

A Stochastic Language for Plant Topology

MengZhen KANG
LIAMA, Institute of Automation
Chinese Academy of Sciences
100080 Beijing - CHINA
mzkang@liama.ia.ac.cn

Jean-Pierre Quadrat
INRIA Rocquencourt - METALAU
78153 Le Chesnay - FRANCE
Jean-Pierre.Quadrat@inria.fr

Paul-Henry Cournède
Ecole Centrale Paris
Laboratory of Applied Mathematics
92295 Châtenay Malabry - FRANCE
paul-henry.cournede@ecp.fr

Philippe de Reffye
INRIA Futurs - DigiPlante
and CIRAD - AMAP
34398 Montpellier - FRANCE
philippe.de_reffye@inria.fr

Abstract

The article describes a stochastic formal language adapted to the botanical concepts underlying the GreenLab organogenesis model. It is based on stochastic L-systems (parallel rewriting grammars) and on multi-type branching processes: stochastic processes control bud productions and at each growth cycle, each new growth unit is the result of a random variable. This formalism allows determining inductively the generating functions of the resulting plant structures and of the numbers of organs, which fully characterizes the plant development resulting from the elementary stochastic processes of bud productions. The moments of the stochastic distributions of the numbers of organs are also explicitly deduced.

1. Introduction

Plant organogenesis models describe the dynamic creation of organs (leaves, internodes, flowers/fruits) and how they arrange to form plant structures (at least the above ground parts). This botanical knowledge is used as a basis for plant geometric models, that is to say to represent realistic 3D plant architectures in computer graphics [4], [5], [17]. To simulate plant structural development, parallel rewriting grammar language also called L-systems [15] have been broadly used since [19], [16]. The stochastic version of this type of grammar gave interesting results from a simulation point of view by increasing the realistic aspect of the geometric plants [16], [8]. However, the formalism of stochastic grammars has not been taken advantage of to

derive the theoretical distributions of plant structures. In this article, we present a stochastic formal language based on L-Systems adapted to the specific botanical concepts underlying the GreenLab organogenesis model ([7], [6]), principally the plant hierarchical organization and the repetition of elementary constructional units, see [1]. The production rules correspond to the meristematic production of growth units and plant development results from multi-type branching processes, [11]. We thus deduce a method to compute inductively the generating functions of plant structures when the organogenesis grammar is known. We also deduce the moments of the stochastic distributions of the numbers of organs. If the method is illustrated on the GreenLab model of plant growth, the same principles would apply in more general contexts.

The article starts with the presentation of the basic botanical concepts that are used in the GreenLab organogenesis model and explains how a stochastic formal language is derived. The inductive relationship to compute the generating functions of plant structures is then introduced and finally applied to obtain the moments of the numbers of organs.

2. Botanical concepts for GreenLab organogenesis

The GreenLab model of plant growth is a functional-structural model, *i.e.* it combines the description of plant architecture at the organ level and that of biomass production and allocation. Literature is already abundant on this model ([7], [20], [10], [3]), and we mainly focus in this section on the basic botanical concepts underlying the GreenLab organogenesis model.

As explained in [1], the architecture of a plant can be seen as a hierarchical branching system in which the axes can be grouped into categories characterized by a particular combination of morphological parameters. Thus, the concept of *Physiological Age* was introduced to represent the different types of axes. For instance, on coffee trees, there are two types: orthotropic trunk and plagiotropic branches. We need less than 5 physiological ages to describe the axis typology of most trees. The main trunk's physiological age is equal to 1 and the oldest physiological age corresponds to the ultimate state of differentiation for an axis, it is usually short, without branches. Except in some very rare cases, an axis of physiological age p always bears axillary branches of physiological ages $\geq p$ (= in the case of reiterations).

Moreover, plants can be considered as modular organisms that develop by the repetition of botanical entities or constructional units at different scales, see for example [2]. The metamer (or phytomer) is one of these entities and it is chosen as the elementary scale to model plant architectural development. A metamer is composed of an internode bearing organs: buds, leaves, flowers. Depending on species, metamers are set in place rhythmically or continuously.

In the rhythmic case, see Figure 1, the plant grows by successive shoots of several metamers produced by buds. The appearance of these shoots defines the architectural *Growth Cycle*. A *Growth Unit* is the set of metamers built by a bud during a growth cycle. These metamers can be of different kinds and ordered according to botanical rules, like acrotony. For example, most temperate trees grow rhythmically, new shoots appearing at spring. For such plants, and if we do not consider polycyclism and neoformation, the architectural growth cycle corresponds to one year.

