ the collection of all subsets that use B. Note that $p(A) = k$ for all $A \in U(B)$. We also let P be the collection of all feasible sets, i.e., $B \in P$ precisely when B is the complement of a rooted subtree. Then we can collect terms in definition 2.4 of f_p as follows:

$$f_p(t, z) = \sum_{A \subseteq E} t^{|E| - p(A)} z^{|A| - p(A)} = \sum_{B \in P} \sum_{A \in U(B)} t^{|E| - |B|} z^{|A| - |B|} = \sum_{B \in P} t^{|E| - |B|} z^{-|B|} \sum_{A \in U(B)} z^{|A|}$$

Now $|E| - |B| = |B^c|$ (where B^c denotes the complement of B). Further, A uses B if and only if $B \subseteq A$ and $A \cap L(B^c) = \phi$, where $L(B^c)$ is the collection of leaves of the rooted subtree B^c. Thus, A uses B means A consists of B together with some internal edges of B^c. Hence, we can rewrite the sum:

$$f_p(t, z) = \sum_{S \in R(T_r)} t^{|S|} \sum_{i=0}^{|S| - |L(S)|} \binom{|S| - |L(S)|}{i} z^i = \sum_{S \in R(T_r)} t^{|S|}(z + 1)^{|S| - |L(S)|}.$$

This completes the proof.

Proposition 11 gives us another way to view the polynomials $f_b(T_r)$ and $f_p(T_r)$; they simply list the ordered pairs $(|S|, |M(S)|)$ and $(|S|, |L(S)|)$, respectively, for every rooted subtree S. Combining this description with Theorem 1 gives us the last result in this section.

Corollary 12. (a) Any rooted tree T_r can be reconstructed from the list of all ordered pairs $(|S|, |M(S)|)$, where S ranges over all rooted subtrees of T_r.

(b) Any rooted tree T_r can be reconstructed from the list of all ordered pairs $(|S|, |L(S)|)$, where S ranges over all rooted subtrees of T_r.

3. UNROOTED TREES

In this section, we define Tutte-like polynomials for unrooted trees $T = (V, E)$. The branching rank function will not satisfy the greedoid rank properties R1–R3. (In particular, if we define $b(A)$ to be the size of the largest subtree contained in A, then R3 will not hold in general.) However, the reader can check that the pruning rank will satisfy R1–R3. In fact, we can define a pruning rank on either V or E and ostensibly obtain two polynomials $f_V(T)$ and $f_E(T)$. We will then show these two polynomials are essentially equivalent, and thereafter concentrate our efforts on $f_E(T)$.

A set $W \subseteq V$ is *feasible* if $V - W$ forms the vertices of a subtree, i.e., W is a collection of vertices that can be pruned from T. Now define the vertex pruning rank function $v(W)$ as follows:

3.1. $v(W) \equiv \max_{S \subseteq W}\{|S| : S \text{ is feasible}\}$.

Similarly, $A \subseteq E$ is *feasible* if $E - A$ forms the edges of a subtree. The edge pruning rank function $e(A)$ is then defined analogously. We can then define two polynomials $f_V(T; t, z)$ and $f_E(T; t, z)$ by using 1.2 exactly as in Section 2. Our first result is analogous to Proposition 11.

Proposition 13. Let $R(T)$ denote the collection of all subtrees of T, $L_V(S)$ denote the vertices of degree one in the subtree S, and $L_E(S)$ denote the leaves (as edges) of S.

(a) $f_V(T; t, z) = \sum_{S \subseteq R(T)} t^{|V(S)|}(z + 1)^{|V(S)| - |L_V(S)|}$

(b) $f_E(T; t, z) = \sum_{S \subseteq R(T)} t^{|E(S)|}(z + 1)^{|E(S)| - |L_E(S)|}$.

The proof of Proposition 13 is essentially the same as the proof of part (b) of Proposition 11. We note that $|L_V(S)| = |L_E(S)|$ if S has more than one edge and $|E(S)| = |V(S)| - 1$ if S is not empty. This leads us to the next result, which allows us to concentrate on $f_E(T)$. We again omit the proof, which is entirely straightforward.

Proposition 14. $f_V(T; t, z) = t(z + 1)f_E(T; t, z) + (n - 1)t + 1 - tz(1 + (n - 1)t)$.

