Modelling pattern formation in plants

M.P. Edwards [a]**, S. Pereverzyev Jr.** [b] **and** <u>**R.S. Anderssen**</u> [c]

[a]*School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW, 2522*
[b]*Industrial Mathematics Institute, Johannes Kepler University Linz, Altenbergerstrasse 69, A-4040 Linz, Austria*
[c]*CSIRO Mathematical and Information Sciences, PO Box 664, Canberra, ACT 2601.*
Email: <u>*Bob.Anderssen@csiro.au*</u>

Abstract: Performing simulations with simple models is a key activity in the recovery of information about the mechanisms which underlie the observed dynamics of biological processes, such as the positioning of the hairs (trichomes) on the leaves of plants. The discovery of a robust representative model is a highly non-trivial process. Without appropriate constraints to regularize the choice of a model, the non-uniqueness of possibilities is vast. As acknowledged by Young (1983) in his modelling of pea leaf development, the degree of non-uniqueness can be reduced by constraining the model to reproduce the patterns observed in mutants as well as on the wild type. Recently, Pereverzyev Jr. and Anderssen (2008) have proposed an explorative combinatorial algebraic ansatz for modelling the genetic signaling, communication and switching controlling the growth of a plant leaf along with the positioning of the hairs (trichomes) on Arabidopsis leaves. The surface of the leaf is modelled as an array of hexagonal cells and the fate of a cell (hair or non-hair) as it grows out from the meristem is determined by a hexagonal recursion anzatz. The application of this simple recursion allows the normal arrangement of trichomes to be generated, while subtle variations of parameters in the model capture mutant forms.

In the earlier work, only the sets of model parameters that lead to symmetric patterns were considered. In this paper, nonsymmetric patterns, which can be obtained using the proposed modelling framework, are investigated. Also, the possibility to capture the patterns with the clumps of trichomes is studied. The leaves of some known mutants show the clumping effects, and reproducing these effects with the existing modeling approaches is problematic. The appropriate modification of the hexagonal algebraic model that gives rise to the clumping patterns is proposed.

The overall goal is an investigation of how modelling and simulation can be utilized to study the genetics of geometry. Here, the geometry of the positioning of trichomes on plant (Arabidopsis) leaves is the model system that is used to formalize and implement ideas.

As in much modelling of complex processes, it is necessary to have appropriate *"link concepts"* that are the stepping stone of the multi-step model that connects the observational data to the information to be recovered. In particular, the role of the link concepts is to identify a logical sequence of steps that connects the observational characterization of the process with the mathematical model that defines how that characterization is related to some assumed internal structure of the process. In the current situation the link concept is the "geometry of the cellular structure where the growth occurs". It is at that location that "the genetics controls the mechanism which in turn orchestrates the cellular growth".

Keywords: Biological modelling, pattern formation in plants, trichomes, meristem, hexagonal recursion

1 INTRODUCTION

Since antiquity, pattern formations in plants have had a fascination for humanity. The lovely patterns of petals on flowers and the arrangement of seeds on sunflowers and pine cones are popular examples. That there were underlying algebraic relationships related to the observed geometric features in such patterns dates from the work of Fibonacci. The study of the regular patterns in the geometrical organization of the branches and flowers of plants now goes under the name of phyllotaxis.

In order to give a rigorous foundation to how biological patterns might have developed, Lindenmayer (1975) proposed a "parallel rewriting" ansatz, which subsequently came to be know as an L-system ("L" for Lindenmayer). Rewriting is an example of a "formal grammar" in that it consists of a set of basic objects along with a set of rules as to how they evolve sequentially. It has been studied and utilized extensively in the study of the architecture of plants, plant organ development and the morphology of various organisms Rozenberg and Salomaa (1980).

Turing (1952), in the year before the publication of the double helix paper by Watson and Crick (1953), established that coupled reaction diffusion equations could generate patterns that could be given a biological interpretation. This has been exploited by a variety of authors in the study of biological pattern formation including plants (Meinhardt (1982, 1984)). Though mathematically important because it lays a foundation for modelling pattern formation from a macroscopic perspective, it avoids the need to model the cellular link between the "genetics" and the "geometry".

