

CENTRIFUGAL-ORDER DISTRIBUTIONS IN BINARY TOPOLOGICAL TREES

- J. VAN PELT, R. W. H. VERWER and H. B. M. UYLINGS
Netherlands Institute for Brain Research,
Meibergdreef 33, 1105 AZ Amsterdam,
The Netherlands

Statistical properties of topological binary trees are studied on the basis of the distribution of segments in relation to centrifugal order. Special attention is paid to the mean of this distribution in a tree as it will be used as a measure of tree topology. It will be shown how the expectation of the mean centrifugal order depends both on the size of the tree and on the mode of growth in the context of modelling the growth of tree structures. Observed trees can be characterized by their mean orders and procedures are described to find the growth mode that optimally corresponds to these data. The variance structure of the mean-order measure appears to be a crucial factor in these fitting procedures. Examples indicate that mean-order analysis is an accurate alternative to partition analysis that is based on the partitioning of segments over sub-tree pairs at branching points.

1. Introduction. The shapes of branching patterns show, like most natural objects, a great variability and its quantification is one of the important objectives in many scientific disciplines. This paper will concern the variability in topological properties of branching patterns. By ignoring all metrical properties, the branching pattern is reduced to a number of points (branching points and terminal points) and the segments between these points. The topological structure of a tree is defined by the number of its segments (its size) and the pattern of interconnection. A finite number of different connectivity patterns, called tree types, are possible for a given size. The study of variability in tree topology essentially concerns the frequency of occurrence of tree types. The number of different tree types, however, increases rapidly with the size of trees which makes a direct comparison of tree-type distributions unfeasible. Particular properties or measures of tree topology have preferably been used rather than the tree types themselves (e.g. Jarvis and Werrity, 1975; Uylings *et al.*, 1989). For instance, one may use the partitions in a tree, i.e. the divisions of the terminal segments by the sub-tree pairs at bifurcation points in the tree. Verwer and Van Pelt (1983) and Van Pelt and Verwer (1984) have shown that in partition analysis a drastic data reduction is obtained with only a minimal loss of information. Another tree property is the distribution of segments versus centrifugal order (order distribution). The centrifugal order of a segment (or just order) denotes its topological distance to the root segment. The segments

in a tree may be concentrated near the root, may extend far from the root or may show intermediate distributions determined by the size of the tree and the tree type. For the characterization of a group of trees one may calculate an averaged distribution by averaging the number of segments at equal order over all the trees. For instance, Sadler and Berry (1983) have used the averaged order-distribution to distinguish Purkinje cells from animals of different age classes. However, the trees in such age classes are usually not of equal size and the averaged distribution is then implicitly also dependent on the size of the trees. If one wants to separate the effects of size and of the segment's connectivity on the order distribution one should average per size class. Because of the relatively small variance in the size distribution of the Purkinje cells in an age class in the experiments of Sadler and Berry (1983), the "size" effect on the shape of the segments' distribution has probably been small. A complicating factor in the comparison of (averaged) order-distributions arises from the dependency between the classes, because in a binary tree a parent segment at a particular order is connected to two daughter segments at the next order (Verwer and Van Pelt, 1986). This problem is circumvented if the mean of the distribution is used rather than the distribution profile itself. A tree type is then represented by a single number. Because of the discrete nature of tree types, the mean-order measure, like any measure of tree topology, can adopt only values from a finite, discrete value domain. This is in contrast to the continuous value domain of metrical properties like length, angles and thickness. Triller and Korn (1986) used the mean order and the maximum order in a study of the topological structure of two groups of neurons. In particular, they plotted the mean and maximum order of each observed tree against the size of the tree and they applied regression with log-functions. The regression lines for both groups of neurons appeared to correspond very closely and the authors concluded that common principles of construction may underly the topological structure of neuronal branching patterns. Unfortunately, an erroneous scaling of the abscissa makes the constants in the regression function unfit for general use.

The frequency of occurrence of tree types depends strongly on the mode of their growth. Different modes of growth will therefore produce different averaged order-distributions. Berry and Pymm (1980) have shown this by simulation studies of two hypothetical growth modes, i.e. random terminal and random segmental growth, assuming random branching exclusively at terminal segments or at all segments with equal probability, respectively. Van Pelt and Verwer (1986, 1987) were able to produce a complete range of distributions of segments by appropriate choices for the (two) parameters in a generalized growth model. The test of growth models may proceed by comparing expected order-distributions or expected mean-orders with

observed ones. For instance, Liao and Scheidegger (1968) have used the maximum, the mean and the standard deviation of the terminal-segment distribution in a study of branching patterns in hydrology. On the basis of the assumption that the topological types of 2-D trees occur with equal probability (random topology model, Shreve, 1966), they obtained the expectation values per size class for these parameters by means of Monte Carlo simulations. However, they did not discuss how to test the model outcomes with observed data.

Summarizing, the mean order (as function of the order distribution in a tree) reflects the topological structure of a tree and may be a good basis for comparative studies. Its expected value for a sample of trees depends on the mode of growth such that it may be used in testing growth models. The mean order represents a tree type by a single number in contrast to the order distribution or the partition set and this aspect may be advantageous for the analysis of trees. For these reasons the mean-order measure will be studied in this paper by its general properties and its expected behaviour for different growth models. It will be shown how to characterize a group of trees of different sizes and how to test growth models as is illustrated by the analysis of dendritic and axonal branching patterns. Especially, its use is evaluated by a comparison with outcomes of partition analysis.

2. Topological Trees and Centrifugal Order. The basic elements of a topological tree are points and segments. A binary rooted tree has terminal segments and intermediate segments ending in bifurcation points where two daughter segments arise. The degree n of a tree denotes its number of terminal segments. The tree has $n - 1$ intermediate segments and its size, denoting the total number of segments, equals $2n - 1$. In the following we shall frequently use the degree to identify a tree. The segments in a tree can be numbered by a centrifugal ordering system. The root segment has order zero and at each branching point the order of daughter segments increases with one (Fig. 1). The number of segments at each order is determined by the size of the tree and the connectivity pattern and thus by the tree topology. Tree types of equal size have different connectivity patterns but some tree types may have equal order distributions. The number of different order distributions is thus maximally equal to the number of tree types. Two extremes of order distributions exist for trees of a given size, one in which the segments extend to the highest possible orders (thin tree) and the other in which the segments are as close as possible to the root (compact trees). Note, that there is only one thin tree type (tree 1 in Fig. 1) but there may be several compact tree types (trees 5 and 6 in Fig. 1). Like the order distribution, also the mean value will show large variations for the different tree types, making it a sensitive measure for tree topology.

		DEGREE 6					
		THIN				COMPACT	
ORDER	5						
	4	Y	Y	Y	Y		
	3	Y	Y	Y	Y	Y	Y
	2	Y	Y	Y	Y	Y	Y
	1	Y	Y	Y	Y	Y	Y
	0	Y	Y	Y	Y	Y	Y
		1	2	3	4	5	6

Figure 1. Centrifugal order assignments to the segments of the six trees of degree 6 (terminal segments). The number of segments at each order are indicated.

3. Mean Centrifugal Order. Let α^n denote a tree of degree n with $s_\alpha(\gamma)$ segments at order γ and let γ_m be the highest non-empty order. The mean order $\bar{\gamma}_\alpha(n)$ of tree α^n is defined as:

$$\bar{\gamma}_\alpha(n) = \frac{1}{2n-1} \sum_{\gamma=0}^{\gamma_m} \gamma s_\alpha(\gamma) \quad \text{with} \quad \sum_{\gamma=0}^{\gamma_m} s_\alpha(\gamma) = 2n-1, \tag{1}$$

because the tree has a total of $2n-1$ segments.

3.1 Thin trees. For a thin tree of degree n we have $s(0)=1$ and $s(\gamma)=2$ for $\gamma=1, \dots, \gamma_m$. The highest possible order in a thin tree is at $\gamma_m=n-1$. The mean order of a thin tree is thus equal to:

$$\bar{\gamma}_{\text{thin}}(n) = \frac{1}{2n-1} \sum_{\gamma=1}^{n-1} 2\gamma = \frac{n(n-1)}{2n-1}. \tag{2}$$

For large n the mean order in thin trees approaches $n/2$.

