MATCH
Communications in Mathematical and in Computer Chemistry

no. 51, April 2004
MATCDY (51) 79- 95 (2004)

ISSN 0340 - 6253

Fibonacci Trees: A Study of the Asymptotic Behavior of Balaban's Index

Ning Jia
School of Mathematics, University of Minnesota
206 Church St. SE, Minneapolis, MN 55455

and

K. W. McLaughlin
Department of Chemistry
University of Wisconsin-River Falls, WI 54022

(Received September 1, 2003)

Abstract

The asymptotic properties of the Balaban index for Fibonacci trees are analyzed. While this topological index diverges for highly branched molecular graphs, the restricted type of branching exhibited by Fibonacci trees allows the origin of this divergence to be identified and corrected for. Further it is demonstrated that the nature of the branching pattern determines the scaled asymptotic limit for this distance matrix based topological index, rather than the initial graph from which the branching emanates. The same type of analysis is illustrated for a typical dendrimer.

1 Introduction

Topological indices have been used extensively for the prediction of physical properties of specific classes of molecules [1]. The two most important distance matrix based topological indices are the Wiener index [2] and the Balaban index [3]. The asymptotic behavior of the Balaban index for linear polymers with well defined (i.e. finite and repetitive) branching patterns along the main chain has been previously analyzed [4]. Unlike the Wiener index which diverges [2], the behavior of the Balaban index mimics the behavior of the melting temperatures and glass transition temperatures for linear macromolecules, which possess an asymptotic limit for these physical properties [5]. Unfortunately, for highly branched molecular graphs (i.e. for structures in which the branching pattern involves the formation of branches off other branches) the Balaban index diverges too.

Polymer molecules with complex branching patterns represent a class of molecular graphs which are known to require special treatment for topological analysis [6,7]. Understanding the asymptotic behavior of a topological index is critical to successfully modeling the physical properties of highly branched macromolecules. Therefore, in order to identify the effect of branching on the Balaban index for highly branched graphs, an analysis of the distance matrix for a class of molecular graphs possessing a well-defined and restricted branching pattern has been undertaken.

It was decided to restrict the branching pattern in a way which would generate a molecular graph that contains some substructure, but was not symmetrical. Therefore, we have undertaken an asymptotic analysis of Fibonacci trees. This is a special case of Fibonacci graphs which have been previously defined [8]. The Balaban index for hydrocarbons is given by

$$J_G = \frac{Q}{C+1} \sum_{i,j} \frac{1}{\sqrt{d_i d_j}}, \tag{1}$$

where G is the molecular graph, Q is the number of edges in G, C is the number of cycles in G (the cyclomatic number) which in the case of trees is zero, and d_i and d_j are the row sums in the distance matrix for vertices i and j, respectively [3], and the sum is taken over all edges i,j. It is how this sum diverges as the molecular graph grows that we are primarily interested in understanding. The following analysis illustrates the key aspect to the behavior of the sum in equation (1) and how it should be scaled to produce a non-zero finite limit that is directly related to the branching pattern in the molecular graph.

2 Definitions

The Fibonacci series is recursively defined to be $f_0 = 1$, $f_1 = 1$, and $f_n = f_{n-1} + f_{n-2}$ for $n = 2, 3, \ldots$ The Fibonacci trees are unlabeled binary trees that can also be defined recursively as F_0 = empty tree, F_1 = a single vertex, and F_n is a tree with F_{n-1} and F_{n-2} as left and right subtrees. It is easy to see that F_n has $f_{n+1} - 1$ vertices.

We now define labeled trees T_n for all $n \geq 1$ (Figure 1), which is the first molecular graph under discussion. We will show later that T_n has the same tree structure as F_n. But we define T_n using a different recurrence relation so that its vertices will be labeled by integers 1 through $f_{n+1} - 1$. This labeling has nice properties that will greatly facilitate later discussions. The vertices will also be identified by their labels in cases where there is no confusion. For reasons that we will see later on, we will define T_{n-1} instead of T_n.

PROPOSITION 2.1

(Construction of T_n) *Let T_1 = a single vertex labelled by 1. Then we are able to get T_{n-1} for $n \geq 3$ by adding vertices $f_{n-1}, f_{n-1}+1, \ldots, f_n - 1$ to T_{n-2} according to the following:*

1. *If $f_{n-1} \le j \le 2f_{n-2} - 1$, attach a new vertex labeled by j to vertex $j - f_{n-2}$ (which is in T_{n-2}) as its right child;*

2. *If $2f_{n-2} \le j \le f_n - 1$, attach a new vertex labeled j to vertex $j - f_{n-2}$ as its left child.*

This means, vertex $j - f_{n-2}$ is a leaf in T_{n-2}.

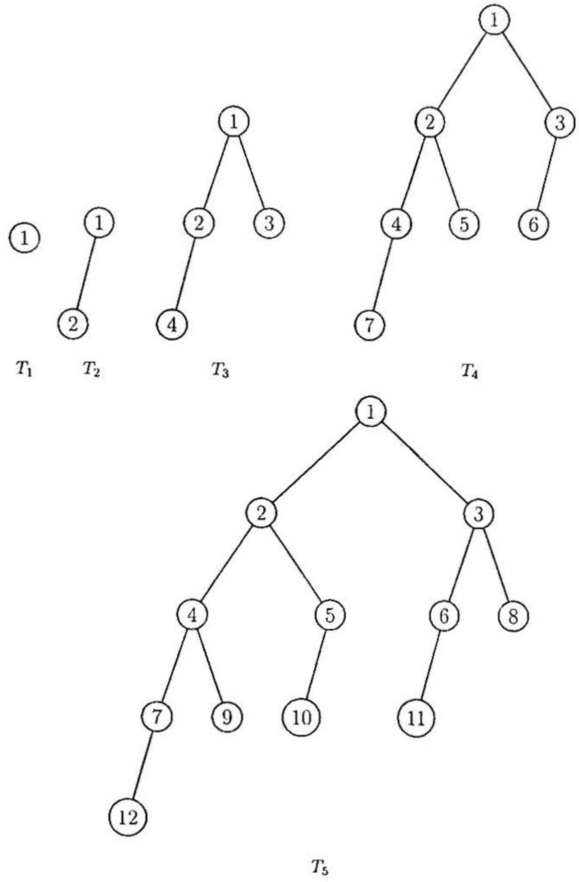

Figure 1: Examples of T_n

Proof.

