Notes on the Tutte Polynomial

1 The recursive definition

We have already seen several graph isomorphism invariants that satisfy (and can by computed by means of) some sort of deletion-contraction recurrence:

$$\begin{aligned}
\tau(G) &= \tau(G-e) + \tau(G/e) && \text{(number of spanning trees)} \\
a(G) &= a(G-e) + a(G/e) && \text{(number of acyclic orientations)} \\
\chi(G;\,k) &= \chi(G-e;\,k) - \chi(G/e;\,k) && \text{(the chromatic polynomial)}
\end{aligned} \tag{1.1}$$

The *Tutte polynomial of a graph* is its "most general deletion-contraction invariant"; all these invariants, and much more useful information, can be obtained from it.

Definition: Let G be a graph. The *Tutte polynomial* $T(G) = T(G;\,x, y)$ is defined by the recurrence

$$T(G;\,x,y) \;:=\; \begin{cases} 1 & \text{if } E(G) = \emptyset, & \text{(a)} \\ x \cdot T(G/e) & \text{if } e \text{ is a cut-edge,} & \text{(b)} \\ y \cdot T(G-e) & \text{if } e \text{ is a loop,} & \text{(c)} \\ T(G-e) + T(G/e) & \text{for any other edge } e. & \text{(d)} \end{cases} \tag{1.2}$$

Note that we have to keep track of loops and parallel edges; it is *not* true in general that the Tutte polynomial of a graph equals that of its underlying simple graph.

The problem with this definition is that it is far from clear clear why the polynomial $T(G)$ is independent of the choice of edge e. We will eventually prove that by giving a closed formula for $T(G)$ that does not depend on any such choice. First, some examples.

2 Examples

Example 2.1. Suppose G is a forest with m edges. Then

$$T(G;\,x,y) = x^m. \tag{2.1}$$

For $r = 0$, this is case (a) of the recurrence (1.2). Otherwise, every edge of G is a cut-edge, and G/e is a forest with $r-1$ edges. By induction, $T(G/e) = x^{r-1}$, so $T(G) = x^r$ by case (b).

Example 2.2. Let L_r be the graph with one vertex and r loops. By a similar argument, using case (c) of (1.2), we have

$$T(L_r;\,x,y) = y^r. \tag{2.2}$$

More generally, any graph with r edges, all of which are loops, has Tutte polynomial y^r.

Example 2.3. It can be shown by induction on n that the Tutte polynomial of the n-cycle is

$$T(C_n;\,x,y) = x^{n-1} + x^{n-2} + \cdots + x^2 + x + y. \tag{2.3}$$

If we swap x and y in this formula, we obtain the Tutte polynomial of a graph with two vertices joined by n parallel edges.

Example 2.4. Let $G = K_3$ and e any edge (it doesn't matter which one). Then case (d) of (1.2) gives

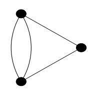
$$\begin{aligned} T(K_3) &= T(K_3 - e) + T(K_3/e) \\ &= T(P_3) + T(C_2). \end{aligned}$$

Now P_3 is a tree with two edges, so its Tutte polynomial is x^2. For the digon C_2, let f be either edge. Then

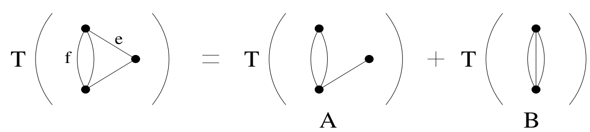
$$\begin{aligned} T(C_2) &= T(C_2 - f) + T(C_2/f) \\ &= T(K_2) + T(L_1) \\ &= x + y, \end{aligned}$$

so $T(K_3) = x^2 + x + y$.