Young (1983) proposed a simple compartment ansatz for the modelling of the growth of a pea. The importance of his contribution is two fold:

(a) The modelling was based on the hypothesis that it is the concentration of key hormone signals which defines the state of the developing plant at the "meristem". It represents validation for the view that an essentially simple system of signalling and switching is all that is required to explain biological development. This is consistent with Wolpert's view (Wolpert (1977)) that it is a simple program that controls (plant) development. This is also consistent with the Lindenmayer approach.

(b) With respect to small changes in the parameters, Young's model was able to reproduce known pea mutants. This is consistent with the fact that the growth of an individual plant is a stable process, with the switching on or off of key control processes being the changes which engender mutants.

The relevance of this to the current paper is that it makes more explicit the cellular process associated with (a). This will be done by first proposing a multistage model for the positioning of trichomes on leaves and then investigating how this model might be utilized.

Biology, via DNA and modern molecular biology, has reached the point were the technology and understanding exists which allow a more detailed investigation of the "genetics of geometry" to be undertaken than was possible in the past. As explained by Green et al. (2010) and Coen et al. (2004), it is now possible to propose mechanisms and then explore among the known genes for a possible role for them in controlling geometry.

In order to study the genetics of geometry, from the perspective of identifying how a plants genetics mechanistically controls the geometric shape of plant organs, such as the positioning and growth of trichomes, the first step is the formulation of a conceptual understanding of the process. Because of the known sequentially recursive nature of genetic activity (e.g. a number of interacting "signal → action → switch → signal" recursions), it is clear that various interacting multistage processes are involved. For the modelling of organ growth, it is not possible to go mechanistically from the genetics directly to the geometry in a single step. It is necessary to invoke appropriate intermediate steps. This is implicit in Green et al. (2010) in that they determined, through simulations, that the evolving shape of a snapdragon could be reproduced if the expansion occurred preferentially along particular orientations. This highlights the deductive nature of biological modelling compared with the traditional inductive nature where scientific conclusions about genetic activity are based on correlating genetic changes with the corresponding changes in geometry (or, more generally, some relevant phenotype). As a minimum, in order to explicitly accommodate geometry, any model of the multistage process must involve the following two global sequential steps: (i) a model of

how the genetics controls the activity of the proteins (enzymes and hormones) which are involved in the control of genetic shape. Such information comes from the current inductive studies; and (ii) a geometric model which encapsulates the biology and allows for a geometric structure to develop. It is here that a mathematical model is required. The goal here is to build on the earlier modelling of the positioning of the trichomes on Arabidopsis leaves (Pereverzyev Jr. and Anderssen (2008), O'Keefe et al. (2011)).

2 HEXAGONAL RECURSION

The basic assumption on which the current modelling is based is that the fate of the cells on the leaf of a plant is determined as it grows out of the meristem as illustrated in Figure 1(a). In the subsequent cell division as the leaf grows and expands, it is only the non-trichome cells which divide. A hexagonal cell is assumed in keeping with the results in Table 1 in the paper by Larkin et al. (1996), where the average number of sides of a cell is 6. In general, plant calls are not hexagonal, however, the cells do tend to have a structure that is topologically similar to a hexagonal array. For example, a running rectangular brick array is topologically equivalent to a hexagonal array

It is assumed that the concentration of the signal, controlling whether a cell becomes a trichome, accumulates according to the additive rule of Figure 1(b). In the sequel this formula will be referred to as the *"hexagonal recursion"*. The numerical values generated by this recursion will be referred to as *"hexagonal concentration values"* (HCV).

For such a model, a number of important biological constraints are automatically taken into account:

  (i) The model is cellular, and not some macroscopic model that has smoothed out the cellular details.

 (ii) The fate of a cell is determined by its neighbours (Pennell et al. (1995)).