3.2 Compact trees. In compact trees the number of segments double at each order. Only the highest order may not be completely filled. Up to the highest order we have 2^γ segments in order γ and in total

$$\sum_{\gamma=0}^{\gamma_m-1} 2^\gamma = 2^{\gamma_m} - 1$$

segments. The number of segments in the highest order then equals $2n-1-(2^{\gamma_m}-1)=2n-2^{\gamma_m}$. The highest order can be obtained by considering that it certainly contains segments ($2n-2^{\gamma_m}>0$) but may not be completely filled ($2n-2^{\gamma_m}\leq 2^{\gamma_m}$). These inequalities can be rewritten as $\gamma_m < {}^2\log(n)+1$ and $\gamma_m \geq {}^2\log(n)$, respectively. Considering that γ_m must be integer, it follows that γ_m equals the smallest integer greater than or equal to ${}^2\log(n)$, denoted by $\lceil {}^2\log(n) \rceil$. This property can also be formulated as $\gamma_m = \lfloor ({}^2\log(n-1)) \rfloor + 1$, if $\lfloor x \rfloor$ denotes the greatest integer smaller than or equal to x . An analogous result was obtained by Werner and Smart (1973) for the minimum possible diameter

in channel networks. They defined the diameter as the largest path length (number of segments from the root to a terminal point). For the mean order in compact trees we can write:

$$\bar{\gamma}_{\text{compact}}(n) = \frac{1}{2n-1} \left\{ \sum_{\gamma=0}^{\gamma_m-1} \gamma 2^\gamma + \gamma_m(2n-2^{\gamma_m}) \right\}. \quad (3)$$

The summation can be evaluated as follows:

$$\begin{aligned} \sum_{\gamma=0}^{\gamma_m} \gamma 2^\gamma &= \sum_{i=1}^{\gamma_m} \sum_{\gamma=i}^{\gamma_m} 2^\gamma = \sum_{i=1}^{\gamma_m} \left\{ \sum_{\gamma=0}^{\gamma_m} 2^\gamma - \sum_{\gamma=0}^{i-1} 2^\gamma \right\}, \\ &= \sum_{i=1}^{\gamma_m} \{2^{\gamma_m+1} - 1 - 2^i + 1\} = \gamma_m 2^{\gamma_m+1} - \sum_{i=1}^{\gamma_m} 2^i = \gamma_m 2^{\gamma_m+1} - 2^{\gamma_m+1} + 2, \end{aligned}$$

and finally we have:

$$\sum_{\gamma=0}^{\gamma_m} \gamma 2^\gamma = 2[1 + 2^{\gamma_m}(\gamma_m - 1)]. \quad (4)$$

Substituting this result into equation (3) gives:

$$\begin{aligned} \bar{\gamma}_{\text{compact}}(n) &= \frac{1}{2n-1} \{2[1 + 2^{\gamma_m-1}(\gamma_m - 2)] + \gamma_m(2n - 2^{\gamma_m})\} \\ &= \frac{2}{2n-1} (1 - 2^{\gamma_m} + n\gamma_m) \quad \text{while} \quad \gamma_m = \lceil 2 \log(n) \rceil. \end{aligned} \quad (5)$$

If also the highest order is completely filled we have $\gamma_m = 2 \log(n)$ and the mean order equals:

$$\bar{\gamma}_{\text{compact}}(n) = \frac{2}{2n-1} [1 - n + n \cdot 2 \log(n)] = \frac{2}{2n-1} [n(2 \log(n) - 1) + 1], \quad (6)$$

and for large n we have:

$$\lim_{n \rightarrow \infty} \bar{\gamma}_{\text{compact}}(n) = 2 \log(n) - 1. \quad (7)$$

3.3 Mean-order expectation values. Equations (2) and (5) show the different relationships between mean order and degree for the extreme tree types. Any other tree type will have a mean-order value between those of the thin and compact tree. For the averaged mean-order for a set of N trees of equal degree n with individual mean-order values $\bar{\gamma}_i$, $i = 1, \dots, N$ we have:

$$\bar{\gamma}(n) = \frac{1}{N} \sum_{i=1}^N \bar{\gamma}_i(n). \tag{8}$$

Substitution of equation (1) gives:

$$\bar{\gamma}(n) = \frac{1}{N} \sum_{i=1}^N \frac{1}{2n-1} \sum_{\gamma=0}^{\gamma_m} \gamma s_i(\gamma).$$

Exchanging both summations results in:

$$\bar{\gamma}(n) = \frac{1}{2n-1} \sum_{\gamma=0}^{\gamma_m} \gamma \frac{1}{N} \sum_{i=1}^N s_i(\gamma) = \frac{1}{2n-1} \sum_{\gamma=0}^{\gamma_m} \gamma \bar{s}(\gamma), \tag{9}$$

if $\bar{s}(\gamma)$ denotes the mean number of segments at order γ for a set of trees. Instead of calculating $\bar{\gamma}(n)$ in terms of actual occurring trees, it can also be expressed in terms of probabilities of occurrence of trees. Let $A^n = \{\alpha_i^n | i = 1, \dots, N_\alpha^n\}$ be the set of the N_α^n different tree types of degree n and let $p(\alpha_i^n)$ be the probability of occurrence of tree α_i^n within this set. The expectation of the mean order in this set is then equal to:

$$E(\bar{\gamma}|n) = \sum_{i=1}^{N_\alpha^n} p(\alpha_i^n) \bar{\gamma}_{\alpha_i^n}(n). \tag{10}$$

Inserting equation (1) gives:

$$\begin{aligned} E(\bar{\gamma}|n) &= \sum_{i=1}^{N_\alpha^n} p(\alpha_i^n) \frac{1}{2n-1} \sum_{\gamma=0}^{n-1} \gamma s_{\alpha_i^n}(\gamma) = \frac{1}{2n-1} \sum_{\gamma=0}^{n-1} \gamma \sum_{i=1}^{N_\alpha^n} p(\alpha_i^n) s_{\alpha_i^n}(\gamma) \\ &= \frac{1}{2n-1} \sum_{\gamma=0}^{n-1} \gamma E(s(\gamma)|n). \end{aligned} \tag{11}$$

Here,

$$E(s(\gamma)|n) = \sum_{i=1}^{N_\alpha^n} p(\alpha_i^n) s_{\alpha_i^n}(\gamma)$$

and denotes the expected number of segments at order γ in a tree of degree n . A tree of degree n has $2n - 1$ segments thus

$$\sum_{\gamma=0}^{n-1} E(s(\gamma)|n) = 2n - 1.$$

For the maximal order $n - 1$ only the thin tree in set A^n contributes with two

terminal segments to the expected value. In analogy to $E(s(\gamma)|n)$ we may also introduce the expected number of intermediate segments $E(s_i(\gamma)|n)$ and of terminal segments $E(s_t(\gamma)|n)$ at order γ . We have:

$$\begin{aligned} E(s(\gamma)|n) &= E(s_i(\gamma)|n) + E(s_t(\gamma)|n) \quad \text{and} \\ E(s(\gamma)|n) &= 2E(s_i(\gamma-1)|n), \end{aligned} \quad (12)$$

because every intermediate segment at order $\gamma-1$ is connected to two segments at order γ . A direct relation between $E(s(\gamma)|n)$ and $E(s_t(\gamma)|n)$ can easily be derived by repeatedly applying the equations in (12) resulting in:

$$E(s(\gamma)|n) = 2^\gamma \left[1 - \sum_{k=1}^{\gamma-1} 2^{-k} E(s_t(k)|n) \right]. \quad (13)$$

A tree of degree n has $n-1$ intermediate and n terminal segments thus:

$$\sum_{\gamma=0}^{n-1} E(s_i(\gamma)|n) = n-1 \quad \text{and} \quad \sum_{\gamma=0}^{n-1} E(s_t(\gamma)|n) = n.$$

Additionally, we may define the expectation of the mean order of intermediate segments $E(\bar{\gamma}_i|n)$ and of terminal segments $E(\bar{\gamma}_t|n)$ as:

$$E(\bar{\gamma}_i|n) = \frac{1}{n-1} \sum_{\gamma=0}^{n-1} \gamma E(s_i(\gamma)|n) \quad \text{and} \quad E(\bar{\gamma}_t|n) = \frac{1}{n} \sum_{\gamma=0}^{n-1} \gamma E(s_t(\gamma)|n). \quad (14)$$

The expectations $E(\bar{\gamma}|n)$, $E(\bar{\gamma}_i|n)$ and $E(\bar{\gamma}_t|n)$ are related by:

$$E(\bar{\gamma}|n) = \frac{n-1}{2n-1} E(\bar{\gamma}_i|n) + \frac{n}{2n-1} E(\bar{\gamma}_t|n), \quad (15)$$

as can be shown by substituting equations (12) and (14) into equation (11). A direct relation exists also between $E(\bar{\gamma}|n)$ and $E(\bar{\gamma}_i|n)$ and between $E(\bar{\gamma}|n)$ and $E(\bar{\gamma}_t|n)$. For instance, using equations (11), (12) and (14), we obtain:

$$\begin{aligned} E(\bar{\gamma}_i|n) &= \frac{1}{n} \sum_{\gamma=0}^{n-1} \gamma E(s(\gamma)|n) - \frac{1}{n} \sum_{\gamma=0}^{n-1} \gamma \frac{E(s(\gamma+1)|n)}{2} \\ &= \frac{2n-1}{n} E(\bar{\gamma}|n) - \frac{1}{2n} \sum_{\gamma=0}^{n-1} (\gamma+1) E(s(\gamma+1)|n) + \frac{1}{2n} \sum_{\gamma=0}^{n-1} E(s(\gamma+1)|n) \\ &= \frac{2n-1}{n} E(\bar{\gamma}|n) - \frac{2n-1}{2n} E(\bar{\gamma}|n) + \frac{2n-2}{2n} \\ &= \frac{2n-1}{2n} [E(\bar{\gamma}|n) + 1] - \frac{1}{2n}. \end{aligned} \quad (16)$$

