1 Introduction

2 Rooted Trees

A *rooted tree* is a triple $t = (r, V, E)$, such that (V, E) is a finite, simple connected graph without cycles with vertex set V and edge set E. A vertex, $r \in V$, is selected from V and called the *root*. By convention each edge is directed away from r. We denote the order of a tree, t, by $|t|$ and the set of order n rooted trees is denoted $\mathcal{R}_n$. In addition we define $R := \bigcup R_n$.

A *labelled rooted tree* is a quadruple $t = (r, V, E, L)$ such that (r, V, E) is a rooted tree and

$$L : V \longrightarrow \{1, 2, \ldots, |t|\}$$

is a bijective map. We denote the set of order n labelled rooted trees by $\mathcal{L}_n$ and we define $\mathcal{L} = \bigcup \mathcal{L}_n$.

For any pair of vertices u and v of a rooted tree we write $u \leq v$ if u lies on the unique shortest path from r to v. A *random recursive tree* is a quadruple $t = (r, V, E, l)$ such that t is a labelled rooted tree and l is a labelling such that if $u \leq v$ then $l(u) < l(v)$. We denote the set of order n random recursive trees by $\mathcal{T}_n$ and we define $\mathcal{T} = \bigcup \mathcal{T}_n$.

Two rooted trees t_1 and t_2 with roots r_1 and r_2 and vertex sets $V(t_1)$ and $V(t_2)$ respectively are said to be isomorphic if there exists a bijection $f : V(t_1) \to V(t_2)$ such that vertices $u, v \in V(t_1)$ are adjacent if and only if $f(u), f(v) \in V(t_2)$ are adjacent and $f(r_1) = r_2$. If $V(t_1) = V(t_2)$ then f is called a rooted tree automorphism. The set of automorphisms of a tree, t, together with composition of maps forms a group denoted $\mathrm{Aut}(t)$. The order of the automorphsim group is

$$\sigma(t) := |\mathrm{Aut}(t)|$$

Consider a map $\phi : \mathcal{L} \to R$ that simply forgets the labels of a labelled by mapping

$$(r, V, E, L) \mapsto (r, V, E)$$

The map ϕ is clearly surjective but not injective hence we define $\beta(t) = |\phi^{-1}(t)|$ to be the number of possible isomorphism classes of labellings of a rooted tree $t \in \mathcal{R}$. There are $|t|!$ possible labellings of a tree $t \in \mathcal{R}$ hence there are

$$\beta(t) = \frac{|t|!}{\sigma(t)} \tag{1}$$

isomorphism classes of labellings of t.

Since $\mathcal{T} \leq \mathcal{L}$ we denote by ψ the restriction of ϕ to random recursive trees and define

$$\alpha(t) := |\psi^{-1}(t)|$$

however, to provide a result analagous to Equation 1 we require several additional definitions.

Let $t_1, t_2, \ldots, t_k$ be a forest of rooted trees: the rooted tree $B^+(t_1, t_2, \ldots, t_k)$ is built from this forest by introducing a new vertex, r (the root of $B^+(t_1, t_2, \ldots, t_k)$), and joining the roots of each tree $t_1, t_2, \ldots, t_k$ to r via an edge. Since every tree t can be written as $B^+(t_1, t_2, \ldots, t_k)$ if $|t| > 1$ we henceforth assume that each rooted tree $t = B^+(t_1, t_2, \ldots, t_k)$ if $|t| > 1$. For conveniance we will denote the rooted tree on 1 vertex by $\bullet$.

We have already seen one function $f : \mathcal{R} \to \mathbb{R}$ on rooted trees ($f(t) = |t|$), let us consider another: We define the tree factorial $t!$ recursively

$$\bullet! = 1 \tag{2}$$

$$t! = |t| \prod_{i=1}^{k} t_i! \tag{3}$$

Example 2.1.

Lemma 2.2. *For a rooted tree t,*

$$\alpha(t) = \frac{|t|!}{t!\sigma(t)} \tag{4}$$

Proof. There are $n!$ ways of labelling a tree $t \in \mathcal{R}_n$, however if l is a random recursive labelling every induced subtree t_v and a totally ordered set, S, of labels there is precisely one possible label $s \in S$ for vertex v (namely $s = \min(S)$). Therefore the factor we should divide out by is precisely $t!$. In addition, to calculate the number of isomorphism classes of random recursive trees we must again divide out by $\sigma(t)$. $\square$