Plant growth is said continuous when meristems keep on functioning and generate metamers one by one, see Figure 1. The number of metamers on a given axis (that is to say generated by the same meristem) is generally proportional to the sum of daily temperatures received by the plant, see [12]. The growth cycle is defined as the thermal time unit necessary for a bud to build a new metamer, it can be quite short, corresponding to a couple of days. The growth of tropical trees, bushes or agronomic plants is often continuous.

In both continuous and rhythmic cases, the *Chronological Age* of a plant (or of an organ) is defined as the number of growth cycles it has existed for. In this work, we do not consider time scales that are smaller than the architectural growth cycle. We refer to [7] for more details on the definition of GreenLab organogenesis cycle.

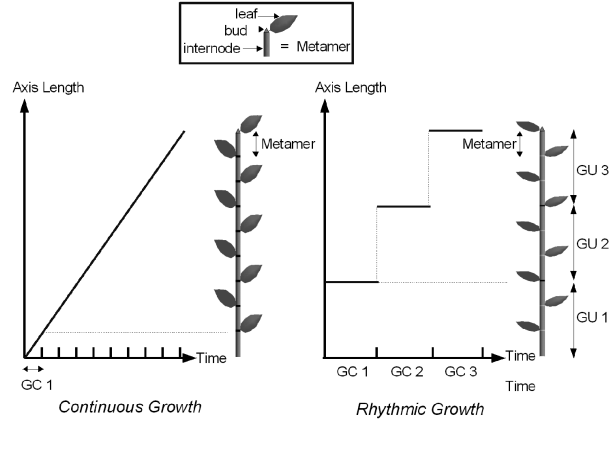


Figure 1. Growth Cycle (GC), Growth Unit (GU) and Metamer in GreenLab. The axis length is given as a function of time for continuous and rhythmic growths, see [3]

3. Describing botanical concepts with a formal grammar

In computational models, plants are generally represented as words in a formal language, more precisely a language based on a generative parallel rewriting grammar also called L-system [15], [19], [16]. In GreenLab, the alphabet \mathcal{A} is given by the set of metamers (terminal symbols) and buds (non-terminal symbols), see [3]. Let m_p and s_p denote respectively a metamer and a bud of physiological age p , with $1 \leq p \leq P$, P being the maximal physiological age of the plant. We have:

$$\mathcal{A} = \{m_1, \dots, m_P, s_1, \dots, s_P\}. \quad (1)$$

We do not consider symbols for organs since the constitution of a metamer is supposed fixed by botanical rules (an internode and a given number of leaves and fruits). If, for example, flowering is particularly studied, symbols denoting flowers would be introduced in the alphabet without difficulty. In the following, we will distinguish $m = (m_1, \dots, m_P)$ (terminal symbols) and $s = (s_1, \dots, s_P)$ (non-terminal symbols).

The set of words endowed with the concatenation operator seen as an internal, non-commutative operation is denoted \mathcal{A}^* . For example, $m_1 s_2 s_1 \in \mathcal{A}^*$ represents a structure composed of an internode of physiological age 1, bearing a lateral bud of physiological age 2 and an apical bud of physiological age 1.

Note that if the physiological age of an axillary bud is always strictly superior to that of its bearing axis (*i.e.* no reiteration), the plant topology can be deduced from the word

unequivocally, without needing additional symbols in the alphabet like brackets which are classically used since [16] to distinguish branches.

$S_p(k)$ denotes a structure of physiological age p and chronological age k . It is defined as the complete plant structure that is generated after k cycles by a bud of physiological age p . In the deterministic case, all the structures with the same physiological and chronological ages are thus identical, they are monomials in \mathcal{A}^* . We recall here that the physiological age of the main trunk is 1. Thus, at growth cycle t , $S_1(t)$, represents the whole plant.

In the stochastic case, the probability distribution of a random structure $S_p(k)$ can be described by its generating function $\mathcal{S}_p(k)$. It is defined as:

$$\mathcal{S}_p(k)(m, s) = \sum_{w \in \mathcal{A}^*} P(S_p(k) = w) w(m, s). \quad (2)$$

where $P(S_p(k) = w)$ is the probability that the random structure $S_p(k)$ is equal to w .