We can also write:

$$E(\bar{\gamma}|n) = \frac{2n(E(\bar{\gamma}_i|n) - 1) + 2}{2n - 1}. \quad (17)$$

In a corresponding way, it can be shown that:

$$E(\bar{\gamma}_i|n) = \frac{2n - 1}{2(n - 1)} E(\bar{\gamma}|n) - 1 \quad \text{or} \quad E(\bar{\gamma}|n) = \frac{2(n - 1)}{2n - 1} (E(\bar{\gamma}_i|n) + 1). \quad (18)$$

4. Order Distributions and Growth Models. The quantification of the variability in the topological structure of trees essentially implies the determination of the frequencies of occurrence of topological tree types. Theoretical studies have shown that the probabilities of occurrence of trees strongly depend on the mode of growth (e.g. Harding (1971), Hollingworth and Berry (1975), Dacey and Krumbein (1976) and Van Pelt and Verwer (1986) for sequential growth models and Horsfield *et al.* (1987) for synchronous growth models). These models offer also the possibility to test the model predictions against the frequencies of observed tree-types. The use of appropriate measures in these tests is essential and we will therefore study the order distribution and the mean order also within the context of these growth models. In particular, we shall use the generalized sequential parameter model of Van Pelt and Verwer (1986) because it encompasses by its parameterization the particular sequential models of the mentioned authors. This so-called Q - S growth model defines the branching rules of the segments in a tree as: (1) a branching event is defined as the division of a segment into two segments by a branching point from which a new terminal segment protrudes; (2) growth is considered as a series of branching events, one at a time; (3) a branching event may occur at every segment in a tree with a certain probability; (4) the branching probability of a terminal segment has been defined as $p_t = C \cdot 2^{-S\gamma}$ and of an intermediate segment as $p_i = R \cdot p_t$ (Van Pelt and Verwer, 1986). The parameter γ denotes the segment order, S defines the strength of the order dependency and the parameter R defines the ratio of branching probabilities of an intermediate and a terminal segment of equal order. The parameter $Q = R/(R + 1) = p_i/(p_i + p_t)$ with value domain $[0, 1]$ is preferably used rather than R which may have values between 0 and ∞ . The normalization constant C must be chosen such that the branching probabilities of all segments sum to one to ensure that only one segment branches at a time. The probability of any tree type can be calculated on basis of these growth rules (Q - S model). If the order-dependence is removed (i.e. $S = 0$) the model reduces to the so-called Q -model. If branching is only allowed for terminal segments (i.e. $R = 0$ or $Q = 0$) the model reduces to the so-called S -model. For instance, for $Q = 0.5$ ($R = 1$) and

$S=0$ we have a mode of growth with equal branching probabilities for all segments, corresponding to the random segmental growth model. For $(Q, S)=(0, 0)$ we have branching of exclusively terminal segments with equal probabilities (random terminal growth). For $(Q, S)=(0, 1)$ the branching probability of a terminal segment at a particular order is twice that of a terminal segment at the next higher order.

The probability of a particular tree can also be expressed in terms of the probabilities of all its partitions (Van Pelt and Verwer, 1985). It is therefore sufficient to calculate the partition probabilities instead of a much larger number of tree-type probabilities as outcome of the growth model. Also the mean-order expectation value should preferably be expressed in terms of partition probabilities rather than following equation (10) that is based on tree-type probabilities. A recursion for $E(s(\gamma)|n)$ can immediately be formulated. A tree of degree n has a probability $p(r, n-r|Q, S)$ to have first-order subtrees of degree r and $n-r$ with a value determined by the Q - S growth model. The expected number of segments at order γ in the n th-degree tree equals the sum of the expected number of segments at order $\gamma-1$ in its two first-order subtrees and thus:

$$E(s(\gamma)|n) = \sum_{r=1}^{[n/2]} p(r, n-r|Q, S) [E(s(\gamma-1)|r) + E(s(\gamma-1)|n-r)], \quad (19)$$

for $n > 1$ and $\gamma > 0$. With $E(s(0)|n) = 1$ for $n > 0$, the recursion in equation (19) and the partition probabilities $p(r, n-r|Q, S)$, the expectation values $E(s(\gamma)|n)$ can be calculated for any Q - S model. Examples of the order distributions are given in Fig. 2 for $(Q, S)=(0.8, 0)$, $(0.5, 0)$, $(0, 0)$ and $(0, 1)$ and for several values of the degree n . The expected number of segments per order may vary considerably by appropriate choices for the mode of growth. A small number of segments for the whole range of orders is expected for $(Q, S)=(0.8, 0)$ (Fig. 2a). The growth mode $(Q, S)=(0, 1)$, on the contrary, predicts trees with a concentration of segments at low orders which are (almost) completely filled (Fig. 2d). The mean of each order distribution also relates to the degree n in a growth-mode dependent way as is shown in Fig. 3. The curves in Fig. 3 are obtained by calculating for each degree the expected mean-order via equations (11) and (19). The curves in Fig. 3a are calculated for several values of the parameter Q with S taken zero (Q -model). The curves for compact and thin trees are included to indicate the total range of mean-order values and are calculated by means of equations (2) and (5). The set of curves shows that the Q -model is able to produce tree sets with expected mean-orders in a large range between compact and thin trees, bounded by the values for the "thin trees" ($Q=1$) and by the values predicted for terminal growth ($Q=0$). The Q -model is not able to predict sets of trees that have on the average more segments at low

order. The gap between the “terminal growth” curve and the “compact tree” curve can be filled by assuming positive values for the parameter S (Fig. 3b). Apparently, the sequential growth model only predicts sets with (nearly) compact trees if the branching probabilities are assumed to be order dependent with dominant branching of low-order segments. Although the chosen parameter values represent only a part of the Q – S parameter space (i.e. the Q -axis and the positive S -axis), they are able to produce a full range of order distributions and mean-order curves. The results of other parameter values appear to closely resemble one out of this range and for ease of description we shall therefore consider only this restricted area in the model parameter space.

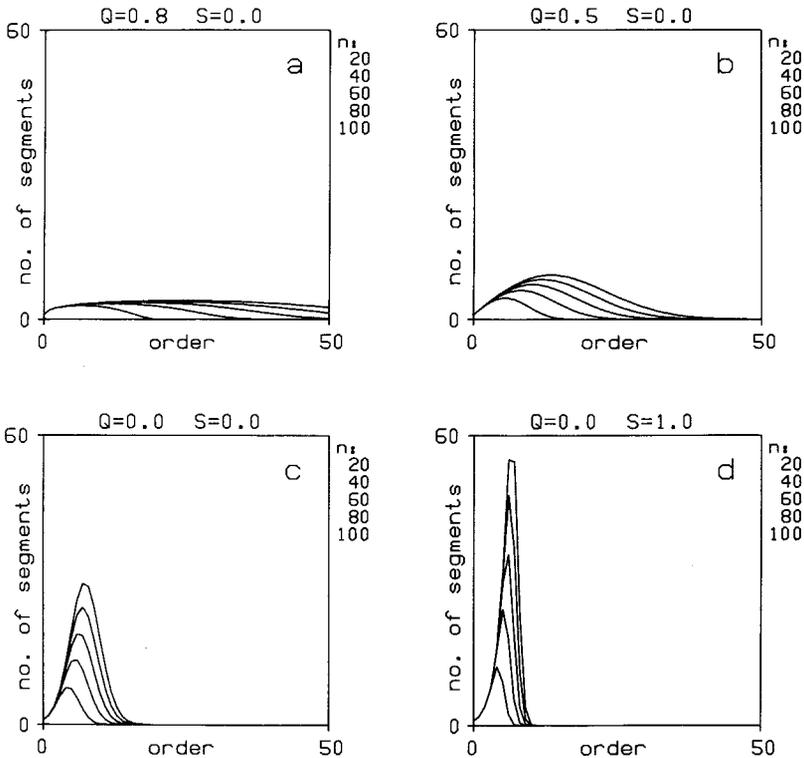


Figure 2. Expectation values for the number of segments *vs* centrifugal order, calculated for four different values of the growth-model parameters Q and S . Each sub-figure contains five successive order-distributions calculated for trees of degree $n = 20, 40, 60, 80$ and 100 , respectively. (a) A rather flat distribution, extending up to high orders, is produced by growth mode $(Q, S) = (0.8, 0)$ indicating that elongated trees with few segments per order have a high probability of occurrence. (d) The power of two shaped rising part of the curves indicates that most orders are completely filled. Thus, compact trees have high probability of occurrence for the growth mode $(Q, S) = (0, 1)$. (b) The order distributions for random segmental growth and (c) for random terminal growth show transitional shapes.