(iii) As implied in various papers by Wolpert such as Wolpert (1969), any cellular model must respect the known positional information behaviour in the biological development of that part of an organism that is being modelled.

 (iv) Because of the directional nature of the hexagonal recursion, in defining how the concentration of the key signal accumulates, the model automatically involves a polar transport mechanism.

  (v) No trichome forms in a boundary cell of an Arabidopsis leaf (Schellmann et al. (2002)).
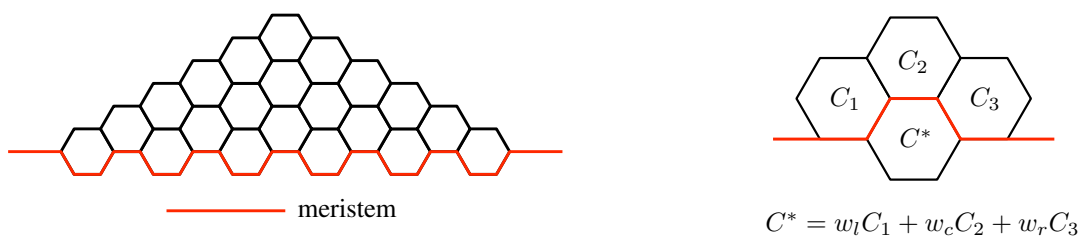

$$C^* = w_l C_1 + w_c C_2 + w_r C_3$$

Figure 1. (a) A hexagonal cell approximation of the epidermal cells on the upper side of a leaf. The top hexagonal cell corresponds to the tip of the leaf. The red line represents the meristem out of which the leaf is growing. Further details can be found in Pereverzyev Jr. and Anderssen (2008). (b) The localized additive relationship that models how the concentration of the signal in the cells above the meristem determine its concentration in the cells forming at the meristem.

Initially, in section 2.1, the basic hexagonal recursion is defined and analysed. It must be robust in that small changes in the value of the defining parameters give only small changes in the patterns generated.

This does not rule out the fact that large changes in the parameter will give large changes in the patterns. The effect of incorporating additional steps that model how the wild type changes are then shown to be able to generate an extensive range of synthetic mutants. Here, the small changes that differentiate between the wild type and mutants are the additional steps which define the modifications of the wild type to produce the mutants. In the sequel, even with only one additional step, a comprehensive set of mutants is generated. More than one step could be used. Through the introduction of additional steps the fact that the molecular biological activity in a mutant is different from that of a wild type.

2.1 The basic hexagonal recursion

For the modelling for the wild type, the basic recursive rule for determining the HCV in a cell and the consequential positioning of trichomes is given by:

B1. Cells located on the periphery of the leaf have the "boundary" HCV P_0.

B2. The leaf starts growing from the 3 tip hexagons in the manner indicated in Figure 1(b). The HCV C^* of the cell forming at the meristem is determined by

$$C^* = w_l C_1 + w_c C_2 + w_r C_3, \quad w_l \geq 0, w_c \geq 0, w_r \geq 0, \tag{1}$$

where $w_l + w_c + w_r \geq 1$, which guarantees that, until reset, the HCV increases in the cell(s) forming at the meristem.

B3. This process continues progressively for each triple of hexagonal cells above the meristem as the leaf step by step grows out of the meristem.

B4. With respect to a specified threshold $T_1 \geq P_0$, if a cell, forming at the meristem as indicated in Figure 1(b), has HCV $C^* \geq T_1$, then the fate of that cell is set to be a trichome.

B5. If, in a triple of hexagonal cells above the one that they are involved in forming at the meristem, the fate of at least one of them has been set to be a trichome, i.e. if at least one of the cells C_1, C_2, C_3 is a trichome, then the HCV C^* is reset to be P_0.