We will denote:

$$\mathcal{S}(k) = \begin{pmatrix} \mathcal{S}_1(k) \\ \vdots \\ \mathcal{S}_P(k) \end{pmatrix}. \quad (3)$$

The generating functions of plant topological structures are multivariate polynomials in the letters of the alphabet. They are non-commutative for the multiplicative (concatenation) operator.

4. Organogenesis and branching processes

We study in this section the dynamics of plant structural development, without considering organ sizes or geometry.

4.1 Production rules

The initial word (start symbol) of the generative grammar describing plant development is given by the plant seed, which is equivalent to s_1 , and the productions rules correspond to the meristematic productions of growth units at each growth cycle, including lateral and apical buds. An example of the deterministic case is given in Figure 2, which illustrates the construction of a simplified coffee tree with two physiological ages and the following deterministic production rules:

$$s_1 \rightarrow m_1 s_2^2 s_1, \quad s_2 \rightarrow m_2 s_2. \quad (4)$$

Note that since the multiplicative sign is used for the concatenation operator, s_2^2 simply means $s_2 s_2$.

In the stochastic case, the production rules may not be

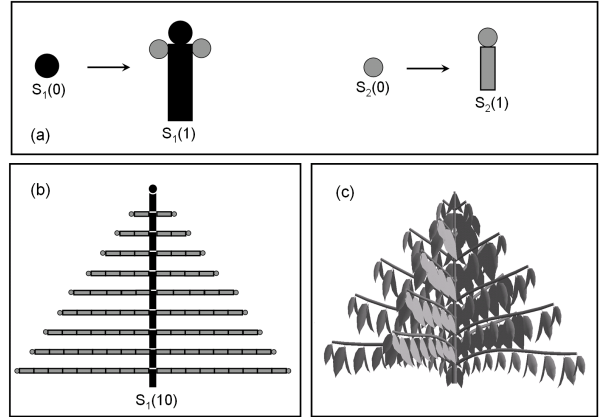


Figure 2. Example of production rules. (a) illustrates the production rules, with rectangles representing metamers and disks representing buds, black is for PA 1 and grey for PA 2. A schematic representation and a 3D drawing of the tree at age 10 are also given in (b) and (c) respectively.

unique for a given symbol and are augmented with probabilities. The sum of probabilities of all the production rules associated to a non-terminal symbol is equal to 1, see for example [16], [14]. In the stochastic organogenesis model of GreenLab, buds of the same type may develop into different growth units [5]. Therefore, several derivation rules are possible for a bud, each with a certain probability. From a botanical point of view, the potential growth units result from the combination of several probabilities concerning bud death or pause, branching and metamer appearance. See [13] for more details. If we go back to the coffee tree example, it was observed [4] that not all the metamers of physiological age 1 bear axillary branches, and the axillary branches of physiological age 2 do not develop at the same rhythm as the main trunk. In a very simplistic way, we could write for example the following rules to take into account these phenomena:

$$\begin{aligned} s_1 &\rightarrow \begin{array}{c|c|c} m_1 s_2^2 s_1 & s_1 & 1 \\ \hline (0.9) & (0.09) & (0.01) \end{array} \\ s_2 &\rightarrow \begin{array}{c|c|c} m_2 s_2 & s_2 & 1 \\ \hline (0.6) & (0.3) & (0.1) \end{array} \end{aligned} \quad (5)$$

where we use the classical notation in language theory: s_1 may produce $m_1 s_2^2 s_1$ with probability 0.9, s_1 with probability 0.09 and may die with probability 0.01. 1 denotes the neutral element for the concatenation operator and may thus be used to characterize bud death.

When a production rule is of the type: $s_p \rightarrow s_p$, it means that the axis elongation pauses, *i.e.* no growth unit is produced by the bud during the growth cycle even if the meristem is still alive. When the probability for the bud to develop into the "complete" growth unit varies with its physiological age, the main and axillary axes will (on average) grow at different rhythms: in the example given in Equation (5), the main trunk grows on average 1.5(= 0.9/0.6) times as quickly as axillary branches.

The parameters of stochastic organogenesis models can be estimated from experimental observations on plant populations and statistical analyses, see [4] for original works or [9] more recently.