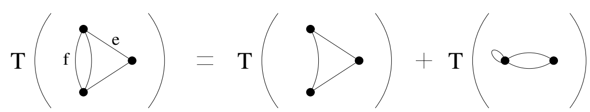
Example 2.5. We will compute the Tutte polynomial of the following graph G:



Applying the recurrence with the edge e gives



while applying the recurrence with f gives



There is no particular reason why these two calculations ought to yield the same answer—but they do. First,

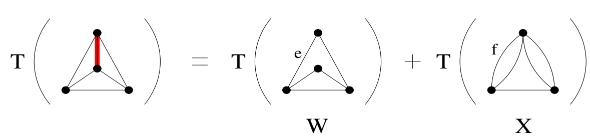
$$\begin{aligned} T(A) &= x \cdot T(C_2) = x(x+y), \\ T(B) &= T(C_2) + T(L_2) = (x+y) + (y^2), \end{aligned}$$

yielding $T(G) = x^2 + xy + y^2 + x + y$. On the other hand,

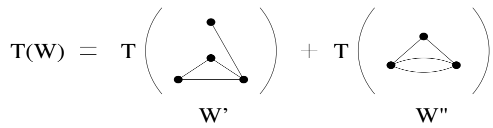
$$\begin{aligned} T(Y) &= T(K_3) = x^2 + x + y, \\ T(Z) &= y \cdot T(C_2) = y(x+y), \end{aligned}$$

giving the same answer for $T(G)$.

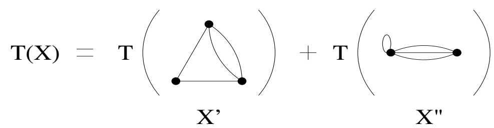
Example 2.6. One more calculation: $T(K_4)$. Of course, it doesn't matter which edge we start with; in the following diagram, we use the one indicated in red.



So we now have to calculate the Tutte polynomials of W and X. Apply the recurrence (1.2) with the labeled edge $e \in E(W)$:



Now $T(W') = x \cdot T(K_3)$, and W'' is the graph G of Example 4. On the other hand, applying the recurrence to $f \in E(X)$ gives



So $X' \cong G$, and $T(X'') = y \cdot T(B)$, with B as in Example 4. Putting it all together:

$$\begin{aligned} T(K_4) &= T(W) + T(X) \\ &= T(W') + T(W'') + T(X') + T(X'') \\ &= x(x^2 + x + y) + 2(x^2 + xy + y^2 + x + y) + y(x + y + y^2) \\ &= (x^3 + y^3) + (3x^2 + 2xy + 3y^2) + (2x + 2y). \end{aligned}$$

There are some interesting things about this polynomial. First, it is symmetric in x and y—that is, if we swap x and y, the polynomial is unchanged. Second, if we plug in various values of x and y, the numbers that come out are rather suggestive:

$$\begin{array}{lll} T(K_4;\, 0,0) = 0, & & \\ T(K_4;\, 0,1) = 6, & T(K_4;\, 1,1) = 16, & \\ T(K_4;\, 0,2) = 24, & T(K_4;\, 1,2) = 38, & T(K_4;\, 2,2) = 64. \end{array}$$

It's not entirely clear what all this means, but $24 = 4!$ is the number of acyclic orientations of K_4 and 16 is the number of spanning trees, among other things. These same substitutions work for K_3; recall that $T(K_3;\, x, y) = x^2 + x + y$, so $T(K_3;\, 1,1) = 3$ is the number of spanning trees and $T(K_3;\, 2,0) = 6 = 3!$ is the number of acyclic orientations.

(These are not coincidences!)

3 A closed formula for $T(G)$

In order to prove that the Tutte polynomial is well-defined by the recurrence (1.2), we will give a closed formula for $T(G;\, x, y)$, one that does not potentially depend on any choiuce of edges and consequently is well-defined.

Definition: Let $G = (V, E)$ be a graph. The *rank* $r_G(F) = r(F)$ of an edge set $F \subseteq E$ is defined as

$$r(F) \;:=\; \max\{|X| :\; X \subseteq F \text{ is acyclic.}\} \tag{3.1}$$

The rank of G itself, denoted $r(G)$, is just the rank of its edge set. Equivalently, this is the size of a spanning forest of G.

The *corank* of F is $r(G) - r(F)$. This can be described as the minimum number of edges that need to be added to F in order to make it span G—that is, contain a spanning forest. That is,

$$r(G) - r(F) \;=\; \min\{|Y| : \; Y \supseteq F \text{ spanning}\} - |F|. \tag{3.2}$$

We can now state the closed formula for the Tutte polynomial:

$$\mathbf{T}(G;\, x, y) \;:=\; \sum_{F \subseteq E} (x-1)^{r(G)-r(F)} (y-1)^{|F|-r(F)}. \tag{3.3}$$

(For the time being, I use the boldface $\mathbf{T}(G)$ to distinguish this polynomial from the recursive definition of $T(G)$. Once we have proved that the two are equal, I'll dispense with the boldface. In lecture, I used $T(G)$ for the closed-form and $\tilde{T}(G)$ for the recurrence.)