We need to prove that in cases 1 and 2, before attaching the vertex labeled by j to vertex $j - f_{n-2}$ as its left or right child, vertex $j - f_{n-2}$ does not already have a left or right child. In other words, we need to prove the construction method is well-defined.

We prove the Proposition by induction on n.

The case when $n = 3$ is trivial. Suppose we can construct trees up to T_{k-2}, or, for all $(n-1) \leq (k-2)$ according to the above rules. So now we have vertices 1 through $f_{k-1} - 1$ in the tree T_{k-2}, and we need to attach vertices $f_{k-1}, \ldots, f_k - 1$ to T_{k-2} to get T_{k-1}.

1. When $f_{k-1} \leq j \leq 2f_{k-2} - 1$, or, $f_{k-3} \leq j - f_{k-2} \leq f_{k-2} - 1$, according to the construction scheme, $j - f_{k-2}$ must have been added to T_{k-4} to form T_{k-3}. Also, $f_{k-2}, \ldots, 2f_{k-3} - 1$ were attached to $f_{k-4}, \ldots, f_{k-3} - 1$ as their right children; and $2f_{k-3}, \ldots, f_{k-1} - 1$ were attached to $f_{k-3}, \ldots, f_{k-2} - 1$ as their left children, to form T_{k-2} from T_{k-3}. But $f_{k-3} \leq j - f_{k-2} \leq f_{k-2} - 1$, so vertex $j - f_{k-2}$ has no right child in T_{k-2}. The recurrence rule (1) is workable.

2. When $2f_{k-2} \leq j \leq f_k - 1$, we have $f_{k-2} \leq j - f_{k-2} \leq f_{k-1} - 1$, from the above argument, we know $j - f_{k-2}$ has no children in T_{k-2}. The recurrence rule (2) is workable.

□

COROLLARY 2.2

For any vertex i such that $f_{k-1} \leq i \leq f_k - 1$, if i has left and/or right child in any T_n, they are labeled by $i + f_{k-1}$ and $i + f_k$, and vertices i, $i + f_{k-1}$ and $i + f_k$ first appear, respectively, in trees T_{k-1}, T_k and T_{k+1}.

Proof. If $f_{k-1} \leq i \leq f_k - 1$, then by definition of T_n, vertex i first appeared in tree T_{k-1}; and $2f_{k-1} \leq i + f_{k-1} \leq f_{k+1} - 1$, from (2) in Proposition 2.1, we know vertex $i + f_{k-1}$ should be the left child of i and it first appeared in T_k. Similarly, $f_{k+1} \leq i + f_k \leq 2f_k - 1$ and from (1) in Proposition 2.1, vertex $i + f_k$ should be the right child of i, and it first appeared in tree T_{k+1}. □

The relation between the labels of vertex i and its two children proved in the corollary will turn out to be of crucial to our later proofs and we call it the *parent-children relation*. Obviously the parent-children relation uniquely determine: the labeling of any T_n.

The *distance* between two vertices in a tree is defined to be the number of edges in the shortest path connecting them. Let the **distance sum**(distasum) of a vertex be the sum of distances from this vertex to all other vertices in the tree they are in. Denote by $d_{i,n}$ the distasum of vertex i in T_n, it will be proved in Corollary 4.4 that

$$\lim_{n\to\infty} \frac{d_{1,n}}{R_n} \sum_{ij} \frac{1}{\sqrt{d_{i,n} d_{j,n}}} = \frac{1}{2},$$

where $R_n = f_{n+1} - 1$ is the number of vertices in T_n, and the sum is taken over all edges $\{ij\}$ in T_n.

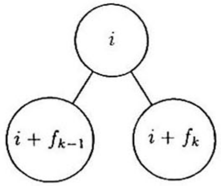

Figure 2: The parent-children relation

3 Properties of T_n

The *descendants* of a vertex in a tree are all the vertices in the subtree with the vertex as the root except the vertex itself. In a binary tree, the left descendants and right descendants of a vertex are simply the vertices in the left and right subtrees of the vertex.

LEMMA 3.1

Let $2 \le i \le n+1$ and $f_{i-1} \le j \le f_i - 1$. Then in T_n, vertex j has $f_{n-i+3} - 2$ descendants, among which $f_{n-i+2} - 1$ are left descendants, $f_{n-i+1} - 1$ are right descendants.

Proof. We prove the lemma by induction on i. First consider $i = n+1$, then the lemma is true since any vertex j such that $f_n \le j \le f_{n+1} - 1$ has no descendant by Corollary 2.2. If $i = n$, then we have, according to the proof of Proposition 2.1 Corollary 2.2, j has one left descendant and no right descendant. Suppose the lemma is true for $i = n+1, n, n-1, \dots\ k+1$, $(k < n-1)$. Consider vertex j such that $f_{k-1} \le j \le f_k - 1$, then according to Corollary 2.2, j has both left and right children. Let them be c_1 and c_2. Then according to the parent-children relation, $c_1 = j + f_{k+1}$, so $2f_{k-1} \le c_1 \le f_{k+1} - 1$, or, $f_k \le c_1 \le f_{k+1} - 1$. Thus by the induction hypothesis, c_1 has $f_{n-(k+1)+3} - 2$ descendants. Similarly, c_2 has $f_{n-k+3} - 2$ descendants. Including c_1 and c_2 themselves, j has $(f_{n-(k+1)+3} - 2) + (f_{n-k+3} - 2) + 2 = f_{n-(k-1)+3} - 2$ descendants. So the case when $i = k - 1$ is true. □

THEOREM 3.2

T_n is isomorphic to F_n for any $n \ge 1$.