4.2 Generating functions

We introduce the generating function $G : \mathcal{A} \rightarrow (\mathbb{R} [\mathcal{A}])^P$ to give the production of buds:

$$G(m, s) = [G_1(m, s), G_2(m, s), \dots, G_P(m, s)], \quad (6)$$

with $G_i(m, s)$ the generating series for s_i , which is a polynomial in (m, s) :

$$G_i(m, s) = \sum_{w \in \mathcal{A}^*} P(s_i \rightarrow w) w(m, s), \quad (7)$$

with $P(s_i \rightarrow w)$, the probability that the production rule ($s_i \rightarrow w$) applies. In the deterministic case, $G_i(m, s)$ are monomials in (m, s) , that is to say $G_i(m, s) \in \mathcal{A}^*$

For the example on the stochastic development of coffee trees given in section 4.1, the components of the generating function G are:

$$G_1(m, s) = 0.9 m_1 s_2^2 s_1 + 0.09 s_1 + 0.01 \quad (8)$$

$$G_2(m, s) = 0.6 m_2 s_2 + 0.3 s_2 + 0.1$$

Plant development can thus be seen as a multitype branching process and the dynamic equations of organogenesis are deduced from the classical equations giving the evolution of these processes [11] as follows.

We can deduce $\mathcal{S}_p(k)$ from $\mathcal{S}_p(k-1)$ by exploring all the possible growth units that the buds of $\mathcal{S}_p(k-1)$ may develop into. It corresponds to compose $\mathcal{S}_p(k-1)$ with the generating function G .

$$\mathcal{S}_p(k)(m, s) = \sum_{w \in \mathcal{A}^*} P(\mathcal{S}_p(k-1) = w) w(m, G(m, s)). \quad (9)$$

This holds for all p , and we can write in a compact way:

$$\mathcal{S}(k)(m, s) = \mathcal{S}(k-1)(m, G(m, s)) \quad (10)$$

From classical properties of multi-type branching processes [11], we also have:

$$\mathcal{S}(k)(m, s) = G(m, \mathcal{S}(k-1)(m, s)) \quad (11)$$

By induction, we can thus compute $\mathcal{S}(k)$ for all k , starting with $\mathcal{S}(0) = s$.

Software developed in Scilab (see www.greenscilab.org) reads the organogenesis parameters of `GreenLab` and computes the generating function of the stochastic plant thus obtained. However, exploring all the possible stochastic plants becomes tedious quite quickly since the number of scenarios increases exponentially with the plant chronological age. The interest of Equation (11) mainly lies in its application to compute the moments of the numbers of metamers or organs in the plant structure as detailed in section 5.

5. Distributions of the numbers of organs

In some cases, the variables of interest are the numbers of organs (for example when studying the functional part of plant growth at the organ level, see [20] for `GreenLab`) and it may prove useful to compute their distributions and their moments. Since we suppose that the composition of each type of metamer is given by botanical rules, we restrict ourselves to studying the numbers of metamers of all physiological ages in the plant.

The generating functions of the numbers of metamers and buds in the structures $\mathcal{S}_p(k)$ is directly deduced from $\mathcal{S}(k)$ by considering the concatenation operator commutative and considering the generating functions $\mathcal{S}(k)$ as real-valued functions of the real vectors m and s .

We can use the classical formula giving the moments of multi-type branching processes from their generating functions [11], [18]. In the following, we will show how to compute the expectations of the numbers of metamers. Moments of higher orders can be derived in the same way.

$\bar{1}$ will denote the vector whose components are all equal to 1 (independently of the vector dimension).

Let $\mathcal{M}(k)$ be a real square matrix of order P giving the expectations of the numbers of metamers of all physiological ages in the plant structures of chronological age k . More precisely, for all $(i, j) \in [1; P]^2$, $[\mathcal{M}(k)]_{ij}$ denotes the expectation of the number of metamers of physiological age j in $\mathcal{S}_i(k)$. Since a structure of physiological age i can not bear metamers of physiological age strictly inferior to i , $\mathcal{M}(k)$ is upper triangular.