Theorem 3.1. *$\mathbf{T}(G;\, x, y = T(G;\, x, y)$.*

Before giving the proof, we use the formula (3.3) to calculate $T(K_3;\, x, y)$. Note that $r(G) = 2$, and that $r(F) = \min(|F|, 2)$ for all $F \subseteq E$. We calculate the Tutte polynomial as follows, using (3.3):

$\lvert F\rvert$	# of sets $F \subseteq E$	$r(F)$	Contribution to $T(G;\, x, y)$		
0	1	0	$(x-1)^2(y-1)^0$	$=$	$x^2 - 2x + 1$
1	3	1	$3(x-1)^1(y-1)^0$	$=$	$3x - 3$
2	3	2	$3(x-1)^0(y-1)^0$	$=$	3
3	1	2	$(x-1)^0(y-1)^1$	$=$	$y - 1$
5					$x^2 + x + y.$

Note that this agrees with the recursive calculation in Example 3.

Proof of Theorem 3.1. We induct on $e(G)$. For the base case $E = \emptyset$, note that $r(\emptyset) = |\emptyset| = 0$, so that (3.3) gives

$$\mathbf{T}(G;\, x, y) = (x-1)^0(y-1)^0 = 1 = T(G;\, x, y)$$

by case (a) of (1.2).

For the inductive step, suppose that the theorem holds for all graphs with fewer edges than G. In particular, if we choose $e \in E$ arbitrarily, then the theorem holds for $G - e$ and (provided that e is not a loop) for G/e.

First, suppose that e is neither a cut-edge nor a loop. Then $r(G) = r(G - e) = r(G/e) + 1$, and for $F \subseteq E$,

$$r_G(F) = \begin{cases} r_{G-e}(F) & \text{if } e \notin F, \\ r_{G/e}(F - e) + 1 & \text{if } e \in F. \end{cases}$$

So we can calculate $T(G;\, x, y)$ as

$$\sum_{\substack{F\subseteq E\\ e\notin F}} (x-1)^{r(G)-r(F)}(y-1)^{|F|-r(F)} + \sum_{\substack{F\subseteq E\\ e\in F}} (x-1)^{r(G)-r(F)}(y-1)^{|F|-r(F)}$$

$$= \sum_{F:\ e\notin F} (x-1)^{r(G-e)-r_{G-e}(F)}(y-1)^{|F|-r_{G-e}(F)} + \sum_{F:\ e\in F} (x-1)^{r(G/e)-r_{G/e}(F-e)}(y-1)^{|F-e|-r_{G/e}(F-e)}$$

$$= \sum_{F\subseteq E-\{e\}} (x-1)^{r(G-e)-r_{G-e}(F)}(y-1)^{|F|-r_{G-e}(F)} + \sum_{F\subseteq E-\{e\}} (x-1)^{r(G/e)-r_{G/e}(F-e)}(y-1)^{|F-e|-r_{G/e}(F-e)}$$

$$= \tilde{T}(G-e) + \tilde{T}(G/e)$$

which agrees with case (d) of (1.2).

Second, suppose that e is a cut-edge. Rewrite $\mathbf{T}(G;\,x,y)$ as a sum over all $F\subseteq E-\{e\}$ by by pairing the summands for F and $F+e$. For all such F, we have $|F+e|=|F|+1$ and $r(F+e)=1+r(F)$ (the latter because e is a cut-edge). Therefore

$$\begin{aligned}
\mathbf{T}(G;\,x,y) &= \left[\sum_{F\subseteq E-\{e\}} (x-1)^{r(G)-r(F)}(y-1)^{|F|-r(F)} + (x-1)^{r(G)-r(F+e)}(y-1)^{|F+e|-r(F+e)}\right]\\
&= \sum_{F\subseteq E-\{e\}} \left[(x-1)^{r(G)-r(F)}(y-1)^{|F|-r(F)} + (x-1)^{r(G)-r(F)-1}(y-1)^{|F|-r(F)}\right]\\
&= \sum_{F\subseteq E-\{e\}} \left[(x-1)^{r(G)-r(F)-1}(y-1)^{|F|-r(F)}\right]\left[(x-1)+1\right]\\
&= x\ \cdot \sum_{F\subseteq E-\{e\}} (x-1)^{r(G)-r(F)-1}(y-1)^{|F|-r(F)}. \qquad (3.4)
\end{aligned}$$