Proof. As seen in the proof of the above lemma, in any T_n, the subtree determined by any vertex and its left and right subtrees (when they exist) have the sizes equal to three consecutive Fibonacci numbers each subtracting 1. This property along with the size of T_n uniquely determines the tree structure of T_n. But the Fibonacci tree F_n also has the same property by its definition, and its size is equal to that of T_n, thus the two trees have the same tree structure. □

Let us denote the set of vertices on the ith horizontal level of vertices (starting with 1 being on the first level) in T_n by $L_{i,n}$, and denote the number of elements in $L_{i,n}$ by $l_{i,n}$. For example, $L_{3,5} =$ {vertices $4, 5, 6, 8$}, and $l_{3,5} = 4$. Obviously T_n has n levels.

THEOREM 3.3

The number of vertices in the $(i+1)$th level in T_{n+1} is

$$l_{i+1,n+1} = \sum_{k=0}^{n-i} \binom{i}{k}.$$

Proof. Because of Theorem 3.2, we can easily write down the recurrence relation

$$l_{i+1,n+1} = l_{i,n} + l_{i,n-1},$$

where $l_{1,m} = 1$ for all $m \geq 1$. If we can prove $\sum_{k=0}^{n-i} \binom{i}{k}$ satisfies the same initial condition and recurrence relation then we are done. The initial condition is easily satisfied. Also we have

$$\sum_{k=0}^{(n-1)-(i-1)} \binom{i-1}{k} + \sum_{k=0}^{(n-2)-(i-1)} \binom{i-1}{k} = \sum_{k=0}^{n-i} \binom{i-1}{k} + \sum_{k=0}^{n-i-1} \binom{i-1}{k}$$
$$= \sum_{k=0}^{n-i-1} \left(\binom{i-1}{k} + \binom{i-1}{k+1} \right) + \binom{i-1}{0} = \sum_{k=0}^{n-i} \binom{i}{k+1} + \binom{i}{0} = \sum_{k=0}^{n-i} \binom{i}{k}.$$

So the recurrence relation is also satisfied. □

An immediate consequence of Theorem 3.2 is

$$d_{1,n} = \sum_{i=1}^{n} (i-1) \times l_{i,n} = \sum_{j=0}^{n-1} \sum_{k=0}^{n-j} j \binom{j-1}{k}.$$

But from the following recurrence relation we can get more information about the asymptotic value of $d_{1,n}$.

LEMMA 3.4

The recurrence relation for $d_{1,n}$ $(n \geq 3)$ is

$$d_{1,n} = d_{1,n-1} + d_{1,n-2} + f_{n+1} - 2$$

and $d_{1,n} = gnr^n + O(r^n)$, where $g = \frac{(\sqrt{5}+1)^3}{40}$ and $r = \frac{1+\sqrt{5}}{2}$ is the golden ratio.

Proof. Because of Theorem 3.2, we can see that $d_{1,n} = (d_{1,n-1} +$ number of vertices in $T_{n-1}) + (d_{1,n-2} +$ number of vertices in $T_{n-2}) = d_{1,n-1} + d_{1,n-2} + f_{n+1} - 2$.

From this, we can write down the generating function $D_1(x)$ of $d_{1,n}$:

$$D_1(x) = \frac{x^2(1+x)}{(1-x)(1-x-x^2)^2}.$$

Therefore, if we let $r_1 = \frac{-2}{1+\sqrt{5}}$, $r = \frac{-2}{1-\sqrt{5}} = \frac{1+\sqrt{5}}{2}$, there is

$$D_1(x) = \frac{a}{1-r_1x} + \frac{b}{(1-r_1x)^2} + \frac{c}{1-rx} + \frac{g}{(1-rx)^2} + \frac{2}{1-x},$$

where a, b, c, g are constants and thus

$$D_1(x) = a\sum_{k=0}^{\infty}(r_1x)^k + b\sum_{k=0}^{\infty}(k+1)(r_1x)^k + c\sum_{k=0}^{\infty}(rx)^k + g\sum_{k=0}^{\infty}(k+1)(rx)^k + 2\sum_{k=0}^{\infty}x^k,$$

Therefore, we have

$$d_{1,n} = 2 + (a + b(n+1))r_1^n + (c + g(n+1))r^n = gnr^n + O(r^n)$$

as $n \to \infty$. □

THEOREM 3.5

For $f_{i-1} \le j \le f_i - 1$, $3 \le i \le n+1$,

$$d_{j,n} = d_{j-f_{i-2},n} + R_n - 2(f_{n-i+3} - 1).$$

Proof. We know that vertex $j - f_{i-2}$ is the parent of vertex j from the definition of T_n and j is the root of a subtree of size $f_{n-i+3} - 1$ from Lemma 3.1. Now consider adding a "phantom" vertex "x" that is connected to j and $j - f_{n-2}$, each by an edge. If we let the distasum of x in T_n be $d_{x,n}$, then there is $d_{x,n} = d_{j-f_n-2,n} + f_{n-i+3} - 1$, and $d_{x,n} = d_{j,n} + R_n - (f_{n-i+3} - 1)$. Thus $d_{j,n} = d_{j-f_{i-2},n} + R_n - 2(f_{n-i+3} - 1)$. □