We know that:

$$\mathcal{M}(k) = \frac{\partial \mathcal{S}(k)}{\partial m} (\bar{1}) \quad (12)$$

Differentiating Equation (11) with respect to m , and since $\mathcal{S}(k-1)(\bar{1}) = \bar{1}$, we get:

$$\mathcal{M}(k) = \frac{\partial G}{\partial m} (\bar{1}) + \frac{\partial G}{\partial s} (\bar{1}) \frac{\partial \mathcal{S}(k-1)}{\partial m} (\bar{1}) \quad (13)$$

and thus:

$$\mathcal{M}(k) = \mathcal{M}(1) + A \mathcal{M}(k-1) \quad (14)$$

where A denotes $\frac{\partial G}{\partial S}(\bar{1})$.

$\mathcal{M}(1)$ and A being directly obtained from the production rules, $\mathcal{M}(k)$ is deduced by induction for all k .

Since the inductive relationship giving $\mathcal{M}(k)$ is arithmetico-geometric, the explicit expression of $\mathcal{M}(k)$ can be given, provided that $A - I$ is invertible. In such case, we have:

$$\mathcal{M}(k) = [I - A^k] (I - A)^{-1} \mathcal{M}(1) \quad (15)$$

For the coffee tree example with the production rules given by 5, we have:

$$A = \begin{bmatrix} 0.99 & 1.8 \\ 0 & 0.9 \end{bmatrix} \text{ and } \mathcal{M}(1) = \begin{bmatrix} 0.9 & 0 \\ 0 & 0.6 \end{bmatrix} \quad (16)$$

6. Discussion

The article describes a stochastic formal language for the GreenLab organogenesis model. The generating functions of plant structures and of the numbers of organs are derived, as well as the moments of the stochastic distributions of the numbers of organs. These variables are of crucial interest when we want to study biomass production and allocation [20]. However, since organs with different chronological ages have different characteristics as sources and sinks, the alphabet \mathcal{A} should be improved to take into account the chronological ages of metamers, as done in [3].

Likewise, the proposed framework can not deal directly with an important phenomenon observed by botanists, the modification of the morphogenetic characteristics of growth units along an axis during the successive stages of its development [1]. For the GreenLab organogenesis model, it corresponds to a semi-Markov process driving the successive changes of the physiological age of the apical meristem along the axis. We refer to [21] for a detailed description of this phenomenon. As a consequence, the generating function G depends on the state occupancy law, and the proposed framework has to be adapted in order to cope with such cases. It can be done quite easily by increasing the number of physiological ages and adapting the generating functions accordingly. As an example, we consider a mono-stem plant, with deterministic development, two physiological ages, and a deterministic transition from physiological age 1 to physiological age 2 after 5 metamers, see Figure 3. This example could correspond to cereal plants for which the morphology of the first internodes that can bear tillers is different from that of the following ones.

We consider the following alphabet:

$$\mathcal{A} = \{m_1, m_2, s_{10}, s_{11}, s_{12}, s_{13}, s_{14}, s_2\}$$

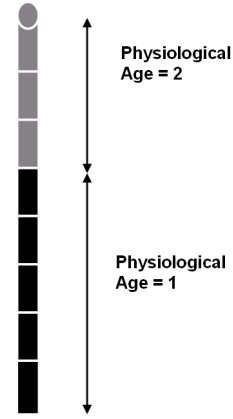


Figure 3. A simple mono-stem plant with transition from physiological age 1 to physiological age 2 after 5 metamers.

and the generating functions:

$$\begin{aligned} G_{10}(m, s) &= m_1 s_{11} \\ G_{11}(m, s) &= m_1 s_{12} \\ G_{12}(m, s) &= m_1 s_{13} \\ G_{13}(m, s) &= m_1 s_{14} \\ G_{14}(m, s) &= m_1 s_2 \\ G_2(m, s) &= m_2 s_2 \end{aligned} \quad (17)$$

With this formulation, the problem of morphogenetic gradients along axes is shown to be a particular case of the general formalism introduced in this article.

A limitation of the stochastic organogenesis model is that the parameters of the stochastic organogenesis processes may strongly be influenced by plant functioning, which is well-known since [4]. This aspect is currently under investigation.

Finally, we restricted ourselves to the study of parallel rewriting grammars. Other languages based on automata that are less formalized but closer to botany have also been developed [5] and could be studied with the same objective of deriving the distributions and moments of the numbers of organs.

7 Acknowledgements

We thank the anonymous reviewers for helping us to improve the manuscript. This work is supported in part by LIAMA (Sino-French Laboratory in Information, Automation and Applied Mathematics), Natural Science Foundation of China (60073007), Chinese 863 program (2006AA10Z229), and PostDoc funding from INRIA granted to M.Z. Kang.

