From Morphogenesis to Embryomorphic Engineering

RENÉ DOURSAT

Institut des Systèmes Complexes, CNRS and Ecole Polytechnique, Paris
E-mail: rene.doursat@polytechnique.edu
http://doursat.free.fr

[Abstract: This paper is an extended abstract]

INTRODUCTION

Faced with a rapid growth in size and complexity of computer systems, whether hardware, software or networks, engineers are gradually led to rethink information and communication technology (ICT) in terms of *complex systems*. ICT needs to transition from a state of exogenously imposed order toward increasing organizational and functional autonomy. Instead of designing systems in every detail, engineers should only "meta-design" them, i.e., focus on generic conditions allowing their endogenous growth and evolution. In this context, understanding natural complex systems, in particular multicellular biological development, could help launch a new generation of artificial complex systems based on paradigms still largely absent from traditional engineering, such as decentralization, autonomy and adaptation.

The objectives of "emergent" engineering, however, seem paradoxical: Can autonomy be planned? Can decentralization be controlled? Can evolution be designed? The most important challenge is not simply to allow self-organization to happen, but also to *guide* it. Therefore, conversely, another open research question is to reintroduce programmability and reproducibility into self-organization. But instead of a top-down enforcement of structures, these new controls would take the form of *local* instructions, i.e., a "genotype", inside every agent of the system. Through their genotype, agents can be steered toward displaying specific

characteristics and gradually modified toward new, improved behaviors. The more sophisticated the genotype, the richer the variety and complexity of the overall performance, or "phenotype". Complex systems engineering can thus be reformulated as meta-designing the genotype (at agent level) as opposed to directly designing the phenotype (at system level).

Self-organized systems of abiotic physical-chemical matter generally form "simple" spatial patterns: spots, stripes, branches, waves, etc. Despite a great diversity of phenomena, these emergent structures essentially consist of repeated motifs. Their order arises from amplified fluctuations and shows statistical uniformity and randomness similar to textures. On the other hand, still outside biology, all complex, intricate structures made of segments and parts arranged in specific ways are the product of human inventiveness: computers, cars, buildings, etc. But contrary to physical systems, human constructions are made of a diversity of modules that are statistically heterogeneous and information-rich. Hence, so far, the only emergent *and* nontrivial architectures that we can see around us are living organisms. This is because biological agents (cells) carry a sophisticated microprogram (DNA) that endows them with a vast repertoire of highly non-trivial behaviors. Cells do not randomly mix but proactively position themselves in precise ways. Regions of genetic expression are not randomly distributed but highly regulated in number and position. An organism's shape dynamically unfolds on the basis of calculations and decisions carried out by each cell at every time step. If not a full-fledged deterministic program, DNA is at least a repository of stimuli-response rules, vastly superior in quantity of functional information to elementary volumes of inert physical matter. Therefore, genetic-like regulation at the agent's level could also be the key to controlling self-organization in complex ICT systems.

Darwinian evolution consists of random *variation* followed by non-random *selection*. New trends in evolutionary engineering, such as artificial embryogeny (AE), emphasize the importance of constituting fundamental laws of *development* and developmental variations before these can be selected on the evolutionary time scale. In the framework of genetic algorithms and evolutionary computation, this means an *indirect* or *implicit* mapping (as opposed to direct or explicit) from genotype to phenotype. Understanding variation by comparing the development of different species is also the concern of "evo-devo", a fast growing field of biology. The genotype-phenotype link cannot remain an abstraction if we want to unravel the generative laws of development and evolution—and ultimately transfer them to artificial self-organized systems. Moreover, fine-grain,

hyperdistributed systems (i.e., many light-weight agents, as opposed to a few heavy-weight agents) such as multicellular organisms might be in a unique position to provide a "solution-rich" space needed for successful selection.

MODEL

In this context, the goal of this work is to understand how complex self-organization can be controlled through a computational model of *programmable* and *reproducible* artificial morphogenesis. I propose that, from an abstract viewpoint, self-organized complex morphologies such as biological development can be best understood as a combination of *self-assembly* (SA) and *pattern formation* (PF) under the control of non-random, structured *genetic regulation* (GR) stored inside each agent of a swarm. To take an artistic metaphor, this is similar to mixing "self-sculpting" and "self-painting" in one composition. The differential properties of cells (adhesion, division) are determined by the regions of gene expression to which they belong, while at the same time these regions further expand and segment into subregions due to the self-assembly of differentiating cells. The model can be construed from two different vantage points: either (a) pattern formation on *moving* cellular automata, in which the cells spatially rearrange under the influence of their activity pattern, or (b) collective motion in a *heterogeneous* swarm, in which the agents gradually differentiate and modify their interactions according to their positions and the regions they form.