From now on, we will write $f(T)$ for $f_E(T)$ and $L(S)$ for $L_E(S)$ for a subtree S to simplify notation.

It is possible to find connections between $f(T)$ and $f_p(T_r)$ from Section 2, where T_r is a rooted tree that is obtained from T somehow. One of the connections involves deletion and contraction, which we give now. Let e be a leaf of T and let $T_r(e)$ be the rooted tree obtained by placing the root at the vertex of e that has degree one. Then, as in Proposition 4, we can obtain a deletion-contraction formula for $f(T)$. The proof of this proposition follows along the same lines as the proof of Proposition 4.

Proposition 15. Let e be a leaf of T. Then $f(T) = f(T/e) + tf_p(T_r(e)/e)$.

The rooted tree $T_r(e)/e$ can naturally be associated with the deletion. The association is made precise in the context of greedoids, where the feasible sets of the deletion and the contraction are defined in a way that is analogous to the corresponding definitions in matroid theory. We refer the reader to [2] for more information on greedoids.

We can generalize Proposition 15 by allowing e to be any edge in T. We state the next proposition without proof.

Proposition 16. Let e be any edge of T, and let R and S be the two subtrees that are formed when e is deleted. Denote by R_r and S_r the trees R and S with roots at the endpoints of e. Then $f(T) = tf_p(R_r)f_p(S_r) + f(R) + f(S) - 1$.

Proposition 15 follows from 16, since, if e is a leaf, we can take $R = T/e$, $S =$ the one vertex tree (so $f(S) = f(S_r) = 1$) and $R_r = T_r(e)/e$. We also note that the recursion given by 15 is essentially complete, since we can use Proposition 4 to continue the recursion for the rooted trees that arise.

The final connection between the rooted and unrooted polynomials we give involves the following procedure. For an unrooted tree T with n edges, let $g(T)$ be defined by $g(T) = \Sigma_{r\in V} f_p(T_r)$. Thus, $g(T)$ is the sum of all $(n + 1)$ possible rooted polynomials T can generate. Then we can compute $g(T)$ from $f(T)$.

Proposition 17. $g(T) = f(T) + t(z + 1)(\partial f/\partial t) - (z^2 + z)(\partial f/\partial z) + n(1 - tz)$.

Proof. Let S be a subtree of T having k edges and m leaves, with $k > 1$. From Proposition 13, S contributes the term $t^k(z + 1)^{k-m}$ to $f(T)$. To compute the contribution S makes to $g(T)$, we use Proposition 11(b) and we note that S will contribute a nonzero term to a summand if and only if the root of that summand is in S. There are then two cases to consider.

Case 1. The root is an internal vertex in S, i.e., the root has degree >1 in S. In this case, S will be correspond to a rooted subtree with k edges and m leaves. Thus, by Proposition 11(b), S contributes the same term to $g(T)$ as it did to $f(T)$, namely $t^k(z+1)^{k-m}$. Since there are $k-m+1$ internal vertices in a tree with $k > 1$ edges, we get a total of $(k-m+1)t^k(z+1)^{k-m}$ from S for this case.

Case 2. The root is an external vertex of S, i.e., the root has degree one in S. This time, S will correspond to a rooted subtree with k edges and $m-1$ leaves. Thus, we get a contribution of $t^k(z+1)^{k-m+1}$ to $g(T)$ from S. Since there are m external vertices, the total contribution S makes to $g(T)$ in this case is $mt^k(z+1)^{k-m+1}$.

The reader can now check that the first three terms in the statement of the proposition will contribute exactly $(k-m+1)t^k(z+1)^{k-m} + mt^k(z+1)^{k-m+1}$ for each term in $f(T)$ of the form $t^k(z+1)^{k-m}$. To see how far this sum is from $g(T)$, we must separately consider the cases $k = 1$ and $k = 0$, since the computations in Cases 1 and 2 are only valid for $k > 1$. If $k = 1$, then S is simply an edge and these three terms contribute $t + t(z+1)$. The actual contribution an edge makes to $g(T)$ is just $2t$, since it contributes t for each endpoint. If $k = 0$, then S is empty and the actual contribution to $g(T)$ is $n+1$, while the three terms contribute 1. Thus

$$g(T) - \left\{f(T) + t(z+1)\frac{\partial f}{\partial t} - (z^2+z)\frac{\partial f}{\partial z}\right\} = 2nt - n(t + t(z+1)) + (n+1) - 1 .$$

The proposition now follows at once.