References

- [1] D. Barthélemy and Y. Caraglio. Plant architecture: a dynamic, multilevel and comprehensive approach of plant form, structure and ontogeny. *Annals of Botany*, 99:375–407, 2007.
- [2] A. Bell. *Plant Form. An illustrated guide to flowering plant morphology*. Oxford University Press, 1991.
- [3] P. Cournède, M. Kang, A. Mathieu, J. Barczy, H. Yan, B. Hu, and P. de Reffye. Structural Factorization of Plants to Compute their Functional and Architectural Growth. *Simulation*, 82(7):427–438, 2006.
- [4] P. de Reffye. *Modélisation de l'architecture des arbres par des processus stochastiques. Simulation spatiale des modèles tropicaux sous l'effet de la pesanteur. Application au Coffea Robusta*. PhD thesis, Université Paris-Sud, Centre d'Orsay, 1979.
- [5] P. de Reffye, C. Edelin, J. Françon, M. Jaeger, and C. Puech. Plant models faithful to botanical structure and development. In *Proc. SIGGRAPH 88, Computer Graphics*, volume 22(4), pages 151–158, 1988.
- [6] P. de Reffye, M. Goursat, J. Quadrat, and B. Hu. The Dynamic Equations of the Tree Morphogenesis Greenlab Model. Technical Report 4877, INRIA, 2003.
- [7] P. de Reffye and B. Hu. Relevant Choices in Botany and Mathematics for building efficient Dynamic Plant Growth Models: Greenlab Cases. In *Plant Growth Models and Applications*, pages 87–107. Tsinghua University Press and Springer, 2003.
- [8] O. Deussen and B. Lintermann. *Digital Design of Nature - computer generated plants and organics*. Springer Verlag New York, 2005.
- [9] Y. Guédon, D. Barthélemy, Y. Caraglio, and E. Costes. Pattern analysis in branching and axillary flowering sequences. *Journal of Theoretical Biology*, 212:481–520, 2001.
- [10] Y. Guo, Y. Ma, Z. Zhan, B. Li, M. Dingkuhn, D. Luquet, and P. de Reffye. Parameter optimization and field validation of the functional-structural model greenlab for maize. *Annals of Botany*, 97:217–230, 2006.
- [11] T. Harris. *The theory of branching processes*. Springer-Verlag, 1963.
- [12] H. Jones. *Plants and Microclimate*. Cambridge University Press, 1992.
- [13] M. Kang, P. Cournède, P. de Reffye, D. Auclair, and B. Hu. Analytical study of a stochastic plant growth model: application to the greenlab model. *Mathematics and Computers in Simulation*, 2007. In press.
- [14] W. Kurth and B. Sloboda. Growth grammars simulating trees—an extension of l-systems incorporating local variables and sensitivity. *Silva Fennica*, 31(3):285–295, 1997.
- [15] A. Lindenmayer. Mathematical models for cellular interactions in development. i. filaments with one-sided inputs. *Journal of Theoretical Biology*, 18:280–289, 1968.
- [16] P. Prusinkiewicz and A. Lindenmayer. *The Algorithmic Beauty of Plants*. Springer-Verlag, Berlin,, 1990.
- [17] P. Prusinkiewicz, A. Lindenmayer, and J. Hanan. Developmental models of herbaceous plants for computer imagery purposes. *Computer Graphics*, 22(4):141–150, 1988.
- [18] R. Sedgewick and P. Flajolet. *Introduction to Algorithm Analysis*. Addison-Wesley, 1996.
- [19] A. Smith. Plants, fractals and foraml languages. *Computer Graphics (SIGGRAPH 84 Conference Proceedings)*, 18(3):1–10, 1984.
- [20] H. Yan, M. Kang, P. De Reffye, and M. Dingkuhn. A dynamic, architectural plant model simulating resource-dependent growth. *Annals of Botany*, 93:591–602, 2004.
- [21] X. Zhao, P. de Reffye, D. Barthélemy, and B. Hu. Interactive simulation of plant architecture based on a dual-scale automaton model. In B. Hu and M. Jaeger, editors, *Plant Growth Models and Applications*, pages 144–153. Tsinghua University Press and Springer, 2003.