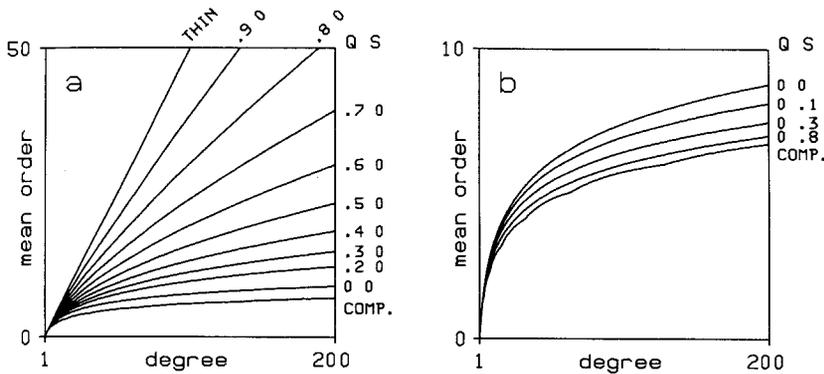


Figure 3. Expectation values of the mean centrifugal order in trees as function of the degree. The curves are calculated for several values of the growth-model parameters Q and S with (a) parameter S taken zero and (b) parameter Q taken zero. The curves for both thin and compact (comp.) individual trees are also plotted to indicate the extreme cases. The non-smoothed compact-tree curve shows the “filling effect” of the highest order. Note that the set of curves in (b) are positioned between the lowest two curves in (a).

5. Numerical Calculations. Unfortunately, we did not succeed in formulating the order distributions in general closed expressions. For specific growth models, however, more direct formulae could be obtained.

5.1 Random segmental growth. A closed expression has only been found for the random segmental growth model ($Q = 0.5; S = 0$). This model results in the same tree-type probabilities as the random topology model (Shreve, 1966). For this model, Werner and Smart (1973) have shown that:

$$E(s(\gamma)|n) = \frac{2^\gamma}{N(n)} \frac{\gamma + 1}{2n - \gamma - 1} \binom{2n - \gamma - 1}{n} \quad \text{with} \quad N(n) = \frac{1}{2n - 1} \binom{2n - 1}{n}. \quad (20)$$

This equation describes the curves in Fig. 2b for $(Q, S) = (0.5, 0)$. Werner and Smart (1973) also present a formula for the expected value of the total path length of terminal segments to be $2^{2^{n-1}}/N(n)$. Realizing that the path length of a segment includes the segment itself and thus is equal to the centrifugal order plus one, the total path length of terminal segments can also be expressed as $n \cdot (E(\bar{\gamma}_i|n) + 1)$ and using equation (17) it follows that:

$$E(\bar{\gamma}|n) = \frac{2^{2^{n-1}}}{(2n - 1) \cdot N(n)} - 2. \quad (21)$$

5.2 Random terminal growth. For random terminal growth ($Q = 0; S = 0$) it was possible to derive a simple recurrent relation. Random terminal growth

implies branching of exclusively terminal segments, each one with equal probability. Any segment in a tree then maintains its order during growth of the tree. The number of segments at order γ only increases (with two segments) if a terminal segment at order $\gamma - 1$ branches. A tree of degree $n - 1$ has an expected number of $E(s_t(\gamma - 1)|n - 1)$ terminal segments at order $\gamma - 1$ and each terminal segment has a probability of $1/(n - 1)$ being selected as the branching segment. Thus:

$$E(s(\gamma)|n) = E(s(\gamma)|n - 1) + 2E(s_t(\gamma - 1)|n - 1) \frac{1}{n - 1}. \tag{22}$$

Using equation (12) we get:

$$E(s(\gamma)|n) = E(s(\gamma)|n - 1) + \frac{2}{n - 1} \left\{ E(s(\gamma - 1)|n - 1) - \frac{E(s(\gamma)|n - 1)}{2} \right\},$$

and finally:

$$E(s(\gamma)|n) = \frac{n - 2}{n - 1} E(s(\gamma)|n - 1) + \frac{2}{n - 1} E(s(\gamma - 1)|n - 1), \tag{23}$$

with $\gamma > 0$ and $E(s(0)|n) = 1$.

A concise expression could be derived by working out the scheme of $E(s(\gamma)|n)$ values that is obtained by repeatedly applying the recursion of equation (23), resulting in:

$$E(s(\gamma)|n) = \frac{2^\gamma}{n - 1} \sum_{k=1}^{n-\gamma} {}^k I_{\gamma-1}^{n-2} \quad \text{with } \gamma > 0 \text{ and } n > 1. \tag{24}$$

The symbol ${}^m I_k^n$ denotes the sum of inverse products of all combinations of k unique integers in the integer interval $[m, n]$. For instance:

$${}^2 I_3^5 = \frac{1}{2 \cdot 3 \cdot 4} + \frac{1}{2 \cdot 3 \cdot 5} + \frac{1}{2 \cdot 4 \cdot 5} + \frac{1}{3 \cdot 4 \cdot 5}.$$

Especially, we have:

$${}^m I_1^n = \frac{1}{m} + \frac{1}{m + 1} + \dots + \frac{1}{n}, \tag{25}$$

and:

$${}^m I_{n-m+1}^n = \frac{1}{m \cdot (m + 1) \cdot (m + 2) \cdot \dots \cdot n} = \frac{(m - 1)!}{n!}. \tag{26}$$

An alternative recursive relation for the calculation of the order distribution of

terminal segments in the case of random terminal growth has been described by Ley *et al.* (1986):

$$E(s_t(\gamma + 1)|n) = \frac{2^\gamma}{n} P_{(\gamma+1,n)} \quad \text{while} \quad P_{(\gamma+1,n)} = \frac{1}{n-1} \sum_{k=\gamma}^{n-1} P_{(\gamma,k)}. \quad (27)$$

The order distribution of segments can subsequently be calculated using equation (13). An expression for the expectation of the mean order in trees, grown by random terminal branching could be obtained from the work of Vannimenus *et al.* (1984). They studied the diffusion controlled process of cluster growth on the Cayley tree and showed that it is equivalent to the Eden model where growth occurs at any boundary site with equal probability. In this way it corresponds to our model of a growing tree via branching of randomly chosen terminal segments. Vannimenus *et al.* (1984) were interested in the size of the cluster, the number of particles at a certain level and the mean level of the particle distribution. To this end, the authors also had to study a recurrent relation, analogous to the one in equation (23). Their results can easily be adapted to growing trees by realizing that there is a one-to-one correspondence between particles in the cluster and branching points (and thus intermediate segments) in the tree. For the mean order of intermediate segments we then have:

$$E(\bar{\gamma}_i|n) = 2 \frac{n}{n-1} [\psi(n) + \gamma_{\text{Euler}}] - 2, \quad (28)$$

and for large n we have:

$$E(\bar{\gamma}_i|n) \simeq 2[\ln(n) + \gamma_{\text{Euler}}] - 2,$$

(Vannimenus *et al.*, 1984) where $\psi(n)$ is the digamma function and γ_{Euler} is Euler's constant ($\gamma_{\text{Euler}} = 0.5772 \dots$).

By means of equations (18) and (28) we get for $E(\bar{\gamma}|n)$ for terminally grown trees:

$$E(\bar{\gamma}|n) = \frac{2}{2n-1} \{2n[\psi(n) + \gamma_{\text{Euler}}] - 3(n-1)\} = \frac{2}{2n-1} \left\{ 2n \sum_{k=1}^{n-1} \frac{1}{k} - 3(n-1) \right\}, \quad (29)$$

(Abramowitz and Stegun, 1965) and for large n we get:

$$E(\bar{\gamma}|n) \simeq \frac{2}{2n-1} \{2n[\ln(n) + \gamma_{\text{Euler}}] - 3(n-1)\} \quad (30)$$

(Abramowitz and Stegun, 1965).

6. *Approximate Analytical Functions.* The mean-order *vs* degree curves in Fig. 3 are obtained by calculating the expected order-distribution and mean order for any degree. For practical applications, however, it is highly desirable to have a closed expression for the functional relationship between mean order and degree including the growth-model parameters Q and S . However, such an expression has only been obtained for random segmental growth. Therefore, we have studied analytical functions that approximate the calculated curves closely. We have found that the function:

$$E(\bar{\gamma}|n, Q, S) = \frac{1}{2n-1} \left[A(Q, S) \cdot n \cdot \frac{n^Q - 1}{Q} - B(Q, S) \cdot (n-1) \right], \quad (31)$$

describes the mean order as function of the degree n for a given growth mode (Q, S) very satisfactory by appropriate choices for the coefficients A and B . These coefficients thus remain functions of the parameters Q and S . For terminal growth and large n , equation (31) is consistent with equation (30) because:

$$\lim_{Q \rightarrow 0} \frac{n^Q - 1}{Q} = \ln(n).$$

It is shown in Fig. 4 how precisely the function of equation (31) can describe mean-order curves for both the Q -model (Fig. 4a) and the S -model (Fig. 4b). The optimized values for the coefficients A and B for each curve are given in Table I. Unfortunately, we did not find an analytical expression for the coefficients A and B as functions of Q and S . The results must therefore be regarded as a start to obtain an analytical expression for $E(\bar{\gamma}|n, Q, S)$ making it of practical value.