We now return to the question of what information about the tree can be reconstructed from $f(T)$. We will show that the polynomial determines the degree sequence of the tree and the number of paths of length k for every k. (As in Corollary 12, we can use Proposition 13 to view the polynomial as simply listing all ordered pairs $(|S|, |L(S)|)$ where S again ranges over all subtrees of T. Obviously then, $f(T)$ also determines the size and number of leaves of every subtree.)

We recall some basic definitions. The *distance* $d(v,w)$ between vertices v and w is just the length of the unique path from v to w. The *center* $Z(T)$ of a tree T is defined by $Z(T) = \{v \in V$: the maximum $d(v,w)$ for $w \in V$ is minimum$\}$. A tree is *central* if its center consists of one vertex; otherwise $|Z(T)| = 2$ and we say T is bicentral.

Proposition 18. (a) The degree sequence of T can be determined from $f(T)$.

(b) The number of paths of length k for all k can be determined from $f(T)$.

(c) Whether T is central or bicentral can be determined from $f(T)$.

Proof. (a) Suppose $f(T)$ is given as a polynomial in t and $(z+1)$ as in Proposition 13. We first determine the maximum degree D of any vertex of T and the number n_D of vertices achieving this maximum. Now a vertex v of degree D corresponds to a subtree S_v having D edges and D leaves. This S_v contributes t^D to the polynomial. Thus $D = \max\{d\colon a_d t^d$ appears in $f(T)$ for some $a_d > 0\}$ and $n_D = a_D$.

Now let $d < D$ and assume, by induction, that we know the number n_k of vertices of degree k for all k with $d < k \leq D$. Then each vertex of degree $k > d$ will contribute to the coefficient a_d of t^d in $f(T)$. More precisely, we get $n_d = a_d - \sum_{k=d+1}^{D} \binom{k}{d} n_k$.

(b) A path in a tree is a subtree that has two leaves. The proof can be constructed along the same lines as part (a), where we consider terms of the form $t^k(z+1)^{k-2}$. As in (a), we first find the longest path and then proceed inductively. We leave the details to the reader.

(c) A tree is central if and only if the longest path has even length. The result follows immediately from (b).

We now give an example analogous to Example 3 that shows that $f(T; t, 0)$ does not determine T.

Example 19. Let T_1 and T_2 be the trees of Figure 3. As in Example 3, we set $y = z + 1$ for ease of notation.

$$\begin{aligned} f(T_1) = {} & t^9y^5 + 2t^8(y^5 + y^4) + t^7(6y^4 + 2y^3) + t^6(3y^4 + 7y^3 + 2y^2) \\ & + t^5(5y^3 + 8y^2 + y) + t^4(8y^2 + 6y) + t^3(10y + 2) + 10t^2 \\ & + 9t + 1. \end{aligned}$$

$$\begin{aligned} f(T_2) = {} & t^9y^5 + 2t^8(y^5 + y^4) + t^7(y^5 + 4y^4 + 3y^3) + t^6(2y^4 + 8y^3 + 2y^2) \\ & + t^5(5y^3 + 8y^2 + y) + t^4(8y^2 + 6y) + t^3(10y + 2) + 10t^2 \\ & + 9t + 1. \end{aligned}$$

