# Project Suggestion

**Title:** Polytopes and consecutive patterns on trees

## Background

### Rooted planar trees

We will work with **rooted planar trees**. A rooted tree is a pair $(T, v)$ where $T$ is a tree and $v$ is a distinguished vertex of $T$.

For each vertex $w$ that is not the root $v$, there is a unique vertex $w'$ that is adjacent to $w$ and is closer to $v$. We call this vertex $w'$ the parent of $w$. Conversely, we say that $w$ is a **child** of $w'$. We denote the set of children of $w'$ as $\mathcal{C}(w')$. A vertex that has no children is callled a **leaf**. Finally, $w$ is said to be a **descendant** of $w'$ if the unique path between $v$ and $w$ contains $w'$.

A **rooted planar tree** is a triple $(T, v, f)$ where $(T, v)$ is a rooted tree such that each vertex $w$ is endowed with an order $f(w)$ on $\mathcal{C}(w)$. This should be thought of as *how to draw the tree*, *i.e.* in which order to draw the children of a vertex when putting down the tree in a piece of paper. The depth of a tree is the size of the longest path starting at the root. Equivalently, it is the largest distance of a leaf to the root. See Fig. 2. To refer to a rooted planar tree $(T, v, f)$ we simply write $T$.

Two rooted planar trees are equivalent if there is a graph morphism that preserves the root and the order of the children. Let $r, \delta \geq 1$ be integers. We define $\mathcal{F}^r_\delta$ as the set of rooted planar trees (simply **trees** in the following) of depth at most $r$ and where each vertex has at most $\delta$ children.

### Patterns in trees

Given a vertex $w$ on a planar rooted tree $T$, and an integer $r > 0$, we define the **ball of radius $r$ centered at $v$** as the planar rooted tree $T'$ resulting from $T$ by restricting to all the vertices that are descendant of $w$ and their distance to $w$ is at most $r$, and letting $w$ be the root. We denote this rooted planar tree by $B^{(r)}_T(w)$. This is also called the **fringe subtree of $T$ at $v$ (at depth $r$)**.

Given two planar rooted trees $T_1$ and $T_2$ and an integer $r > 0$, the number of $r$-occurrences of $T_2$ in $T_1$ is the number of vertices $v$ of $T_1$ such that $B^{(r)}_{T_1}(v) = T_2$. This number is denoted by $\text{c-occ}^{(r)}(T_2, T_1)$, and let $\widetilde{\text{c-occ}}^{(r)}(T_2, T_1) = \frac{\text{c-occ}^{(r)}(T_2,T_1)}{|T_1|}$ denote

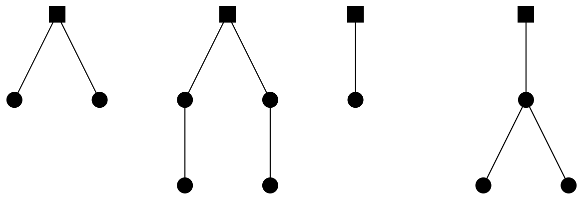

Figure 1: Examples of rooted trees, where a root is represented as a square vertex. Two of height 1, and two of height 2.

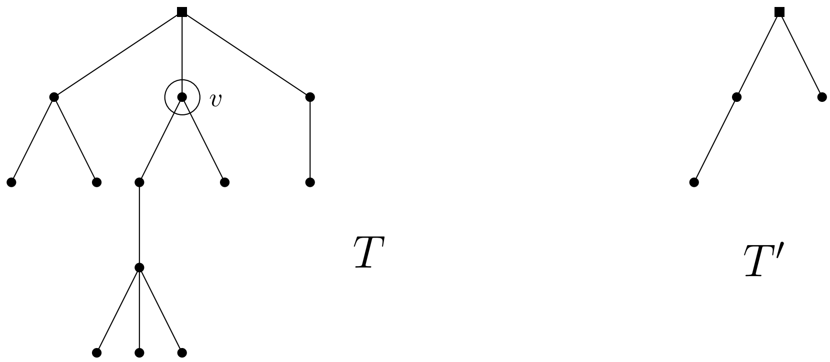


Figure 2: On the right a tree $T$ and a vertex $v$, corresponding to the pattern of radius 2 of the tree $T'$, on the left.

the proportion of such occurrences. For instance, the number of occurrences of the tree with one vertex is exactly the number of leaves.