A brief description of the model follows. First, the motion of a homogeneous swarm (pure SA, Fig. 1a) and the patterning by gradient propagation on a fixed swarm (pure PF, Fig. 1b-c) are introduced separately. Then, these two components are combined to form reproducible growing patterns (SA + PF; Fig. 1d). The genetic program controlling these arrangements inside every agent is also explained. Finally, this combination is repeated as modules ($SA^{(k)} + PF^{(k)}$, Fig. 2e-f) inside a larger, heterogeneous system to create complex morphologies by recursive refinement of details (Fig. 2g).

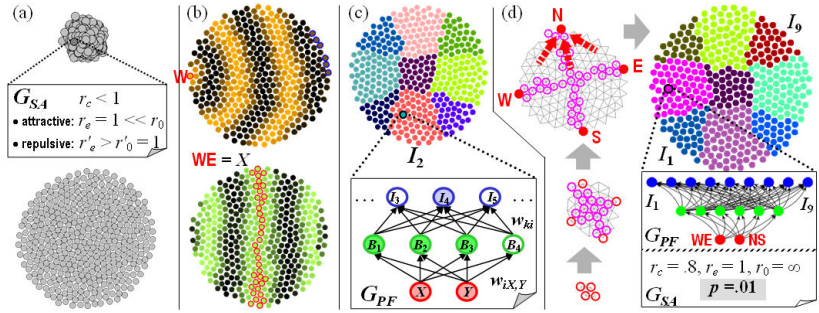

FIGURE 1
Combining self-assembly and pattern formation on one level (see text).

Deployment of a Homogeneous Swarm (SA)

The model consists of a 2-D swarm of agents that integrate two major aspects of cellular biomechanics: cell *adhesion*, in the form of elastic rearrangement, and cell *division*. Agent "shapes" correspond to mutual adhesion affinities, which are implemented by local interaction potentials V among pairs of agents. V consists of three parts: (i) infinite repulsion for $r < r_c$, (ii) quadratic (elastic) attraction around r_e, and (iii) flat potential for $r > r_0$. Starting from a compressed swarm, agents quickly relax to a resting state, in which they tend to form a quasi-regular hexagonal mesh (Fig. 1a). At this stage, each agent possesses fixed "genetic" parameters, r_c, r_e and r_0, denoted by G_{SA}.

Propagation of Positional Information (PF-I) and Programmed Patterning (PF-II)

Pieces of a jigsaw puzzle are defined not only by their position and shape but also by the "image" they carry. In the self-organized swarm, this translates into state variables inside each agent that determine their PF activity. The present model distinguishes between two kinds of PF-specific state variables: *gradient* variables (PF-I) and *pattern* variables (PF-II).

Gradient values propagate from neighbor to neighbor and establish *positional information* across the swarm (Fig. 1b). For example, an agent W passes a counter variable $n_W = 0$ to its neighbors, changing it to 1 on the way. These neighbors in turn instruct their neighbors to set n_W to 2, and so on. The result is a roughly circular wave pattern of n_W values centered on W. Together with W, three other gradients, E, N and S form a 2-D coordinate system.

Pattern values are then calculated on top of gradient values, thereby creating different *agent types*, which in turn will affect the SA behavior (see SA + PF integration below). This process represents the emergence of *heterogeneity*, i.e., the segmentation of the swarm into "identity regions" (Fig. 1c). Each cell contains a *gene regulatory network* G_{PF} that represents the genetic parameters of the PF process, and each identity region corresponds to a high level of expression of a particular gene I, function of variables W, E, N, S above.

Simultaneous Growth and Patterning (SA + PF)

After describing the self-assembly of a non-patterned swarm and the patterning of a fixed swarm, SA and PF are combined to create growing patterns (Fig. 1d). Agents continually adjust their positions according to the elastic SA constraints, *while* continually exchanging gradient values and PF signals over the same dynamic links. This dual dynamics is guided by both genotypes G_{SA} and G_{PF}. Another mechanism, *cell division*, is also introduced at this point. Any agent A may produce a new agent B with probability p. Agent B inherits all of A's attributes, including genotype G_{SA+PF} and internal PF variables. It immediately starts contributing to SA forces and PF gradients, which maintain the pattern's consistency at all times in the swarm.