4 Asymptotic values

Let (i) = the label of the left-most vertex in the ith level in T_n. Then according the to parent-children relation, it is easy to see that

$$(i) = 1 + f_1 + f_2 + \ldots + f_{i-1} = f_{i+1} - 1,$$

So $f_i \le (i) \le f_{i+1} - 1$, and from Theorem 3.5, we have

$$\begin{aligned}
d_{(i),n} &= d_{(i-1),n} + R_n - 2(f_{n-(i+1)+3} - 1) \\
&= d_{(i-2),n} + R_n - 2(f_{n-i+3} - 1) + R_n - 2(f_{n-(i+1)+3} - 1) \\
&= \ldots. \\
&= d_{(1),n} + (i-1)R_n - 2(f_{n-(i+1)+3} + f_{n-i+3} + \ldots + f_{n-3+3} - (i-1)) \\
&= d_{1,n} + i \cdot R_n - 2(-f_{n+3-i} + f_{n+2} - (i-1)) + R_n.
\end{aligned}$$

Remembering that $d_{1,n} = gnr^n + O(r^n)$, $R_n = f_{n+1} - 1$, and

$$f_n = \frac{(1+\sqrt{5})^{n+1} - (1-\sqrt{5})^{n+1}}{2^{n+1}\sqrt{5}} = \frac{r^{n+1}}{\sqrt{5}} + o(r^n),$$

we have

$$d_{(i),n} = gnr^n + i\frac{r^{n+2}}{\sqrt{5}} + O(r^n).$$

Therefore,

$$\frac{d_{1,n}}{d_{(i),n}} = \frac{1}{1 + \frac{i}{n}\frac{r^2}{g\sqrt{5}}} + O(\frac{1}{n}).$$

After some routine calculation, we get

LEMMA 4.1

For large enough n and $1 \le i \le n$,

$$\sqrt{\frac{d_{1,n}^2}{d_{(i),n}d_{(i-1),n}}} = \frac{1}{1 + \frac{i}{n}\frac{r^2}{g\sqrt{5}}} + O(\frac{1}{n}).$$

THEOREM 4.2

Let $\bar{d}_{(i),n}$ be the distasum of any vertex on ith level T_n , there is

$$1 \le \frac{\bar{d}_{(i),n}}{d_{(i),n}} \le 1 + O(\frac{1}{n}).$$

Proof. From Theorem 3.5, for any $f_{i-1} \le j \le f_i - 1$, $i = 2, 3, ..., n+1$,

$$d_{j,n} - d_{j-f_{i-2},n} = R_n - 2(f_{n-i+3} - 1).$$

So the differences between the distasums of any vertex and its parent can only be chosen from the values $R_n - 2(f_2 - 1)$, $R_n - 2(f_3 - 1)$,..., $R_n - 2(f_n - 1)$; which we will call the **Differences** (of T_n). Obviously, $\bar{d}_{(i),n}$ is the sum of $d_{1,n}$ and $(i-1)$ Differences. Observe that $d_{(i),n}$ is the smallest among all $\bar{d}_{(i),n}$, since it is ($d_{1,n}$ + the smallest $(i-1)$ Differences). Therefore,

$$d_{(i),n} \le \bar{d}_{(i),n} \le d_{1,n} + \text{the largest } (i-1) \text{ Differences}$$

$$= d_{1,n} + (i-1)R_n + 2(i-1) - 2\sum_{k=2}^{i} f_k = d_{1,n} + (i-1)R_n + 2(i+2) - 2f_{i+1}.$$

Thus, for large enough n,

$$1 \le \frac{\bar{d}_{(i),n}}{d_{(i),n}} \le \frac{gnr^n + (i-1)R_n + 2(i+2) - 2f_{i+1}}{gnr^n + (i-3)R_n + 2f_{n+3-i}}$$

$$= 1 + O(\frac{1}{n})$$

□

THEOREM 4.3

Let $h_n(\frac{i}{n}) = \sum_{k=0}^{n-i}\binom{i}{k}$ and $S_n = \sum_{i=1}^{n} h_n(\frac{i}{n})$. Let $f(x)$ be any continuous function defined on $[0,1]$. Then we have

$$\lim_{n\to\infty}\sum_{i=1}^{n} f(\frac{i}{n})\frac{h_n(\frac{i}{n})}{S_n} = f(\frac{r}{\sqrt{5}}).$$

Proof. According to [10, Section 1.1], we just need to check the cases when $f(x) = x$ and x^2. Notice that $S_n = R_{n+1} = f_{n+2} - 1 = \frac{r^{n+3}}{\sqrt{5}} + o(r^n)$.

When $f(x) = x$, the numerator of the sum on the left is equal to

$$\sum_{i=0}^{n} \frac{i}{n} \sum_{k=0}^{n-i} \binom{i}{k} = \frac{1}{n}\left(\sum_{i=0}^{n}(i+1)\sum_{k=0}^{n-i}\binom{i}{k} - S_n\right).$$

It is easy to see that the first term in the parenthesis has generating function $\frac{1}{(1-x)(1-x-x^2)^2}$, after breaking down this into partial fractions as in the proof of Lemma 3.4, we see the leading term is equal to $\frac{nr^n}{(1-\frac{1}{r})(1+\frac{1}{r}^2)^2}$, multiplying by $\frac{1}{n}$ and dividing by S_n, the limit goes to $\frac{r}{\sqrt{5}}$.

When $f(x) = x^2$, the numerator of the sum on the left is equal to

$$\sum_{i=0}^{n}\left(\frac{i}{n}\right)^2 \sum_{k=0}^{n-i} \binom{i}{k} = \left(\frac{1}{n}\right)^2\left(\sum_{i=0}^{n}(i+1)(i+2)\sum_{k=0}^{n-i}\binom{i}{k} + O(nr^n)\right).$$

The generating function of the first sum in the parenthesis is $\frac{2}{(1-x)(1-x-x^2)^3}$, and it has leading term $\frac{n^2r^n}{(1-\frac{1}{r})(1+\frac{1}{r}^2)^3}$. After multiplying by $\frac{1}{n^2}$ and dividing by S_n, the limit goes to $\frac{r^2}{5}$.