Now $E-\{e\}=E(G/e)$, and for every subset F, we have $r_G(F)=r_{G/e}(F)$. Also, $r(G)=r(G/e)+1$. So (3.4) becomes

$$x\ \cdot \sum_{F\subseteq E(G/e)} (x-1)^{r(G/e)-r(F)}(y-1)^{|F|-r(F)}.$$

By induction, this is just $x\cdot T(G/e;\,x,y)$. This agrees with case (b) of the recurrence (1.2).

The case that e is a loop is similar, and is left as an exercise. □

We now drop the notation $\mathbf{T}(G)$ and use $T(G;\,x,y)$ for both the polynomial defined recursively by (1.2), and the same polynomial defined in closed form by (3.3).

Corollary 3.2. *The polynomial $T(G;\,x,y)$ has nonnegative integer coefficients*

Proof. By induction on $e(G)$, using the recurrence (1.2). □

4 A few applications

All the other deletion-contraction invariants that we know about (see (1.1)) can be obtained as *specializations* of the Tutte polynomial—that is, by setting the parameters x and y to other values. In some cases, such as the chromatic polynomial, we may need the additional correction of multiplying by a monomial (because $T(G;\, x, y)$ does not keep track of the number of vertices).

4.1 Spanning forests

Theorem 4.1. *The number $\tau(G)$ of spanning forests of G is*

$$\tau(G) = T(G;\, 1, 1).$$

At the end of Example 2.6), we observed that this formula holds for the graphs K_3 and K_4. It is also easily seen to hold for forests by Example 2.1. The formula can be proved using either the recurrence (1.2) or the closed form (3.3).

Proof of Theorem 2. Plugging $x = y = 1$ into (1.2), we find that

$$T(G;\, 1,1) \;=\; \begin{cases} 1 & \text{if } E(G) = \emptyset, \\ T(G/e;\, 1,1) & \text{if } e \text{ is a cut-edge}, \\ T(G-e;\, 1,1) & \text{if } e \text{ is a loop}, \\ T(G-e;\, 1,1) + T(G/e;\, 1,1) & \text{otherwise.} \end{cases}$$

This is precisely the recurrence defining $\tau(G)$. □

Another proof of Theorem 2. Plug $x = y = 1$ into (3.3). It looks as though this will kill every term, but actually some of the terms—namely, those with both $r(G) - r(F) = 0$ and $|F| - r(F) = 0$—are identically 1, and will be unaffected by the substitution $x = y = 1$. Every other term will indeed be killed. Therefore

$$T(G;\, 1,1) \;=\; \#\{F \subseteq E \mid r(G) = r(F),\; |F| = r(F)\}. \tag{4.1}$$

But $r(F) = r(G)$ if and only if F contains a spanning forest of G, and $r(F) = |F|$ if and only F is acyclic. Hence the edge sets F counted in (4.1) are precisely the spanning forests of G. □

By similar arguments, the closed formula (3.3) implies that

$$\begin{aligned} T(G;\, 2,1) &= \text{number of acyclic subgraphs of } G, \\ T(G;\, 1,2) &= \text{number of spanning subgraphs of } G, \\ T(G;\, 2,2) &= 2^{e(G)}. \end{aligned}$$

4.2 The chromatic polynomial and acyclic orientations

Theorem 4.2. *The chromatic polynomial of G is*

$$\chi(G;\, k) \;=\; (-1)^{n(G)-c(G)}\, k^{c(G)}\, T(G;\, 1-k, 0). \tag{4.2}$$

To prove this, one can use the recurrence (1.2) to show that the right-hand side of (4.2) is given by the chromatic recurrence (West, Thm. 5.3.6). In particular, if $n(G) = n(H)$ and $T(G;\, x, y) = T(H;\, x, y)$ then $\chi(g;\, k) = \chi(H;\, k)$. The converse is false—there exist two connected graphs on five vertices with the same chromatic polynomial but different Tutte polynomials.