Two trivial asymptotic situations can arise from this basic recursion. When the threshold value $T_1 = P_0$, the boundary value, the threshold is always met and so trichomes are produced in every cell, including the boundary. The other trivial case arises when the threshold value is so high that it is never met and so no trichomes are ever activated. In all other situations, as shown in Figures 2 and 3, no trichomes are located on adjacent cells. This is a direct consequence of the structure of the basic recursion.
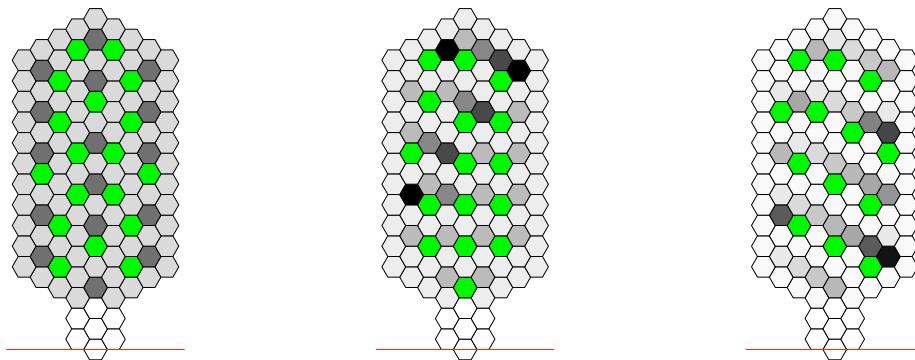

Figure 2. Final distributions generated by the basic recursion with (a) $T_1 = 11$, (b) $T_1 = 22$, (c) $T_1 = 66$, with weights $w_l = 1$, $w_c = 2$, $w_r = 3$ and boundary value $P_0 = 1$. (Note that the shading changes as the value of T_1 changes.)
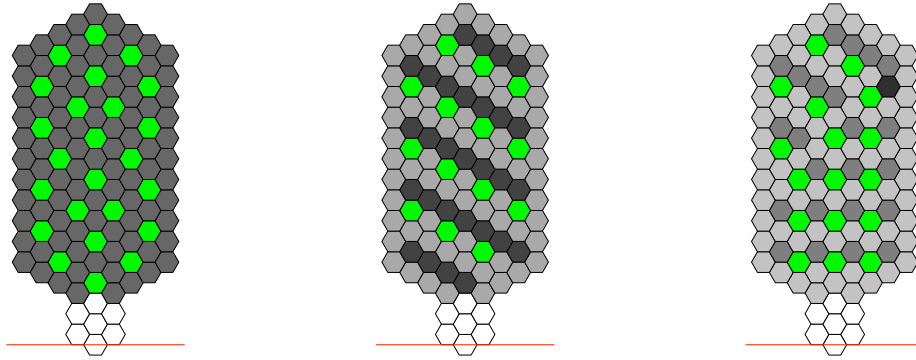
Figure 3. Final distributions generated by the basic recursion with (a) $T_1 = 3$, (b) $T_1 = 5$, (c) $T_1 = 7$, with weights $w_l = 0$, $w_c = 1$, $w_r = 2$ and boundary value $P_0 = 1$. (Note that the shading changes as the value of T_1 changes.)

In Figures 2-3, the HCVs have been determined according to the basic recursion. The colour of the tile represents the corresponding HCV with white corresponding to zero and black to the maximal value $T_1 - 1$. The cells which become trichomes where the HCV matches or exceeds the threshold value T_1 are marked green.

Even if non-symmetric weights $w_l \neq w_r$ are chosen, symmetric patterns (e.g. Figure 2-3 (a)) can form. It is also possible to generate diagonal patterns (e.g. Figure 2(c) and 3(b)). Other patterns start out in an irregular manner but settle into a stable steady-state horizontal pattern (e.g. Figure 2(b) and 3(c)), with the pattern persisting irrespective of the length of the leaf.