□

Intuitively, h_n behaves like a delta function with pulse at $\frac{r}{\sqrt{5}}$ when $n \to \infty$.

COROLLARY 4.4

$$\lim_{n\to\infty} \frac{d_{1,n}}{R_n} \sum_{\{ij\}} \frac{1}{\sqrt{d_{i,n}d_{j,n}}} = \frac{1}{2}$$

Proof. Notice in Theorem 4.3, $h_n(\frac{i_n}{n}) = l_{i_n+1,n+1}$ and $S_n = R_{n+1}$. Let $f(x) = \frac{1}{1 + x\frac{r^2}{g\sqrt{5}}}$. According to Theorem 4.2 and Theorem 4.3 ,

$$\begin{aligned}
\lim_{n\to\infty} \frac{d_{1,n}}{R_n} \sum_{\{ij\}} \frac{1}{\sqrt{d_{i,n}d_{j,n}}} &= \lim_{n\to\infty} \sum_{i=1}^{n} \frac{l_{i,n}}{R_n} \frac{1}{1+\frac{i}{n}\frac{r^2}{g\sqrt{5}}} \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \frac{h_{n-1}(\frac{i-1}{n-1})}{S_{n-1}} \frac{1}{1+\frac{i}{n}\frac{r^2}{g\sqrt{5}}} \\
&= \lim_{m\to\infty} \sum_{j=0}^{m} \frac{h_m(\frac{j}{m})}{S_m} \left(\frac{1}{1+\frac{j+1}{m+1}\frac{r^2}{g\sqrt{5}}}\right) \\
&= \lim_{m\to\infty} \sum_{j=0}^{m} \frac{h_m(\frac{j}{m})}{S_m} \left(\frac{1}{1+\frac{j}{m}\frac{r^2}{g\sqrt{5}}} + O(\frac{1}{m})\right) = \frac{1}{1+\frac{r}{\sqrt{5}}\frac{r^2}{g\sqrt{5}}} = \frac{1}{2}.
\end{aligned}$$

□

5 Other examples

Now we study the trees of another two types of molecular graphs.

The "even-symmetric" trees (Figure 3) E_n ($n \geq 1$) are formed by connecting the roots of two copies of T_n's using an edge.

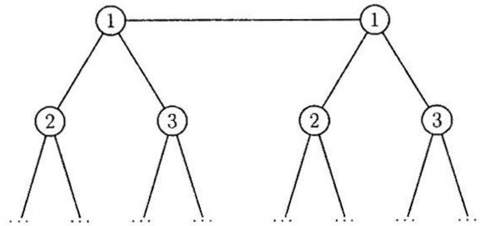

Figure 3: Even-symmetric tree

If we denote by $d'_{j,n}$ the distasum of vertex j (there are two of them now) in E_n, it is easy to see that $d'_{1,n} = 2d_{1,n} + R_n$ and $d'_{j,n} = d'_{j-f_{i-2},n} + R_n - 2(f_{n-i+3-1})$ for $f_{i-1} \leq j \leq f_i - 1$ and $3 \leq i \leq n$; this is the same recurrence relation for $d_{j,n}$. As $n \to \infty$, since R_n is of a smaller order than that of $d_{1,n}$, we see that $\frac{d'_{1,n}}{2d_{1,n}} \to 1$. Let R'_n = number of vertices in $E_n = 2R_n$. We have

$$\lim_{n\to\infty}\frac{d'_{1,n}}{R'_n}\sum_{\{ij\}\text{ in }E_n}\frac{1}{\sqrt{d'_{i,n}d'_{j,n}}} = \lim_{n\to\infty}\frac{2d_{1,n}}{2R_n}\sum_{\{ij\}\text{ in }T_n}\left(\frac{1}{\sqrt{2d_{i,n}2d_{j,n}}} + \frac{2d_{1,n}}{2R_n}\frac{1}{\sqrt{2d_{1,n}^2}}\right) = \frac{1}{2}.$$

The "odd-symmetric" trees (Figure 4) O_n ($n \geq 1$) are formed also by connecting two copies of T_n's, but this time by connecting each root to another vertex by an edge. We label this vertex by 0.

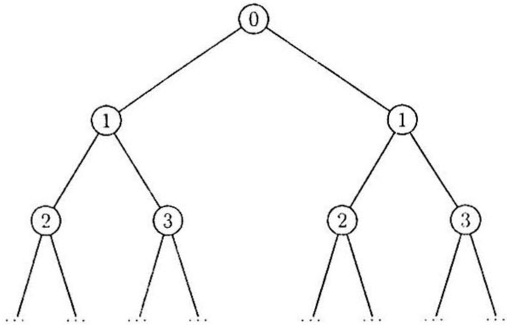

Figure 4: Odd-symmetric tree

If we denote by $d''_{j,n}$ the distasum of vertex j (there are two of them now) in O_n, then there is $d''_{0,n} = 2d_{1,n} + 2R_n$ and $\frac{d''_{0,n}}{2d_{1,n}} \to 1$; $d''_{1,n} = 2d_{1,n} + 2R_n + 1$ and $\frac{d''_{1,n}}{2d_{1,n}} \to 1$. We also have $d''_{j,n} = d''_{j-f_{i-2},n} +$

$R_n - 2(f_{n-i+3-1})$. Similarly as the even-symmetric case, let R''_n = number of vertices in $O_n = 2R_n + 1$, we have

$$\lim_{n\to\infty} \frac{d''_{1,n}}{R''_n} \sum_{\{ij\} \text{ in } O_n} \frac{1}{\sqrt{d''_{i,n}d''_{j,n}}} = \lim_{n\to\infty} \frac{2d_{1,n}}{2R_n+1} \sum_{\{ij\} \text{ in } T_n} \left(\frac{1}{\sqrt{2d_{i,n}2d_{j,n}}} + \frac{2}{\sqrt{d''_{0,n}d''_{1,n}}}\right) = \frac{1}{2}.$$