Finally, let $\chi : \mathcal{F} \to \mathcal{R}$ be the map that "forgets" the embedding of a rooted plane tree. It is clear that χ is surjective but not injective hence we define

$$\gamma(t) = |\chi^{-1}|$$

the number of isomorphism classes of embeddings of a rooted tree t. In order to obtain a third relation analagous to Equations 1 and 4 we will describe a third function $w : \mathcal{R} \to \mathbb{R}$ recursively:

$$w(\bullet) = 1 \tag{5}$$

$$w(t) = k! \prod_{i=1}^{k} w(t_i) \tag{6}$$

Example 2.3. Do the same example set as before

Lemma 2.4.

$$\gamma(t) = \frac{w(t)}{\sigma(t)} \tag{7}$$

Proof. Let $B^+(t_1^{n_1}, t_2^{n_2}, \ldots, t_k^{n_k})$ denote a tree in which the root is incident to n_1 isomorphic copies of a tree t_1, n_2 isomorphic copies of t_2 and so forth. We remark that every rooted tree with at least 2 vertices can be written in this way. Let $t \in R$ and consider an induced subtree $t_v = B^+(t_1^{n_1}, t_2^{n_2}, \dots, t_k^{n_k})$. Vertex v contributes a factor of $\left(\sum_{i=1}^k n_i\right)!$ to $w(t)$ since it has a recursive definition. The order of automorphism group, $\sigma(t)$ can also be expressed recursively:

$$\sigma(\bullet) = 1 \tag{8}$$

$$\sigma(t) = n_1!n_2!\dots n_k! \prod_{i=1}^{k} \sigma(t_i)^{n_i} \tag{9}$$

Therefore vertex v contribute a factor of $\prod_{i=1}^k n_i!$ to $\sigma(t)$. Finally consider, $\gamma(t)$, the number of somorphism classes of embeddings of t. There are

$$\frac{\left(\sum_{i=1}^k n_i\right)!}{\prod_{i=1}^k n_i!}$$

non-isomorphic possibilities for the ordering of the children of v. $\square$

3 Automorphisms of Trees

In this section we will describe a direct product decomposition of the automorphism group of a tree, t, in which factors of the direct product can be associated with particular induced subtrees of t.

Recall that a (rooted) tree automorphism is a permutation of vertices that preserves adjacency and the root. It is a result of Pólya that automorphism groups of trees belong to the class, $\mathcal{W}$, of permutation groups which contains the symmetric groups and is closed under taking direct and wreath products. Let t be a rooted tree and consider the following decomposition,

$$\mathrm{Aut}(t) \cong A_1 \times A_2 \times \dots A_p \times B_1 \times B_2 \times \dots \times B_q \tag{10}$$

where each factor A_i is isomorphic to a symmetric group and each B_i is isomorphic to the wreath product of a group $G \in \mathcal{W}$. In [] MacArthur, Sanchez-Garcia and Anderson showed that the decomposition of $\mathrm{Aut}(t)$ described in Equation has a geometric interpretation. An automorphism group, $\mathrm{Aut}(t)$ may be decomposed by partitioning the set of generators S of $\mathrm{Aut}(t)$ into support-disjoint subsets $S = S_1 \cup S_2 \cup \dots \cup S_r$ and writing

$$\mathrm{Aut}(t) = H_1 \times H_2 \times \dots \times H_r \tag{11}$$

where each H_i is generated by S_i. This, *geometric decomposition* is shown [] to be unique and irreducible (each H_i cannot be written as the direct product of support-disjoint subgroups) hence the geometric decomposition is well defined. We call each H_i in the geometric decompostion of $\mathrm{Aut}(t)$ a *geometric factor*. A symmetric subtree is the induced subtree of t on the support of a geometric factor.

Example 3.1. (i) A k-star is an induced subtree consisting of a vertex adjacent to k vertices of outdegree 0. A k-star is a symmetric subtree that corresponds to a geometric factor S_k (the symmetric group on k objects).

(ii)

By Pólya's Theorem the geometric decomposition (Equation 11) can be written in the form,

$$\mathrm{Aut}(t) \cong A_1 \times A_2 \times \dots A_p \times B_1 \times B_2 \times \cdots \times B_q \tag{12}$$

such that each A_i corresponds to a symmetric subtree.