7. *Variances.* The mean-order curves in Fig. 3 represent for any degree n the average of mean-order values of all individual trees in the set A^n , weighted for their probability of occurrence [equation (10)].

The mean order of individual trees may deviate considerably from these averaged values because in our probabilistic scheme any tree type has a non-zero probability of occurrence. To get an impression of the deviations of mean-order values of individual trees from the calculated mean-order curves we have created a sample of trees by means of Monte Carlo (MC) simulations, one tree per degree, and plotted the mean-order values of all the trees *vs* their degree (Fig. 5). The spread of the simulated mean-order values appears to depend both on the growth mode and on the degree. It becomes very small if the curve approaches the extreme lines for compact and thin trees, respectively. The shape of the distribution of individual mean-order values can be studied by creating a large sample of trees of equal degree via MC simulation. Figure 6

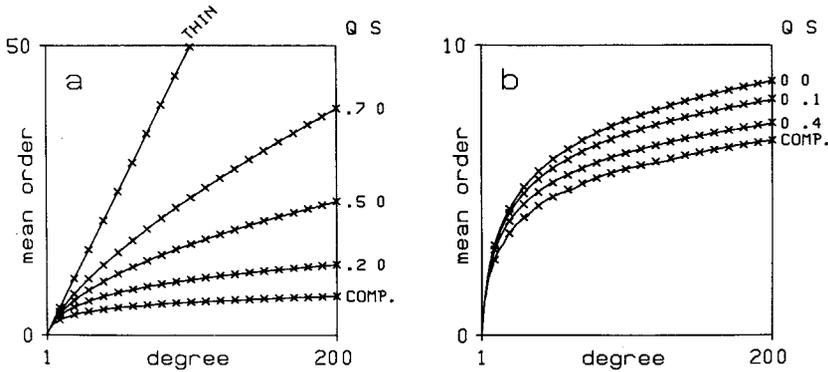


Figure 4. Plots of several mean-order *vs* degree curves for (a) the *Q*-model and for (b) the *S*-model. The crosses indicate the outcomes of the analytical equation (31) if the coefficients *A* and *B* are optimally chosen. These values are given in Table I. The plots show that the model curves are very accurately described by the approximate analytical equation.

TABLE I
 Values for the Coefficients *A* and *B* that Optimize the Fit of the Analytical Function of Equation (31) to the Mean-order *vs* Degree Model Curves, Displayed in Fig. 4

(<i>Q</i> , <i>S</i>)	Thin	(0.7, 0)	(0.5, 0)	(0.2, 0)	(0, 0)	(0, 0.1)	(0, 0.4)	Compact
<i>A</i>	1	1.38	1.78	2.80	3.99	3.55	3.05	2.89
<i>B</i>	0	0.22	0.74	2.13	3.64	2.53	1.56	1.91

shows the frequency distribution of individual mean-order values for samples of 10,000 trees of degree 10, 25, 50 and 100, obtained for growth modes (*Q*, *S*) = (0, 0), (0.5, 0) (0.8, 0) and (0.99, 0). Values for mean, standard deviation (SD) and coefficient of variation are listed in Table II. The coefficient of variation depends on the mode of growth, as expected, and also on the degree, although slightly weaker than the standard deviation does.

8. *Analysis of Observed Trees.* Observed trees can be represented by their mean order and degree and a plot results in a scattergram. The variability in tree types can subsequently be expressed by the variances in both the degree and the mean-order values. But, as has been shown, the averaged mean-order per degree depends on the degree in a way that is characteristic for the set of trees. Appropriate regression then allows for a separation between this characteristic relation and the residual variance in mean-order values. The most educated approach is by using the model curves for the regression because the model predicts the relation between degree and expected mean-order while

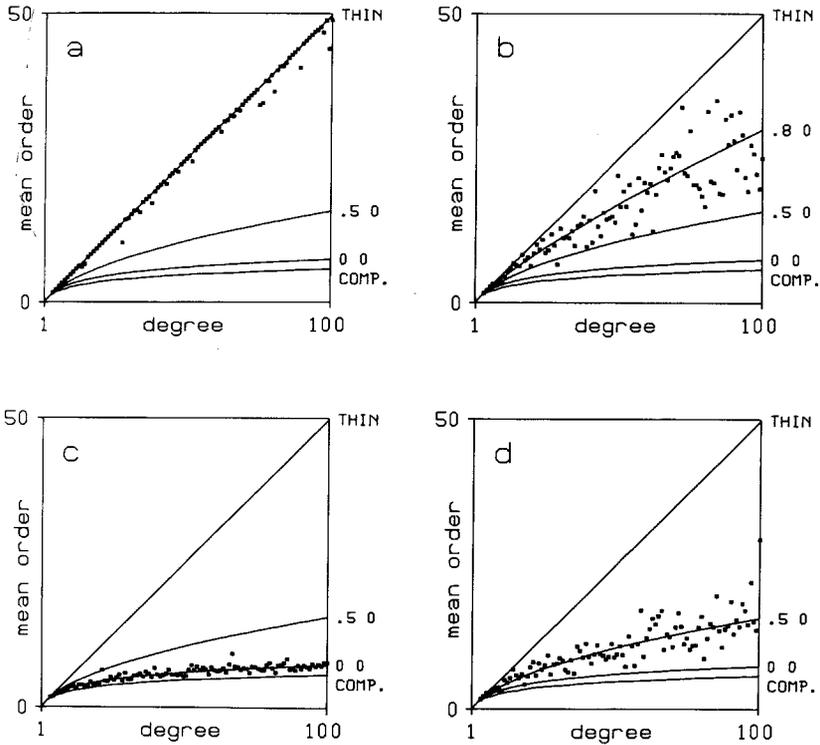


Figure 5. Mean order of individual trees, plotted *vs* their degree. The trees, one per degree, are produced by means of Monte Carlo simulations on basis of growth mode $S=0$ and: (a) $Q=0.99$; (b) $Q=0.8$; (c) $Q=0$; (d) $Q=0.5$. Note that the variance in mean-order values depends strongly on the mode of growth and increases with the degree in a growth-mode dependent way. Some model curves for the mean-order expectation values are drawn for purpose of reference.

the residual variance only arises from the random choices for the branching segments during growth.

The best-fitting model curve will be found by a least-squares procedure, i.e. by minimizing the chi-square distributed quantity:

$$T = \sum_{i=1}^N \{(\gamma_i - E(\bar{\gamma}|n_i))/SD(n_i)\}^2.$$

The chi-square value T is the sum over all N observations (trees) of the squared, weighted differences between the mean order γ_i of tree i (with degree n_i) and the expected mean-order $E(\bar{\gamma}|n_i)$ for trees of degree n_i . The inverse expected mean-order standard deviation $1/SD(n_i)$ is used as weight. The goodness of fit can be calculated from the minimal chi-square value T_{\min} , the number of data points (trees) N and the number of free parameters (par) in the model curve.

TABLE II
 Distributional Properties of Mean Order of Trees Obtained by Monte Carlo Simulation (Fig. 6)

(Q, S)	(0, 1)		(0, 0)		(0.5, 0)		(0.8, 0)		(0.99, 0)	
Degree	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
10	2.77	0.14	3.11	0.35	3.68	0.54	4.22	0.52	4.71	0.13
25	4.02	0.10	4.77	0.49	6.91	1.30	9.44	1.59	12.08	0.55
50	4.98	0.07	6.07	0.56	10.59	2.15	16.82	3.15	24.25	1.21
100	5.96	0.05	7.42	0.59	15.79	3.46	29.94	5.91	48.42	2.52

(Q, S)	(0, 1)	(0, 0)	(0.5, 0)	(0.8, 0)	(0.99, 0)
Degree	Coefficient of Variation				
10	0.05	0.11	0.15	0.12	0.03
25	0.02	0.10	0.19	0.17	0.05
50	0.02	0.09	0.20	0.19	0.05
100	0.01	0.08	0.22	0.20	0.05

Most tables in the literature use the reduced chi-square value as entry that is defined as T_{\min}/df . The number of degrees of freedom (df) is equal to $df = N - \text{par}$. A programme for calculating the goodness-of-fit can be found in Bevington (1969).

In finding the best fitting model curve a problem arises because the model curves are not available in analytical form. Each curve must be calculated again for each value of the parameters Q and S . We have therefore calculated the chi-square values for only a limited number of curves ($Q = 0, 0.1, \dots, 1.0$, and $S = 0, 0.1, \dots, 1.0$). A more precise estimate of the second decimal in the model parameter is obtained by interpolating the chi-square values by means of a cubic-spline function (Press *et al.*, 1986). The minimum of this function gives the parameter value for the best-fitting model.