2.2 A modified hexagonal recursion

It is clear that in the basic recursion model there is no possibility of generating clumps of trichomes, since as soon as a cell reaches the threshold value to produce a trichome, any adjacent cells below have their value reset to the initial value P_0, and the possibility of a trichome is "switched off". Clumps of trichomes can be generated by adding the following rule. Assume a cell C^* has an upper neighboring cell that became a trichome. If the HCV in C^*, which is obtained by the basic hexagonal recursion, is less then a second threshold $T_2 > T_1$, then the cell C^* becomes a trichome. This establishes that two constraints are required to produce the clumping and, consequently, a modified recursion is needed to generate mutants.

Thus, with the mentioned rule the modified recursion is given by:

M1. Cells located on the periphery of the leaf have the "boundary" HCV P_0.

M2. The process evolves as defined in B2, B3 and B4.

M3. With respect to a specified threshold $T_2 \geq T_1$, if, in a triple of hexagonal cells above the one that they are involved in forming at the meristem, at least one of them has HCV greater than or equal to T_2, then the HCV C^* is reset to be P_0.

Setting $T_2 = T_1$ in the modified recursion model yields the same results as basic recursion. A trivial case arises from the modified recursion model when $T_2 \gg T_1$ and $T_1 > P_0$. The first threshold T_1 is eventually reached, however, the second threshold T_2 is never reached. For example, if $T_1 = P_0 + 1$, the initial threshold is reached immediately and so trichomes are produced in all internal cells.

Figures 4-5 both have $w_l \neq w_r$ and the resulting patterns are generally not symmetric, although the lower part of Figure 4(a) does appear symmetric. It is obvious that clumping of trichomes is possible with the modified recursion model. Holding the threshold value T_1 fixed and increasing the value of T_2 leads to a
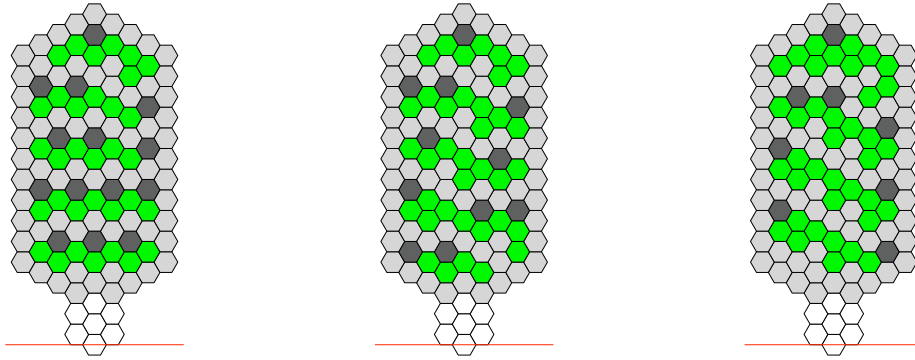
Figure 4. Final distributions generated by the modified recursion when $T_1 = 10$ and (a) $T_2 = 30$, (b) $T_2 = 50$, (c) $T_2 = 100$, with weights $w_l = 1$, $w_c = 2$, $w_r = 3$ and boundary value $P_0 = 1$.
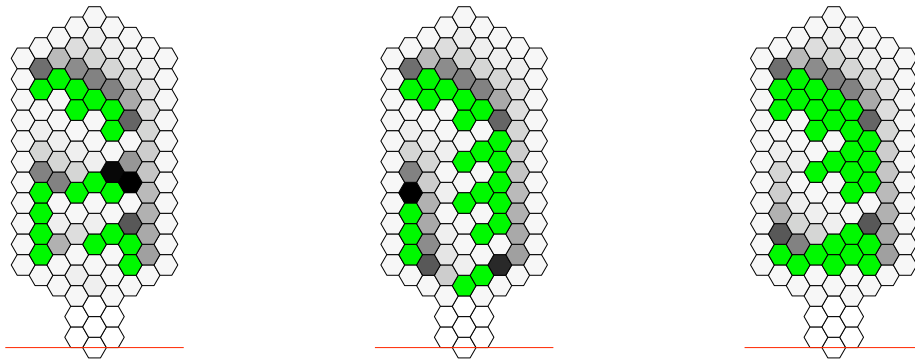