There is a natural way to split the geometric decomposition of $\mathrm{Aut}(t)$ into two subgroups. We define the direct product of symmetric groups to be the *elementary subgroup*:

$$\mathcal{E}(t) = A_1 \times A_2 \times \cdots \times A_p$$

The direct product of wreath products of symmetric groups form the *complex subgroup*:

$$\mathcal{C}(t) = B_1 \times B_2 \times \cdots \times B_q$$

The order, $\sigma(t)$, of an automorphism group can also be split as follows:

$$\sigma(t) = |\mathcal{E}(t)||\mathcal{C}(t)|$$

This begs the question: does the order of either the elementary or the complex subgroup dominate the other? MacArthur [**?**] made the following additional conjecture:

Conjecture 3.2. *Let $\{T_t\}_{t=1}^{n}$ be a RRT. In the limit as $t \to \infty$, $|\mathcal{E}(T_t)|^{\frac{1}{t}} = \mathcal{V}$, while in the limit as $t \to \infty$, $|\mathcal{C}(T_t)|^{\frac{1}{t}} = 1$.*

We claim that the elementary subgroup captures the contribution that (n, k)-stars make to the automorphism group and the complex subgroup captures the contribution that the extended symmetric branches make to the automorphism group.

4 Functions on Trees

In this section we will build up a 3-part tool kit of functions $f : \mathcal{R} \to \mathbb{R}$ which will be used to calculate properties of rooted trees such as path length and order.

4.1 Inductive Maps

Definition 4.1. Let $s = \{s_r\}_{r=0}^{\infty}$ be a sequence such that each $s_r \in \{0, 1\}$. A *variety*, $\mathcal{V}$, of trees is a collection of random recursive trees such that each vertex is permitted to have outdegree r only if $s_r = 1$.

Remark 1. *For a more general setting see [**?**].*

The *degree function* associated with a sequence $s = \{s_r\}_{r=0}^{\infty}$ is the exponential generating function (EGF) defined as follows:

$$\phi(w) = \sum_{r \geq 0} s_r \frac{w^r}{r!}$$

Example 4.2. The collection, $\mathcal{B}$, of increasing binary trees are the set of all random recursive trees such that each vertex has outdegree either 0,1 or 2. The degree function for increasing binay trees is

$$\phi(w) = 1 + w + \frac{w^2}{2}$$

The degree function for random recursive trees is

$$\phi(w) = \exp(w).$$

Fix a variety, $\mathcal{V}$, and define V_n to be the number of trees of order n in $\mathcal{V}$. The EGF of the variety of trees is

$$V_{\mathcal{V}}(z) = \sum_{n \geq 1} V_n \frac{z^n}{n!}$$

For example,

$$V_{\mathcal{T}}(z) = \sum_{n \geq 1} \frac{z^n}{n} \tag{13}$$

$$= \log\left(\frac{1}{1-z}\right) \tag{14}$$

since the number of random recursive trees of order n is $(n-1)!$.

Definition 4.3. Let $f = \{f_n\}_{n \geq 1}$ be a sequence of real numbers. A function $s : \mathcal{R} \to \mathbb{R}$ is called an *inductive map* if it is defineable by a relation,

$$s(t) = f_{|t|} + \sum_{i=1}^{k} s(t_i)$$

where $t = B^+(t_1, t_2, \ldots, t_k)$.

Given a tree function s and a variety $\mathcal{V}$ the EGF of s over $\mathcal{V}$ is

$$S(z) = \sum_{t \in \mathcal{V}} s(t) \frac{z^{|t|}}{|t|!}$$

Example 4.4. Let $\delta_{i,j}$ be the usual kronecker delta. The following is a list of possible sequences f and the tree paramater measured by the corresponding function s

(i) $f_n = 1$ for all n counts tree size.

(ii) $f_n = \delta_{n,i}$ counts the number of induced subtrees of order i.

(iii) $f_n = n$ counts path lengths.

In order to calculate $S(z)$ easily and effectively we appeal to the following Theorem of Bergeron [**?**].

Theorem 4.5.

$$S(z) = Y'(z) \int_0^z \frac{F'(t)}{Y'(t)} dt$$

where $F(z)$ is defined from $\mathcal{V}$ and the sequence f

$$F(z) = \sum_{n \geq 0} f_n V_n \frac{z^n}{n!}$$

4.2 Butcher Series

Throughout this section let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a suitably differentiable function and consider differential equations of the form

$$\frac{d}{ds} y(s) = f(y(s))$$

where $y(s_0) = y_0$. We use the notation $f'(y)$ for the derivative $\frac{d}{dy} f(y)$ and note that $f'(y)$ is a linear map (the Jacobian), the second derivative $f''(y)$ is a billinear map and so on.