Fixed integers $r, \delta$, given a tree we denote $\widetilde{\text{c-occ}}_{\delta}^{(r)}(T_1)$ to be the vector in $\mathbb{R}^{\mathcal{F}_{\delta}^{r}}$ that in the coordinate indexed by a tree $T_2$ has the value $\widetilde{\text{c-occ}}^{(r)}(T_2, T_1)$. Finally, we define the set of all points that are feasible by trees as

$$FR_{\delta}^{(r)} = \{\nu \in [0,1]^{|\mathcal{F}_{\delta}^{r}|} | \exists (T^{(n)}) \text{ s.t. } \widetilde{\text{c-occ}}_{\delta}^{(r)}(T^{(n)}) \to \nu,\, |T^{(n)}| \to \infty\}$$

**Polytope conjecture.** *The feasible region $FR_{\delta}^{(r)}$ is always a polytope.*

# Aim

The goal is for you to give strong evidence for this conjecture for the case $\delta = r = 2$. Notice that in this case you are going to be working in a 13-dimentional vector space, so don't try to draw it!

We conjecture that this is a polytope that satisfies the following equations: if $(a_T)_{T \in \mathcal{F}_{\delta}^{(r)}} \in FR_{\delta}^{(r)}$, then by using the notation described in Fig. 3 to refer to the

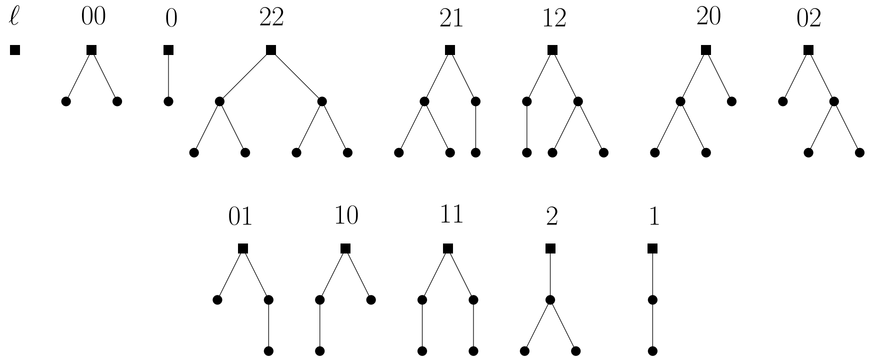


Figure 3: All distinct trees of at most depth two, along with some helpful notation.

different trees of depth at most two, we have

$$
\begin{aligned}
\sum_{T \in \mathcal{F}_\delta^{(r)}} a_T &= 1\,, \\
\sum_{T \in \mathcal{F}_\delta^{(r)}} \deg(T) a_T &= 1\,, \\
a_l &= a_0 + a_{01} + a_{02} + a_{20} + a_{10} + 2a_{00}\,, \\
a_0 + a_2 &= a_{01} + a_{21} + a_{10} + a_{12} + 2a_{11}\,, \\
a_2 + a_{22} &= a_{00} + a_{01} + a_{10} + a_{11}\,,
\end{aligned}
\tag{1}
$$

and furthermore, $a_T \geq 0$ for any tree $T \in \mathcal{F}_\delta^{(r)}$.

# Method

A suggestion is to split the project into two parts. First you will devise a method to generate big trees randomly, and computing its corresponding vector. This should give you a cloud of points that are close to the feasible region. Presumably, the method of generating these trees should be general and robust enough so that this cloud of points really covers the whole feasible region.

The second part of the project is to check if there are other equations that the feasible region satisfies. This can be done by computing the vertices of the polytope that arise from Eq. (1), and then figuring out if there are points in the feasible region arbitrarily close to these vertices.

# Requirements

You are strongly encouraged to use a computer to aid you in this project, in whichever language you feel comfortable with. Just make sure it has a robust geometry package, as it will be handy to go from the H-polytope description to the V-polytope description. Some suggestion is the python packages pycddlib, see [1].

# References

[1] Matthias Troffaes, https://pypi.org/project/pycddlib/