Like the model curve, also the expected standard deviations are unknown functions of both the model parameters Q and S and the degree n (Fig. 6 and Table II). The expected SD's may be obtained from computer simulations, provided the growth model is known. However, the procedure was just applied to find the most optimal fitting model. A two-pass procedure may solve this problem. In the first pass, the best-fitting model curve is searched on basis of a primary estimate of the standard deviations, taken proportional to the mean order (constant coefficient of variation). The mean-order SD's are then estimated for this model by means of MC simulations for each degree that is

present in the original sample. In the second pass, these mean-order SD's are used to obtain a new best-fitting model curve. This procedure ensures that the final result is obtained with an accurate estimate of the weight factors. Even in the case of a "true" description by the model, the deviations of the actual observations from the model curve may by chance be smaller or larger than the expected SD's. Consequently, the reduced chi-square value may also be smaller or larger than one.

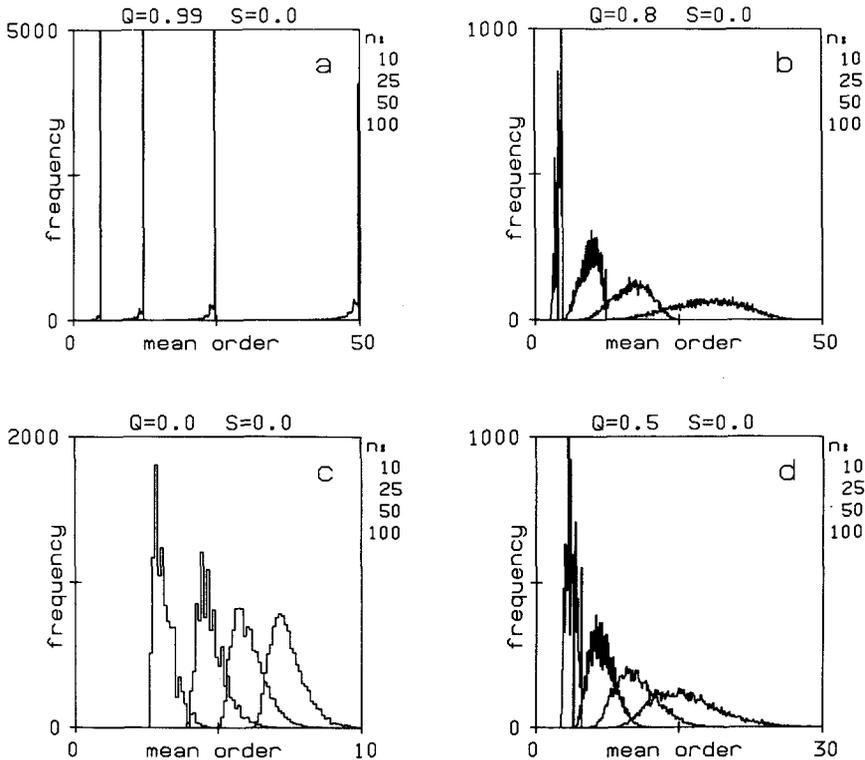


Figure 6. Frequency distributions of mean-order values. Each distribution is calculated for a set of 10,000 trees, obtained by simulating random trees according to a particular growth mode (Q, S) . Each subfigure contains four successive distributions for trees of degree $n=10, 25, 50$ and 100 , respectively. Note that the figures have different scales. (a) Growth mode $(Q, S)=(0.99, 0)$ results in predominantly thin trees and the mean-order values concentrate around the value $n/2$. Figures (a)–(d) show how the shape, the mean and variance of the mean-order distribution depends on the mode of growth.

Alternatively, the mean-order SD's may be estimated by the variance in the data points themselves, which is only possible if the observed sample contains a sufficient number of trees per degree. For a perfect fit one expects that the squared distance between data points and model curve is equal to the variance

in the data points themselves. This means that the reduced chi-square value for a perfect fit has a minimum value of one.

8.1 Example I of mean-order analysis. The first example will use dendritic trees of rat pyramidal (pyr) and multipolar non-pyramidal (mpnp) cells from a study of McConnell and Uylings (e.g. Uylings *et al.*, 1983). The pyr set contains 441 trees in the degree range 4–11, one tree of degree 13 and one of degree 14. The mpnp set contains 483 trees in degree range 4–10 and one tree of degree 11, 12, 13 and 14, respectively. The averaged mean-order values and the standard deviations per degree are displayed in Figs 7c and d. The standard deviations are used for weighting the deviations of all individual mean-order values from the model curve in the calculation of the chi-square value. A best fit for the pyr set was found for the model $Q=0$, $S=0.59$ with a reduced chi-square value of 1.00 (Fig. 7a) and a goodness-of-fit determined by a level of significance of 0.50. A best fit for the mpnp set was found for the model $Q=0$, $S=0.19$ with a reduced chi-square value of 1.01 (Fig. 7b) and a corresponding level of significance ~ 0.50 . A reduced chi-square value of 1.00 in these examples means that the deviations of the observed data points from the model curve are completely explained by the variance in the data points themselves. The model curve itself has not introduced any additional deviations. Therefore, we may conclude that the (Q, S) -model describes both the pyr and the mpnp data set perfectly. The observed averaged mean-order values and the expected ones for the best-fitting model curves are also given in Table III for comparison. Additionally, the expected SD's have been calculated by generating many trees (MC) according to these models. A comparison with the observed SD's shows a very good agreement indicating that also the observed variances are completely explained by the variance sources in the model, i.e. the random choices for the branching segments.

Both sets of pyr and mpnp cells have also been analysed by means of partition analysis resulting in best-fitting model parameter values $(Q, S) = (0, 0.58)$ for the pyr cells and $(Q, S) = (0, 0.18)$ for the mpnp cells. The outcomes of the partition analysis and the mean-order analysis are in very good correspondence. Only complete basal dendrites of pyr and mpnp cells (without cut segments) have been used for the present mean-order and partition analysis. In the experiments of McConnell and Uylings (Uylings *et al.*, 1983), the group of pyr and mpnp cells consists of several sub-groups. In a previous paper (Van Pelt *et al.*, 1986) these sub-groups have already been analysed separately by means of partition analysis in which only subtree pairs with cut segments were discarded. The mean of the best-fitting parameters at $Q=0$ for all the sub-groups was calculated to be $\bar{S} = 0.65 \pm 0.11$ (SEM) for the pyr sub-groups and $\bar{S} = 0.24 \pm 0.09$ (SEM) for the mpnp sub-groups. These results correspond also very well to the present outcomes.

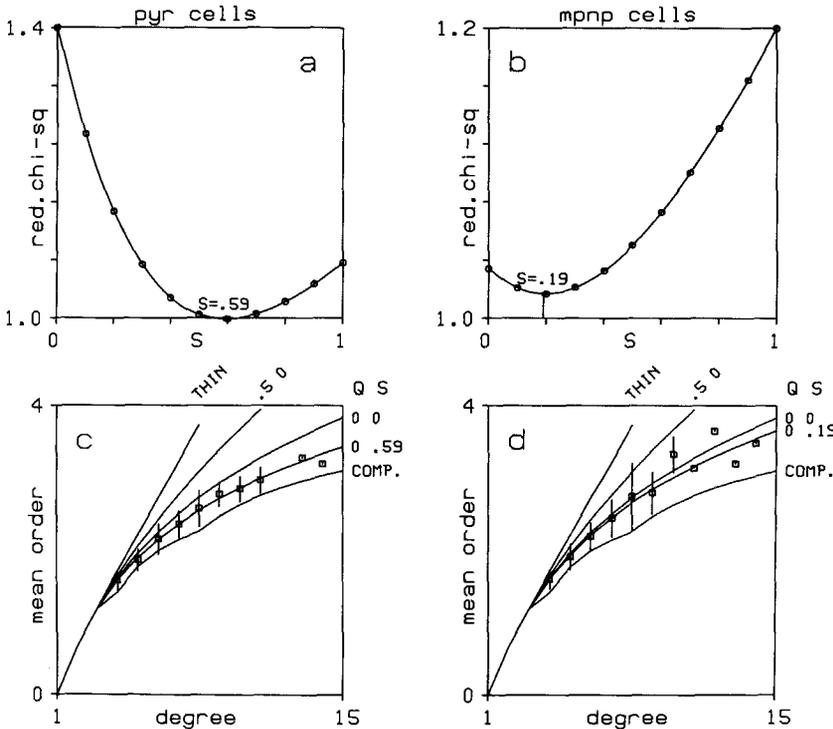


Figure 7. (c) Averaged mean-order values of a set of 443 pyramidal (pyr) cell dendritic trees. The error bars indicate the standard deviations. Degree 13 and 14 were only represented by one tree. Some model curves are included for reference. (a) Reduced chi-square values (circles), obtained from the comparison of the mean-order values of all the individual (pyr) trees with several model curves with $Q=0$. The weights at each degree in the chi-square calculation are taken from the standard deviations in the data subsets themselves. The continuous line is a cubic-spline interpolation and its minimum gives the parameter value of the best-fitting curve, subsequently calculated and plotted in (c). Corresponding results for a set of 487 multipolar non-pyramidal (mpnp) cell dendritic trees are displayed in (d) and (b). Degrees 11, 12, 13 and 14 were only represented by one tree.