Substituting $k = -1$ in (4.2) yields

$$\chi(G;\, -1) \;=\; (-1)^{n(G)} T(G;\, 2, 0).$$

By Stanley's theorem (West, Thm. 5.3.27), it follows that $T(G;\, 2, 0)$ is the number of acyclic orientations of G.

5 A spanning tree expansion of $T(G)$

We have observed that $T(G;\, 1, 1) = \tau(G)$ and that $T_G(x, y)$ has nonnegative integer coefficients. This suggests that $T(G)$ might be given by a sum of monomials corresponding to spanning trees: that is, for each spanning tree T there are nonnegative integers $a(T)$, $b(T)$ such that

$$T(G;\, x, y) = \sum_{\text{spanning trees } T} x^{a(T)} y^{b(T)}. \tag{5.1}$$

This is equivalently to expanding the Tutte polynomial as

$$T(G;\, x, y) = \sum_{i \geq 0,\, j \geq 0} t_{ij} x^i y^j \tag{5.2}$$

(note that this sum is finite!), where

$$t_{ij} = \#\{\text{spanning trees } T \text{ with } a(T) = i \text{ and } b(T) = j\}. \tag{5.3}$$

So, what are $a(T)$ and $b(T)$?

Order the edges $E = E(G)$ as $e_1,\ e_2,\ \ldots,\ e_s$, and let T be a spanning tree. An edge $e_i \in T$ is said to be *internally active* (with respect to T) if it is the smallest edge of the edge cut between the two components of $T - e_i$. Equivalently,

$$e_j \notin T,\ T - e_i + e_j \text{ a tree} \quad \Longrightarrow \quad j \geq i.$$

An edge $e_j \in E - T$ is said to be *externally active* (with respect to T) if it is the smallest edge of the unique cycle in $T + e_j$. Equivalently,

$$e_i \in T,\ T - e_i + e_j \text{ a tree} \quad \Longrightarrow \quad i \geq j.$$

Theorem 5.1. *Fix an ordering of E. For each spanning tree T, let $a(T)$ be the number of internally active edges of T and let $b(T)$ be the number of externally active edges. Then the formula* (5.1) *holds.*

We omit the proof, which involves verifying the recurrence (1.2) and is straightforward if tedious.

For a particular tree T, the numbers $a(T), b(T)$ depend on the choice of ordering, hence are not isomorphism invariants of T (or of the pair T, G). However, the numbers t_{ij} do not depend on such a choice, since they are isomorphism invariants of G (by virtue of being coefficients of the Tutte polynomial). It is rather surprising that all $e(G)!$ possible orderings of the edges produce the same numbers t_{ij}.

In some cases, the coefficients t_{ij} have combinatorial interpretations independent of (5.3), and are known to satisfy certain equalities. For example, $t_{01} = t_{10}$ for all graphs; this number is called the *chromatic invariant* $\theta(G)$. It can be shown $\theta(G)$ depends only on the *homeomorphism type* of G (for example, (2.3) implies that $\theta(C) = 1$ for all cycles C). Moreover, there is the amazing fact

$$\theta(G) = |\chi'_G(1)|,$$

where $\chi'(G)$ is the *derivative* of the chromatic polynomial.

6 Further references

The foregoing is just the tip of the iceberg. The Tutte polynomial can be used to count colorings of G by the number of improper edges, to obtain information on group-valued flows on G, to compute the probability that a random subgraph of G is connected, and to study knot theory, statistical mechanics, and topology. It is also one of the most useful tools in studying *matroids*, which are more general combinatorial objects than graphs. There are still many open problems concerning the Tutte polynomial; for instance, there is much more to be said about the combinatorial meaning of the coefficients t_{ij}.

The canonical reference on the Tutte polynomial is the survey article "The Tutte polynomial and its applications" by T. Brylawski and J. Oxley, which appears as pp. 123–225 of *Matroid applications*, N. White, ed. (Cambridge University Press, 1992). Another nice exposition is the final chapter of *Modern Graph Theory* by B. Bollobás (Graduate Texts in Mathematics 184, Springer, 1998). W.T. Tutte first constructed the polynomial (in a slightly different but equivalent form) in "A contribution to the theory of chromatic polynomials", *Canadian J. Math.* **6** (1954), 80–91, and developed some of its basic properties in "On dichromatic polynomials", *J. Comb. Theory* **2** (1967), 301–320.