6 Conclusions

6.1 Balaban Index of T_n

Recall the Balaban Index that was defined in equation 1. For T_n, let the Balaban index be denoted by J_{T_n}. Since T_n is a tree, there is no cycle in it and the total number of edges is just $R_n - 1$. Thus we get

$$J_{T_n} = (R_n - 1)\sum_{\{ij\}} \frac{1}{\sqrt{d_{i,n}d_{j,n}}}.$$

In T_n, since as $n \to \infty$, $d_{1,n} \sim gnr^n$ and $R_n \sim \frac{r^{n+1}}{\sqrt{5}}$, Corollary 4.4 is equivalent to

$$\lim_{n\to\infty} \sum_{\{ij\}} \frac{n}{\sqrt{d_{i,n}d_{j,n}}} = \frac{\sqrt{5}}{2r^2}.$$

Therefore,

$$\lim_{n\to\infty} \frac{n}{(R_n-1)} J_{T_n} = \frac{\sqrt{5}}{2r^2},$$

or, $J_{T_n} \sim \frac{r^{n-1}}{2n}$. Similarly, $J_{E_n} \sim \frac{r^{n-1}}{2n}$ and $J_{O_n} \sim \frac{r^{n-1}}{2n}$. Thus although the Balaban indices for these molecules diverge, we are able to predict their asymptotic behavior. The asymptotic behavior of the three cases are the same because the even and odd symmetric cases preserved the delta-function like behavior of the Fibonacci tree.

6.2 The Reason for a Different Factor

In the calculation of the Balaban index for a graph, the factor $\frac{Q}{C+1}$ is usually simple, and it is the sum $\sum_{\{ij\}} \frac{1}{\sqrt{d_i d_j}}$ that may be hard to find. In our study of T_n, after determining the structure of T_n, we find $d_{i,n}$ can be found recursively from $d_{1,n}$. Furthermore, we find all $d_{i,n}$ are roughly of the same order of $d_{1,n}$ (or, up to a factor of $\frac{1}{1+\frac{i}{n}\frac{r^2}{g\sqrt{5}}}$: see Lemma 3.1). We can therefore predict that we will get a finite positive constant when we multiply the sum by the factor $\frac{d_{1,n}}{R_n}$ and eventually obtain the asymptotic behavior of the Balaban index.

The following two figures show the true value of the Balaban index for the Fibanacci molecules and the index with different scaling versus the generations. The Balaban index grows exponentially (so the logarithm of it grows linearly) and after rescaling, it goes to $\frac{1}{2}$ (see figures 5 and 6).

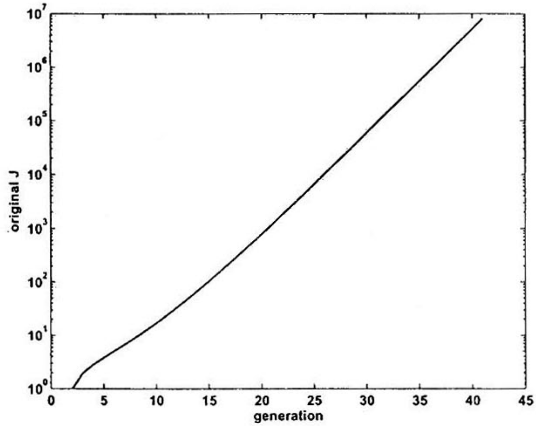

Figure 5: The logarithm of J versus generation

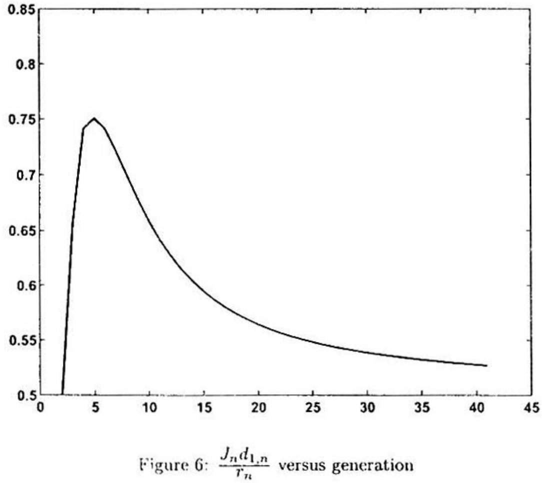

Figure 6: $\frac{J_n d_{1,n}}{r_n}$ versus generation

6.3 Dendrimer

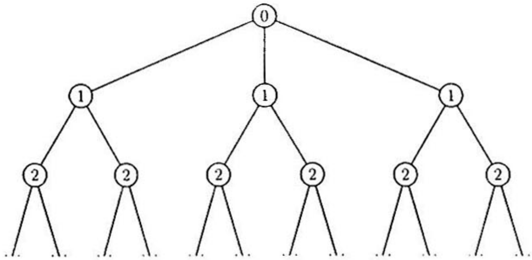

Figure 7: Dendrimer

Our method of studying the Balaban index can also be applied to other trees. For example, consider the tree of a dendrimer molecule, which is three identical complete binary trees with their roots connected to another vertex. Let us define N_n $(n \geq 1)$ be the Dendrimer tree with the root having three subtrees , each a complete binary tree with n levels of vertices (with their roots being on the first level). Label the root of N_n with 0, all its three children on the first level with 1, all the vertices on the second level with 2,..., and all the vertices on the nth level with n. It is easy to see that the size of N_n is $\hat{R}_n = 3(2^n - 1) + 1$, and the distasum of 0 is $\hat{d}_{0,n} = 3\sum_{j=1}^{n} 2^{j-1}j = 3((n-1)2^n + 1) = 3n2^n + O(2^n)$. Let $\hat{d}_{i,n}$ be the distasum of any vertex on the ith level, by an argument completely parallel to that in the proof of Theorem 3.5, we get