8.2 Example II of mean-order analysis. As a second example we have analysed retino-tectal axon terminal arbors from the peripheral, intermediate and central area of the tectum of the adult goldfish displayed in Figs 13–15 of Stuermer (1984). These data have already been used by Verwer *et al.* (1987) to illustrate the estimation of model parameters in the partition analysis. The degree and mean order of the observed trees are listed in Table IV. Best-fitting model curves for the three data sets have been obtained by means of the two-pass procedure and are displayed in Figs 8a–c for the peripheral, intermediate and central area, respectively. The results of both the partition analysis and the mean-order analysis are shown in Table V. The goodness-of-fit of the best-fitting mean-order model curves is illustrated by the levels of significance

TABLE III
Comparison of Averaged Values and Standard Deviations of the
Mean-Orders of Observed Trees and Model Predictions

Degree	no.	pyr cells		$Q=0$ $S=0.59$		no.	mpnp cells		$Q=0$ $S=0.19$	
		Mean	SD	Mean	SD		Mean	SD	Mean	SD
4	111	1.59	0.14	1.61	0.14	247	1.61	0.13	1.60	0.14
5	105	1.87	0.15	1.89	0.17	139	1.92	0.18	1.95	0.20
6	83	2.16	0.22	2.16	0.20	46	2.20	0.20	2.19	0.21
7	64	2.36	0.19	2.36	0.21	30	2.45	0.26	2.47	0.25
8	41	2.58	0.25	2.55	0.21	14	2.75	0.47	2.66	0.29
9	19	2.76	0.17	2.71	0.21	5	2.80	0.29	2.85	0.29
10	13	2.85	0.18	2.87	0.21	6	3.33	0.25	3.02	0.29
11	5	2.97	0.18	3.00	0.20					

(P -values) and shows that the model gives statistically acceptable descriptions of all three samples. From the reduced chi-square values (1.06; 0.56; 0.39) can be concluded that the variance in the observed mean-order values is equal to or somewhat smaller than, but certainly consistent with the variance obtained from MC simulations of the Q - S growth model. There is also a good correspondence in the outcome of both procedures indicating that even for small samples of trees the mean-order and partition analysis give consistent results.

TABLE IV
Axonal Trees from Three Areas of the Goldfish
Tectum (Obtained from Stuermer, 1984)

Peripheral		Intermediate		Central	
Degree	Mean order	Degree	Mean order	Degree	Mean order
25	6.29	9	2.59	12	3.96
30	6.83	10	3.05	15	4.55
33	6.55	18	4.97	18	4.74
34	10.08	19	4.92	24	5.32
37	6.52	20	4.36	25	4.53
40	9.53	27	5.02	32	6.31
42	6.49				

The accuracy of the two-pass procedure has been tested systematically by applying it to sets of simulated trees. Seven trees per run were simulated on basis of the growth model $(Q, S) = (0.40, 0)$, each tree with a degree, at random chosen from the interval [25, 40]. Subsequently, the two-pass procedure was

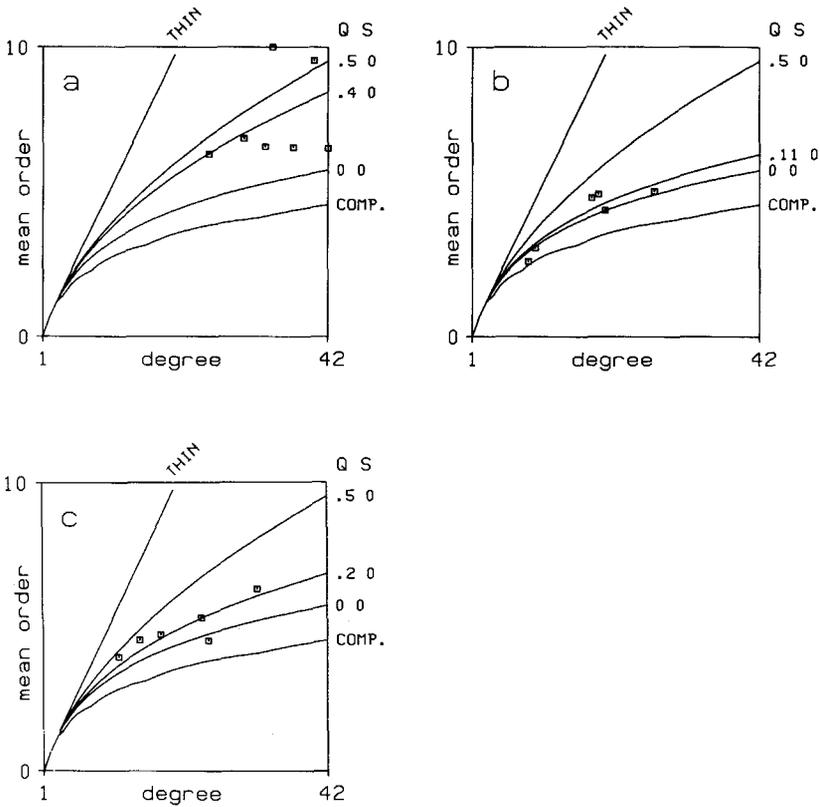


Figure 8. Mean-order values of trees from three sets of axonal branching patterns. The best-fitting model curves for these sets have been obtained by means of a two-pass fitting procedure (described in the text). The axonal branching patterns originate from (a) the peripheral, (b) the intermediate and (c) the central area of the goldfish tectum and are obtained from the work of Stuermer (1984).

used to find the best-fitting growth mode. The results of ten runs are shown in Table VI. The mean value of the estimated \hat{Q} -values is equal to 0.41 ± 0.02 (SEM). The “true” model parameter value ($Q = 0.40$) is within the uncertainty interval, indicating that the two-pass procedure is accurate, without introducing bias. The spread in \hat{Q} -values illustrates to what extent samples of random generated trees for one particular growth mode may differ in their topological properties and how far the estimated \hat{Q} -value for one sample may deviate from the “real” growth mode. The standard deviation (0.08) of the \hat{Q} -value distribution of this test can be used as an estimate of the confidence interval for the outcome of the mean-order analysis of the *P*-area data set. Confidence intervals for the best-fitting model parameters for the *I*-area and *C*-area tree sets are obtained in a corresponding way (Table V, last column).

TABLE V

Estimated \hat{Q} -values by Mean-Order and Partition Analysis of Three Sets of Axonal Trees from the Goldfish Tectum, Obtained from Stuermer (1984). The confidence intervals for the \hat{Q} estimates of the mean-order analysis (last column) are obtained by means of simulations (see text for explanation) and expressed as \pm SD (see also Table VI)

Area	No. of trees	Partition Analysis Verwer <i>et al.</i> (1987)		Mean-Order Analysis			
		\hat{Q}	\hat{Q}	Reduced chi-sq.	<i>df</i>	<i>P</i> -value	Conf. \pm SD
'P'	7	0.42	0.40	1.06	6	0.38	\pm 0.08
'T'	6	0.08	0.11	0.56	5	0.73	\pm 0.09
'C'	6	0.24	0.20	0.39	5	0.86	\pm 0.08

TABLE VI

Mean-order Analysis of 10 Sets of Simulated Trees (7 per Set) According to Growth Mode (Q, S) = (0.40, 0)

Run	1	2	3	4	5	6	7	8	9	10
\hat{Q}	0.48	0.27	0.38	0.38	0.45	0.31	0.45	0.45	0.53	0.35
Reduced chi-square	0.83	1.12	0.69	1.14	0.50	1.62	0.07	0.87	0.53	1.26
Mean - \hat{Q} = 0.41; SD = 0.08; SEM = 0.02										

9. *Measures Related to Mean Centrifugal Order.* There is a one-to-one correspondence between centrifugal order and topological path length. The path from the root up to and including a segment of centrifugal order γ has a length of $\gamma + 1$. Werner and Smart (1973) have studied path lengths in river networks and they introduced mean pathlength and mean exterior pathlength as measures for network topology. Ley *et al.* (1986) use the term generation for the same concept as pathlength and they use the mean generation of terminal segments as a measure of topological structure in the analysis of microvascular networks.

Jarvis (1972) introduced the *E* index as a measure of the topologic structure of dendritic drainage networks. The *E* index is defined as $E = \sum_i n_i p_i / \sum_i n_i$. The numerator sums the product of degree (n_i) and path length (p_i) of all intermediate segments and the denominator sums the corresponding products of all terminal segments.

According to its definition the E index not only incorporates the distribution of order but also of degree of the segments. The order- and degree-distribution of segments, however, are not independent of each other (Uylings *et al.*, 1989) which makes the E index a less easily interpretable measure. This is further shown by rewriting the E index only in terms of the order distribution of the segments. It can be shown that for a tree α the E index can be written as:

$$E_{\alpha} = (3\bar{\gamma}_{\alpha} + \overline{\gamma_{\alpha}^2}) / (2\gamma_{\alpha} + 4).$$

The mean squared order $\overline{\gamma_{\alpha}^2}$ of all segments arises in the definition because of the intermingling of two dependent distributions.