Figure 5. Final distributions generated by the modified recursion with $T_1 = 50$ and (a) $T_2 = 100$, (b) $T_2 = 200$, (c) $T_2 = 500$, with weights $w_l = 0$, $w_c = 1$, $w_r = 2$ and boundary value $P_0 = 1$.

thickening of the groups of cells producing trichomes. Increasing the threshold value T_1 will result in the initial production of trichomes occurring further down the leaf. Setting the left hand weight $w_l = 0$ leads to a trichome pattern which is pushed towards the left hand side of the leaf. Assuming that $w_l = w_c = 0$ (that is, only $w_r \neq 0$) leads to a trichome pattern that is located only on the left hand side of the leaf.

3 CONCLUSIONS

It has been shown that a very simple modification to the basic (wild type) recursion is able to reproduce an extensive variety of mutant patterns. It therefore validates the Wolpert hypothesis (1969,1977) that patterns in biology are the result of simple recursion formulas. Such results are not inconsistent with the large number of genes identified to be involved, as the actual implementation of such simple recursions will involve many components such as one sees in the number of machine language steps to perform simple algebraic manipulations on a computer. Even for the current modified recursion there are more things that can be varied, such as resetting of HCV back to a value other than P_0. Another possibility would be to choose the weights $w_l, w_c, w_r \in \mathbb{R}$. That is, remove the constraint that the weights are non-negative.

REFERENCES

Coen, E., A. Rolland-Lagan, M. Matthews, J. Bangham, and P. Prusinkiewicz (2004). The genetics of geometry. *PNAS 101*, 4728–4735.

Green, A. A., R. Kennaway, A. I. Hanna, J. A. Bangham, and E. Coen (2010). Genetic Control of Organ Shape and Tissue Polarity. *PLOS Biology 8*.

Larkin, J., N. Young, M. Prigge, and M. Marks (1996,). The control of trichome spacing and number in Arabidopsis. *Development 122*(3), 997–1005.

Lindenmayer, A. (1975). Developmental algorithms for multicellular organisms - Survey of L-systems. *J. Theor. Biology 54*, 3–22.

Meinhardt, H. (1982). *Models of Biological Pattern Formation*. Academic Press.

Meinhardt, H. (1984). Models of pattern formation and their application to plant development. *Positional Control in Plant Development*, 1–32.

O'Keefe, C., S. Pereverzyev Jr., and R. Anderssen (2011). The algebra of hexagonal numbers. *The Mathematical Scientist 36*, (1–9).

Pennell, R., Q. Cronk, S. Forsberg, C. Stohr, L. Snogerup, P. Kjellbom, and P. McCrae (1995,). Cell-Contex Signalling. *Phil. Trans Roy. Soc. London, B-Biological Sci. 350*, 87–93.

Pereverzyev Jr., S. and R. Anderssen (2008). Recursive algebraic modelling of gene signalling, communication and switching. *RICAM Report 24*.

Rozenberg, G. and A. Salomaa (1980). *The Mathematical Theory of L-Systems*. Academic Press.

Schellmann, S., A. Schnittger, V. Kirik, T. Wada, K. Okada, A. Beermann, J. Thumfahrt, G. Jurgens, and M. Hulskamp (2002,). TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in Arabidopsis. *EMBO J. 21*, 5036–5046.

Turing, A. (1952). The chemical basis of morphogenesis. *Phil. Trans. R. Soc. London Ser. B-Biol. Sci. 237*, 37–72.

Watson, J. and F. Crick (1953). Molecular structure of nucleic acids - A structure for deoxyribose nucleic acid. *NATURE 171*(4356), 737–738.

Wolpert, L. (1969). Positional information and spatial pattern of cell differentiation. *J. Theoretical Biol. 25*, 1–47.

Wolpert, L. (1977). *The Development of Pattern and Form in Animals*. Carolina Biology Readers, No. 51.

Young, J. (1983). Pea leaf morphogenesis - a simple-model. *Annals Botany 52*, 311–316.