$$\hat{d}_{i,n} = \hat{d}_{i-1,n} + \hat{R}_n - 2(2^{n-i+1} - 1)$$

So we see $\hat{d}_{i,n}$ can also be found recursively from $\hat{d}_{1,n}$. after expanding the above recurrence relation and plugging in $\hat{d}_{1,n}$, we get

$$d_{i,n} = 2^n(3(n+i) + 2^{2-i} - 7) + 3 = 3(n+i)2^n + O(2^n).$$

So following the same argument as it in the proof of Lemma 4.1, we get

$$\sqrt{\frac{\hat{d}_{0,n}^2}{\hat{d}_{i,n}\hat{d}_{i-1,n}}} = \frac{1}{1+\frac{i}{n}} + O(\frac{1}{n}),$$

and since the number of vertices on the ith level is $3 \cdot 2^{i-1}$, we have

$$\frac{\hat{d}_{0,n}}{\hat{R}_n}\sum_{\{ij\}}\sqrt{\frac{1}{\hat{d}_{i,n}\hat{d}_{i-1,n}}} = \frac{1}{2}\sum_{i=1}^{n}\frac{2^{i-n}}{1+\frac{i}{n}} + O(\frac{1}{n}).$$

This sum is easier than the one we analyzed for T_n, since as $n \to \infty$,

$$\sum_{i=1}^{n}\frac{2^{i-n}}{1+\frac{i}{n}} \geq \sum_{i=1}^{n}\frac{2^{i-n}}{2} = 1 - 2^{-n} \to 1,$$

and if we let $k = [n - \sqrt{n}]$, then

$$\sum_{i=1}^{n} \frac{2^{i-n}}{1+\frac{i}{n}} = \sum_{i=1}^{k} \frac{2^{i-n}}{1+\frac{i}{n}} + \sum_{i=k+1}^{n} \frac{2^{i-n}}{1+\frac{i}{n}}$$
$$< \sum_{i=1}^{k} 2^{i-n} + \sum_{i=k+1}^{n} \frac{2^{i-n}}{1+\frac{k}{n}}$$
$$= 2^{k+1-n} - 2^{1-n} + \frac{2-2^{1+k-n}}{2-\frac{1}{\sqrt{n}}}$$
$$= 1 + O(\frac{1}{\sqrt{n}}) \to 1.$$

Therefore, $\sum_{i=1}^{n} \frac{2^{i-n}}{1+\frac{i}{n}} \to 1$ as $n \to \infty$. or,

$$\lim_{n\to\infty} \frac{\hat{d}_{0,n}}{\hat{R}_n} \sum_{\{ij\}} \sqrt{\frac{1}{\hat{d}_{i,n}\hat{d}_{i-1,n}}} = \frac{1}{2}.$$

Thus, the Balaban index J_{N_n} of the dendrimer molecule will be approximately $\frac{3 \cdot 2^{n-1}}{n}$ for large n.

Therefore we have $\frac{J_{T_n}}{J_{N_n}} \sim \frac{1}{6}(\frac{r}{2})^{n-1} \approx \frac{1}{6}(0.809)^{n-1}$, which illustrates how the asymptotic behavior of these two different branching patterns can be compared.

Appendix: Another Labeling Scheme for T_n

From the proof of Theorem 4.2, we know for an arbitrary vertex (i) on the ith level in T_n, $\tilde{d}_{(i),n}$ is the sum of $d_{1,n}$ and i Differences among $R_n - 2(f_{n-j+1} - 1)$ $(j = 1, 2, ..., n-1)$. We will be able to see more clearly which i Differences they should be by labeling T_n differently, or equivalently, by labeling F_n. We will call the newly labeled trees P_n, for $n = 1, 2, ...$ (Figure 8)

1. Label the vertex in F_1 by an empty string and let it be P_1. Label the root in F_2 by 0 and let its left child be labeled by 1 and let the new tree be P_2.

2. To get P_n for $n \geq 3$, label the root with a string of $n-1$ 0's, add 1 at the beginning of each label of P_{n-1} and let this newly labeled tree be the left subtree of the root of P_n; add 01 at the beginning of each label of P_{n-2} and let this newly labeled tree be the right subtree of the root P_n.

From its construction we know that P_n is isomorphic to F_n and therefore to T_n. Now we define $\vec{v} = (a_1, a_2, ..., a_{n-1})$, where $a_i = R_n - 2(f_{n-i+1} - 1)$ are the Differences.

Claim. *For each vertex j in T_n, if we regard its label in P_n as a vector and call it $\vec{u}_{j,n}$, then we have*