10. Discussion. The order distribution of segments and the mean order appear to be sensitive topological measures of a tree, indicating its elongation with a compact and a thin tree as extremes. The mean order is therefore a useful measure for describing the topological variability in sets of trees. Because the mean order depends on both the size and the connectivity pattern of a tree it is important to separate these aspects, i.e. by averaging only the mean orders of trees of equal size. A sample of trees is then characterized by a size distribution and by a particular relationship between the averaged mean-order values and size (or degree). These characteristics subsequently can be used for mutual comparison of different samples of trees or for testing growth models by comparison with model-predicted mean-order functions. In this paper, only the model approach has been worked out in detail and applied to two experimental data sets. It is described how the mean-order expectation values are obtained and are used in finding the best-fitting growth model by means of a minimum chi-square procedure. Of particular interest was the comparison with outcomes from a partition analysis on the same data sets because partition analysis provides currently in our opinion the most sensitive tool for the analysis of trees. On the basis of the analysis of the experimental data sets in this study we can conclude that both the mean-order analysis and the partition analysis give consistent results with respect to the best-fitting growth model. The mean-order analysis also allows the estimation of the confidence interval of the model parameter value by means of MC simulations. The usefulness of the (Q - S) growth model has further been confirmed by its ability to predict perfectly both the mean-order *vs* degree relation and the mean-order variances for the pyramidal and multipolar non-pyramidal cells in the first example. That means that the variability in topological structures of these branching patterns is completely explained by assuming a growth process, in which branching events occur at randomly chosen terminal segments.

The question if the mean-order analysis is easier to apply than the partition analysis is difficult to answer. The mean order of a tree is a single-valued

measure and a plot of mean orders *vs* the size of the trees gives an immediate visual indication of the tree properties. Moreover, a qualitative interpretation in terms of growth models is possible by visual comparison with model curves. A quantitative analysis, however, requires computer assistance for the calculation of the model curves, (eventually) the estimation of the mean-order variances and the calculation of the minimum chi-square parameter values with their confidence intervals. The mean-order analysis would profit considerably from the availability of analytical expressions for the mean-order expectation values and the standard deviations as functions of the growth mode and size. A start has been made with this paper but further study is required.

In partition analysis the tree is decomposed into all its partitions and loses by this its individual identity. A set of trees is then characterized by a number of partition frequency distributions, one per degree (Van Pelt and Verwer, 1984). These partition distributions do not allow a qualitative visual interpretation in terms of growth models. The quantitative analysis requires extensive computer assistance (Van Pelt *et al.*, 1986; Verwer *et al.*, 1987) although the procedure may be simplified by applying drastic lumping schemes (Verwer and Van Pelt, 1986). On the other hand, the data reduction of trees into partitions involves only a minimal loss of information and the partition analysis may therefore serve as a reference for other analysis procedures.

In conclusion, the mean order is an appropriate measure for the analysis of topological trees and mean-order analysis has some favourable aspects in comparison with partition analysis. Further study is ongoing to decide which measure of tree topology has optimal characteristics for the analysis of trees (e.g. Uylings *et al.*, 1989).

We thank J. van Neerven for valuable suggestions and we are grateful to G. van der Meulen for making the photographs.

LITERATURE

- Abramowitz, M. and I. A. Stegun (Eds). 1965. *Handbook of Mathematical Functions*. New York: Dover.
- Berry, M. and D. Pymm. 1981. "Analysis of Neural Networks." In *Advances in Physiological Science*, Vol. 30. *Neural Communications and Control*, G. Szekely, E. Labos and S. Damjanovitch (Eds). 28th International Congress of Physiological Science, Budapest, 1980. Budapest: Akadémiai Kiadó/Oxford: Pergamon Press.
- Bevington, P. R. 1969. *Data Reduction and Error Analysis for the Physical Sciences*. New York: McGraw-Hill.
- Dacey, M. F. and W. C. Krumbein. 1976. "Three Growth Models for Stream Channel Networks." *J. Geol.* **84**, 153–163.
- Harding, E. F. 1971. "The Probabilities of Rooted Tree-Shapes Generated by Random Bifurcation." *J. appl. Prob.* **3**, 44–77.

- Hollingsworth, T. and M. Berry. 1975. "Network Analysis of Dendritic Fields of Pyramidal Cells in Neocortex and Purkinje Cells in the Cerebellum of the Rat." *R. Soc. London Phil. Trans.* **B270**, 227–264.
- Horsfield, K., M. J. Woldenberg and C. L. Bowes. 1987. "Sequential and Synchronous Growth Models Related to Vertex Analysis and Branching Ratio." *Bull. math. Biol.* **49**, 413–430.
- Jarvis, R. S. 1972. "New Measure of the Topologic Structure of Dendritic Drainage Networks." *Water Resources Res.* **8**, 1265–1271.
- and A. Werrity. 1975. "Some Comments on Testing Random Topology Stream Network Models." *Water Resources Res.* **11**, 309–318.
- Ley, K., A. R. Pries and P. Gaehtgens. 1986. "Topological Structure of Rat Mesenteric Microvessel Networks." *Microvascular Res.* **32**, 313–332.
- Liao, K. H. and A. E. Scheidegger. 1968. "A Computer Model for Some Branching-Type Phenomena in Hydrology." *Int. Assoc. Sci. Hydrology Bull.* **13**, 5–13.
- Press, W. H., B. P. Flannery, S. A. Teukolsky and W. T. Vetterling (Eds). 1986. *Numerical Recipes, The Art of Scientific Computing*. Cambridge University Press.
- Sadler, M. and M. Berry. 1983. "Morphometric Study of the Development of Purkinje Cell Dendritic Trees in the Mouse Using Vertex Analysis." *J. Microsc.* **131**, 341–354.
- Shreve, R. L. 1966. "Statistical Law of Stream Numbers." *J. Geol.* **74**, 17–37.
- Stuermer, C. A. O. 1984. "Rules for Retinotectal Terminal Arborizations in the Goldfish Optic Tectum." *J. comp. Neurol.* **229**, 214–232.
- Triller, A. and H. Korn. 1986. "Variability of Axonal Arborizations Hides Simple Rules of Construction: A Topological Study from HRP Intracellular Injections." *J. Comp. Neurol.* **253**, 500–513.
- Uylings, H. B. M., P. McConnell, A. Ruiz-Marcos, J. Van Pelt and R. W. H. Verwer. 1983. "Differential Changes in Dendritic Patterns of Pyramidal and Non-Pyramidal Neurons After Undernutrition and its Subsequent Rehabilitation." In *The Cell Biology of Plasticity*, Fidia Research Series—Frontiers in Neuroscience. Abst. Bk 1, 295–297.
- , J. van Pelt and R. W. H. Verwer. 1989. "Topological Analysis of Individual Neurons." In *Computer Techniques in Neuroanatomy*, J. J. Capowski (Ed.), pp. 215–240. New York: Plenum Press.
- Vannimenus, J., B. Nickel and V. Hakim. 1984. "Models of Cluster Growth on the Caley Tree." *Phys. Rev.* **B30**, 391–399.
- Van Pelt, J. and R. W. H. Verwer. 1983. "The Exact Probabilities of Branching Patterns under Terminal and Segmental Growth Hypotheses." *Bull. math. Biol.* **45**, 269–285.
- and ———. 1984. "New Classification Methods of Branching Patterns." *J. Microsc.* **136**, 23–34.
- and ———. 1985. "Growth Models (Including Terminal and Segmental Branching) for Topological Binary Trees." *Bull. math. Biol.* **47**, 323–336.
- and ———. 1986. "Topological Properties of Binary Trees Grown With Order-Dependent Branching Probabilities." *Bull. math. Biol.* **48**, 197–211.
- , ——— and H. B. M. Uylings. 1986. "Application of Growth Models to the Topology of Neuronal Branching Patterns." *J. neurosci. Meth.* **18**, 153–165.
- and ———. 1987. "Mean Centrifugal Order as a Measure for Branching Pattern Topology." *Acta Stereol.* **6**, 393–397.
- Verwer, R. W. H. and J. Van Pelt. 1983. "A New Method For the Topological Analysis of Neuronal Tree Structures." *J. neurosci. Meth.* **8**, 335–351.
- and ———. 1986. "Descriptive and Comparative Analysis of Geometrical Properties of Neuronal Tree Structures." *J. neurosci. Meth.* **18**, 179–206.
- , ——— and A. J. Noest. 1987. "Parameter Estimation in Topological Analysis of Binary Tree Structures." *Bull. math. Biol.* **49**, 363–378.
- Werner, C. and J. S. Smart. 1973. "Some New Methods of Topologic Classification of Channel Networks." *Geog. Analysis* **5**, 271–295.

Received 8 June 1988

Revised 3 January 1989