Rigorous Development of Sylvaris: Mathematical Study of Tree-like Branching Patterns

Pu Justin Scarfy Yang

July 25, 2024

Abstract

Sylvaris is a newly proposed field of mathematics dedicated to the study of tree-like, branching patterns and their mathematical models. This paper rigorously develops the foundational notations, definitions, and formulas of Sylvaris, providing a comprehensive framework for exploring these complex structures.

1 Introduction

Sylvaris focuses on the properties and dynamics of branching structures, which can be seen in various mathematical and real-world phenomena, such as fractals, phylogenetic trees, and decision trees.

2 Notations and Definitions

- **Tree Structure**: A tree T is a connected acyclic graph. A tree can be represented as $T = (V, E)$, where V is the set of vertices (nodes) and E is the set of edges.
- **Rooted Tree**: A rooted tree is a tree in which one vertex is designated as the root. The root is typically denoted by r.
- **Height of a Tree**: The height $h(T)$ of a tree T is the length of the longest path from the root to any leaf. If $d(v)$ denotes the depth of vertex v, then:

$$h(T) = \max_{v \in V} d(v)$$

- **Branching Factor**: The branching factor $b(v)$ at a vertex v is the number of children of v. For a tree T with vertices V:

$$b(v) = \deg(v) - 1 \quad \text{for } v \neq r, \quad b(r) = \deg(r)$$

- **Fractal Dimension**: For fractal-like branching structures, the fractal dimension D can be defined using the similarity dimension:

$$D = \frac{\log(N)}{\log(S)}$$

where N is the number of self-similar pieces and S is the scaling factor.

- **Phylogenetic Tree**: A phylogenetic tree is a branching diagram showing the inferred evolutionary relationships among various biological species or entities.

3 Formulas and Models

- **Tree Growth Model**: A simple model for tree growth can be described recursively. Let T_n represent the tree at stage n:

$$T_{n+1} = T_n \cup \bigcup_{v \in \text{Leaves}(T_n)} \{(v, v_i) \mid i = 1, 2, \ldots, k\}$$

where k is the fixed branching factor.

- **Recursive Branching Process**: Consider a branching process where each node has a random number of children according to a probability distribution P. The expected size of the tree $E(T_n)$ after n generations can be given by:

$$E(T_n) = \sum_{i=0}^{n} \left(\prod_{j=0}^{i-1} E(b(v_j)) \right)$$

- **Optimal Branching Strategy**: For decision trees, an optimal branching strategy can minimize the expected cost. Let $C(v)$ be the cost function for a node v. The optimal cost $C^*(v)$ can be found using dynamic programming:

$$C^*(v) = \min_{b \in B(v)} \left(c(v, b) + \sum_{v' \in \text{children}(v,b)} C^*(v') \right)$$

where $B(v)$ is the set of possible branching choices at node v and $c(v, b)$ is the immediate cost of branching choice b.

4 New Mathematical Notations and Formulas

- **Branch Weight Function**: Define a branch weight function $w : V \times V \to \mathbb{R}^+$ to assign weights to branches:

$$w(u, v) = \text{weight of branch } (u, v)$$

- **Total Tree Weight**: The total weight $W(T)$ of a tree T is the sum of the weights of all branches:

$$W(T) = \sum_{(u,v) \in E} w(u, v)$$

- **Branching Entropy**: For a branching process with probabilities p_i for each branch, define the branching entropy H_b as:

$$H_b = -\sum_i p_i \log(p_i)$$

- **Branch Length Distribution**: If L denotes the length of a branch, the distribution of branch lengths $P(L)$ can be modeled using a probability density function $f_L(l)$:

$$P(L \leq l) = \int_0^l f_L(x)\, dx$$

5 Applications and Examples

- **Fractals and Self-Similarity**: Tree-like structures can be used to model fractals. For example, the fractal dimension of a self-similar tree can be analyzed using the aforementioned fractal dimension formula.
- **Phylogenetics**: The study of phylogenetic trees involves understanding evolutionary relationships. Algorithms such as UPGMA (Unweighted Pair Group Method with Arithmetic Mean) and neighbor-joining are used to construct phylogenetic trees.
- **Decision Trees in Machine Learning**: Decision trees are used for classification and regression tasks. The optimal branching strategy helps in minimizing the prediction error or cost associated with decision nodes.

6 Conclusion

Sylvaris offers a rich and versatile framework for understanding and modeling tree-like, branching structures. By formalizing these concepts with rigorous mathematical notations and formulas, we open new avenues for research and application in various disciplines.

References

[1] B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman and Co., 1982.

[2] J. Felsenstein, *Inferring Phylogenies*, Sinauer Associates, 2004.

[3] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, *Classification and Regression Trees*, Wadsworth, 1984.