$$d_{j,n} = d_{1,n} + \langle \vec{v}, \vec{u}_{j,n} \rangle.$$

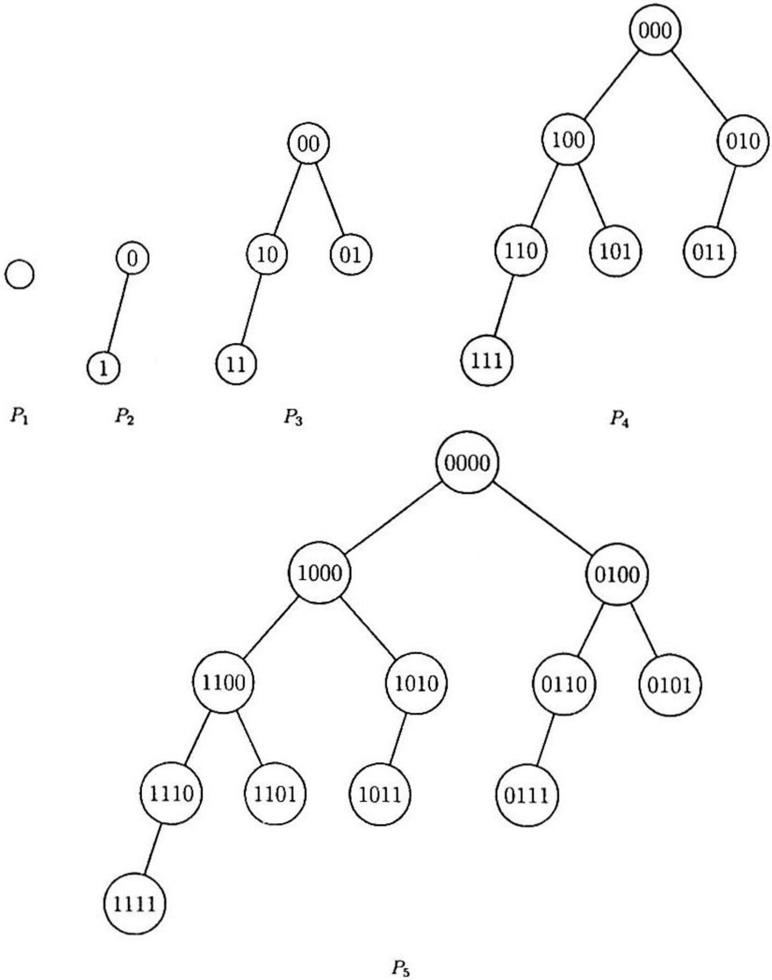

Figure 8: Examples of P_n

Or, the 1's in $\vec{u}_{j,n}$ indicate the Differences to be added to $d_{1,n}$ to get $d_{j,n}$.

Proof. Because of Corollary 2.2, any vertex, its left and right child should first appear in three consecutive T's. Also notice that for any n, a_i in $\vec{u}$ is the difference between the distasums of all the vertices added to T_i in the construction of T_{i+1} and their parents. Thus the Differences between the distasums of left child and parent and that of the right child and the parent should be adjacent in $\vec{v}$, and they should follow the Difference between the distasum of the parent and its parent. This means if our labeling in P_n has the property stated in the claim, it should have and only need to have the following properties:

For any vertex p,

1. if p is the root, then it should be labeled by a string of $n-1$ zeros, its left child should be labeled by 1 followed by $n-2$ zeros and its right child should be labeled by 01 followed by $n-3$ zeros;
2. if p has both left and right children, say, c_1 and c_2, then p should be labeled as $x...x100x...x$, where the 1 is the last 1 in the label of p. Then c_1 and c_2 should be labeled respectively as $x...x110x...x$ and $x...x101x...x$;
3. if p has only left child, then p should be labeled as $x...x10$ and its left child should be labeled as $x...x11$;
4. if p has no children, then p should be labeled as $x...x1$.

Here x's can be 0, 1, or empty string, but are the same for p and its children. Our labeling does satisfy the above property. This can be easily checked because of the recursive way P_n was labeled. Thus the claim is true. □

Notice that if we let $\vec{w} = (f_1, f_2, ...f_{n-1})$, then by parallel reasoning as the above, $j = 1 + \langle \vec{w}, \vec{u}_{j,n} \rangle$. So we have another way to get the labeling of T_n.

Acknowledgments The authors would like to thank Prof. Dennis Stanton for reading the first version of the paper, and Prof. Victor Reiner for helpful conversations.

References

[1] Z. Mihalic & N. Trinajstic, J. Chem. Ed., 69, 9, (1992) 701–712.

[2] H. Wiener, JACS 69 (1947) 17–20.

[3] (a) A. A. Dobrynin, MATCH Commun. Math. Comput. Chem., 41 (2000) 119–134.

(b) A. T. Balaban, Chem. Phys. Lett. 80 (1982) 399–404;

(c) A. T. Balaban, Pure Appl. Chem. 55 (1983) 199–206;

(d) A. T. Balaban & L. V. Quintas, MATCH Commun. Math. Chem. 14 (1983) 213–233;

(e) A. T. Balaban & P. Filip, MATCH Commun. Math. Chem. 16 (1984) 163–190;

(f) A. T. Balaban, MATCH Commun. Math. Chem. 21 (1986) 115–122;

(g) O. Ivanciuc, T. Ivanciuc & A. T. Balaban, J. Chem. Inf. Comput. Sci. 38 (1998) 395–401.

[4] N. Trinajstic, *Chemical Graph Theory*, 2nd Ed., (CRC Press, Roca Raton, 1992) 246–248.

[5] A. T. Balaban, N. Ionexu-Palla & T. S. Balaban, MATCH Commun. Math. Comput. Chem., 17(1985)121–146.

[6] P. C Painter & M. M. Coleman, *Fundamentals of Polymer Science*, 2nd Ed. (Technomic Pub., Lancaster, PA, 1997) Ch. 8.

[7] D. Bonchev, O. Mekenyan & V. Kamenska, J. Math. Chem. 11(1992)107–132.

[8] D. Bonchev, E. Markel & A. Dekmezian, J. Chem. Inf. Comp. Sci., 41(2001)1274–1285.

[9] T. S. Balaban, A. T. Balaban & D. Bonchev, J. Mol. Struct. (Theochem), 535 (2001) 81–92.

[10] I. Gutman & S. El-Basil, MATCH Commun. Math. Comput. Chem., 20(1986)81–94.

[11] R. Grimaldi, Congr. Numer. 84(1991)21–32.

[12] P. P. Korovkin, *Liner Operators and Approximation Theory*, (Hindustan Publishing Corp. (India), 1960)Ch.1

[13] L. Roger & S. Besrukov, Rostock Math. Kolloq. No. 43, (1991)17–22.

[14] T. L. Hill, *Introduction to Statistical Thermodynamics*, (Addison-Wesley, London, 1960)Ch.1