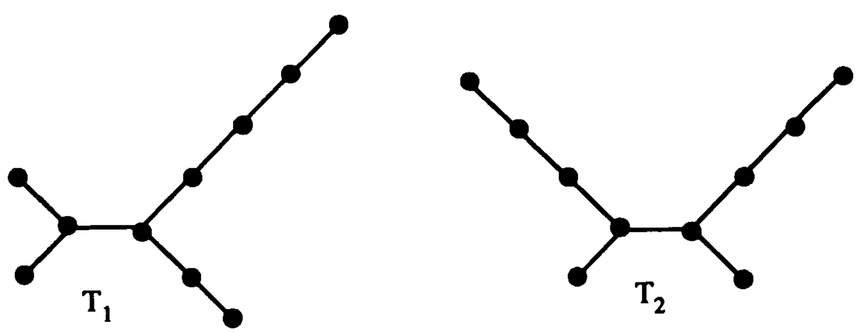

FIGURE 3

We conclude by briefly considering the connection between the lattice of all subtrees of a tree and $f(T)$. As in Section 2, let $M(S) = \{e \in E - S: S \cup \{e\}$ is a subtree$\}$ for a subtree S. Let $L(T)$ be the lattice whose elements are the (edge-sets of) subtrees of T ordered by inclusion. In [11], several properties of the lattice are verified. For example, it is clear that if S is a subtree, then the lattice rank $r(S) = |S|$, the number of subtrees that S covers in the lattice is just $|L(S)|$ and the number of subtrees that cover S is $|M(S)|$. Thus, we can compute $f(T)$ directly from the lattice $L(T)$:

$$f(T) = \sum_{x \in L(T)} t^{r(x)} z^{r(x)-c(x)}$$

where $r(x)$ is the rank of x in the lattice and $c(x)$ is the number of elements x covers. It is also clear that the tree T can be reconstructed from $L(T)$, since the entire edge incidence relation can be recovered from the rank 1 and rank 2 levels of $L(T)$.

The main open question remaining is whether $f(T)$ completely determines T or not. We see three possible avenues of attack on this problem.

1. Try to reconstruct the lattice $L(T)$ from $f(T)$. It is possible to determine both sets of ordered pairs $(|S|, |M(S)|)$ and $(|S|, |L(S)|)$ directly from $f(T)$, where S ranges over all subtrees of T.

Computing $(|S|, |M(S)|)$ essentially follows from problem 21 on page 157 of [7], while the computation for $(|S|, |L(S)|)$ is immediate from Proposition 13. Thus, some of the information in $L(T)$ can be reconstructed from $f(T)$.

2. Try to use Theorem 1 in connection with one of Propositions 15, 16, or 17. Theorem 1 tells us how to reconstruct rooted trees from $f_p(T_r)$, while 15–17 relate $f_p(T_r)$ to $f(T)$ in various ways. Thus if we could recover $f_p(T_r)$ from $f(T)$ for any rooted tree associated with T, then we could reconstruct T from $f(T)$.
3. Use the fact that the "standard" reconstruction problem is solved for trees. Here, the reconstruction problem refers to reconstructing a graph from the deck of edge-deleted subgraphs. It is not clear how to generate the deck of edge-deleted subgraphs of T from $f(T)$, however.

Finally, we mention that many applications of $f(T)$ remain to be explored, especially for greedoids T, which are not trees or rooted trees. For instance, many of the results given here have been extended to posets and antimatroids [5].

ACKNOWLEDGMENT

The work of SC was partially supported by a Lafayette College Dana Scholarship student program.

References

[1] G. D. Birkhoff, A determinant formula for the number of ways of coloring a map. *Ann. Math.* (2) 14 (1912–3) 42–46.

[2] A. Björner and G. Ziegler, Introduction to Greedoids. *Theory of Matroids, Advanced Theory,* Cambridge University Press, Cambridge, to appear.

[3] T. Brylawski, The Tutte polynomial, part I: General theory. *Matroid Theory and Its Applications.* Proceedings of Third International Mathematical Summer Centre (C.I.M.E.) 1980. Ligouri, Naples, Italy (1982) 125–276.

[4] T. Brylawski and J. Oxley, The Tutte polynomial and its applications. *Theory of Matroids, Advanced Theory.* Cambridge University Press, Cambridge, to appear.

[5] G. Gordon, A Tutte polynomial for partially ordered sets. Submitted.

[6] G. Gordon and E. McMahon, A greedoid polynomial which distinguishes rooted arborescences. *Proc. Am. Math. Soc.* 2 **107** (1989) 287–298.

[7] R. P. Stanley, *Enumerative Combinatorics, Volume I.* Wadsworth & Brooks/Cole, CA (1986).

[8] W.T. Tutte, A contribution to the theory of chromatic polynomials. *Can. J. Math.* **6** (1954) 80–91.

[9] O. Veblen, An application of modular equations in analysis situs. *Ann. Math.* 2 **14** (1912) 86–94.

[10] H. Whitney, A logical expansion in mathematics. *Bull. Am. Math. Soc.* **38** (1932) 572–579.

[11] B. Zelinka, The lattice of all subtrees of a tree. *Math. Slovaca* **27** (1977) 277–286.