Definition 4.6. Let $t = B^+(t_1, t_2, \ldots, t_k)$ be a rooted tree. An *elementary differential* is a map $F(t) : \mathbb{R}^n \to \mathbb{R}^n$ defined recursively by:

$$F(\bullet)(y) = f(y) \tag{15}$$

$$F(t)(y) = f^m(y)(F(t_1)(y), \ldots, f(t_k)(y)) \tag{16}$$

Theorem 1. *The solution of*

$$y(s) = y_0 + \int_{s_0}^{s} f(y(s'))ds'$$

is

$$y(s) = y_0 + \sum_{t \in \mathcal{R}} \frac{(s - s_0)^{|t|}}{|t|!} \alpha(t) F(t)(y_0)$$

Proof. See [**?**, **?**] □

Example 4.7. Let $y_0 = s_0 = 0$ and consider the case $f(s) = \frac{1}{1-s}$ so that $f^{(k)}(0) = k!$ then $F(t) = w(t)$.

4.3 Bare Green Functions

5 Bounds on the expected value of $\mathrm{Aut}(t)$

5.1 A lower bound

5.2 A disproof of Conjecture 3.2

5.3 An upper bound

In this section we will use B-series and Theorem **??** to calculate an upper bound for the expected order $\mathbb{E}_{\mathcal{T},n}(\sigma(t))$ of an automorphism group of a random recursive tree t.

Recall from equation 7 that $\gamma(t) = \frac{w(t)}{\sigma(t)}$ hence

$$\mathbb{E}_{\mathcal{T},n}(\sigma(t)) = \frac{1}{|\mathcal{T}_n|} \sum_{t \in R_n} \alpha(t)\sigma(t) \tag{17}$$

$$< \frac{1}{|\mathcal{T}_n|} \sum_{t \in R_n} \alpha(t)w(t) \tag{18}$$

We define $f(s) = \frac{1}{1-s}$ as in Example 4.7 so that $f^{(k)}(0) = k!$. By Theorem 1 the solution to

$$x(s) = \int_0^s = \frac{1}{1 - x(s')} ds' \tag{19}$$

is

$$x(s) = \sum_{t \in \mathcal{R}} \frac{s^{|t|}}{|t|!} \alpha(t)w(t) \tag{20}$$

We may restate Equation 19 as $x'(s) = \frac{1}{1-x(s)}$ which is a first order, non-linear differential equation with solution

$$x(s) = 1 \pm (c_1 - 2s + 1)^{\frac{1}{2}} \tag{21}$$

where c_1 is a constant to be determined. Since (by Equation 19) $x(0) = 0$ we may deduce that $c_1 = 0$. By using the binomial expantion of Equation 21 and comparing with Equation 20 we find that,

$$\sum_{|t|=n} \alpha(t)w(t) = \prod_{i=1}^{n-1} (2i - 1)$$

Note that $\prod_{i=1}^{n-1}(2i-1) = (2n-3)!!$ hence using factorial identities and Stirling's approximation

$$\frac{(2n-3)!!}{(n-1)!} = \frac{(2n-3)!}{(n-1)!^2 2^{n-1}} \tag{22}$$

$$= b \tag{23}$$

Therefore,

$$\mathbb{E}_{\mathcal{T},n}(\sigma(t)) < blah$$

By Jensen's inequality,

$$\mathbb{E}_{\mathcal{T},n}(\sigma(t)^{\frac{1}{n}}) < blah^{\frac{n-1}{n}} = \text{constant as required}$$

It is interesting to compare the expected value of $\sigma(t)^{\frac{1}{n}}$ with the typcial value of $\sigma(t)^{\frac{1}{n}}$ for random recursive trees since this will illuminate the nature of the distribution of automorphism group orders.

Consider the following alternative decription of random recursive trees as a growing tree process: A *random recursive tree*, $t \in \mathcal{T}_n$ can be written as a nested sequence of random recursive trees:

$$t = t^1 \subset t^2 \subset \cdots \subset t^n.$$

such that at initial time $s = 1$ the tree t^1 is the random recursive tree with 1 vertex and no edges and subsequently at times $s = 2, 3, \ldots, n$ a vertex, v, is chosen from the set vertices, $V^{(s-1)}$, of tree $t^{(s-1)}$ uniformly at random and a new vertex labelled t is attached to v via an edge.

Remark 2. *The interpretation, above, makes it clear why we name this family;* random recursive trees. *It also makes obvious the observation that* $|\mathcal{T}_n| = (n-1)!$ *since at time* s *there are* $s-1$ *possible vertices for* v *to be attached.*

Let $t = \{t^s\}_{s \geq 1}$ be a random recursive tree and $X_{s,i}$ be the number of vertices of outdegree $i-1$ in t^s.

Theorem 5.1. *In the limit as* $s \to \infty$*, almost surely* $\frac{X_{s,i}}{s} \to 2^{-(i+1)}$.

Proof. See Janson **??** □

6 Conclusions