Modular, Recursive Patterning (PF[k])

Embryological patterns, however, do not develop in one shot but in numerous incremental stages. An adult organism is produced through modular, recursive patterning. In the case of *Drosophila*, regions that acquire leg, wing or antenna identity ("imaginal discs") start developing local coordinate systems of morphogen gradients to form the planned organ. In the present model, the above gene network G_{PF} is extended to include a pyramidal *hierarchy* of network modules able to generate patterns in a recursive fashion (Fig. 2e, g). First, the base network G_{PF} establishes main identity regions as before, then a few subnetworks $G_{PF}^{(k)}$ triggered by nodes I_k in G_{PF} further partition these regions into smaller identity compartments at a finer scale. Modularity is a feature as desirable in the genotype as in any other software architecture or evolvable system. It seems that biological evolution also discovered this principle naturally.

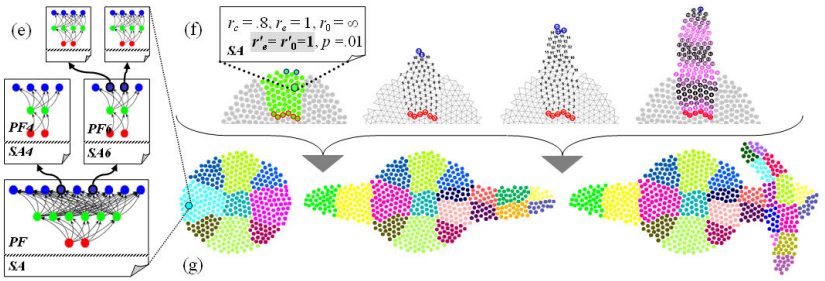

FIGURE 2
Modular growth and patterning of whole systems (see text).

Modular, Anisotropic Growth (SA[k])

So far missing from the model is a true topological *deformation* dynamics, or "morphodynamics", that can confer non-trivial shapes to the organic system beyond simple blobs. To this aim, agents must be able to diversify their SA characteristics, *depending on their PF type and spatial position*, thus closing the feedback loop between SA and PF. In particular, they have to exhibit *inhomogeneous*, *anisotropic* cell division (varying p) and *differential* adhesion (varying V). For example, the growth of limb-like structures can be achieved by a coarse imitation of meristematic plant offshoots (Fig. 2f). In this process, only the tip or "apical meristem" of the organ is actively dividing at any time. Moreover, potential V is attractive only among agents within the limb region (green in Fig. 2f), while it is repelling between the limb and other areas (gray). Just like inhomogeneous division, differential adhesion is an essential condition of complex shape formation.

Modular Growth and Patterning (SA[k] + PF[k])

Putting everything together, full morphologies can develop and self-organize from a few agents (Fig. 2g). These morphologies are *complex*, *programmable* and *reproducible*. They are architecturally complex because they can be made of any number of various modules and parts that are not necessarily repeated in periodic or trivial ways. They are programmable phenotypes emerging from the same genotype carried by every agent of the swarm. They are also reproducible, as their morphological structures are not left to chance but dictated by the genotype. The exact agent positions at the microscopic level are still random, but not the mesoscopic and macroscopic regions that they form. The modularity of the phenotype is also a direct reflection of the modularity of the genotype: the hierarchical SA + PF

dynamics recursively unfolds inside the different regions and subregions that it creates. Each $SA^{(k)} + PF^{(k)}$ block can be reused by exact *duplication*. It can also *diverge* from other blocks, i.e., receive different internal genetic SA and PF parameters potentially giving each region a different morphodynamic behavior and activity landscape. Duplication followed by divergence is the basis of *serial homology* (e.g., vertebrae, teeth, digits), a major natural evolutionary mechanism. The integration between SA and PF is controlled through the identity nodes I_k: just as these nodes turn on gene expression activity in subordinate $G_{PF}^{(k)}$ modules to create new local segmentation patterns, they also simultaneously turn on behavioral changes in subordinate $G_{SA}^{(k)}$ modules to create new morphodynamical behaviors.

CONCLUSION

This study is inherently interdisciplinary, as it closely follows biological principles at an abstract level, but does not attempt to model detailed data from real genomes or organisms. Thus, it lies at crossroads between different families of works, from developmental and systems biology to artificial life, in particular spatial computing, evolutionary programming and swarm robotics. It is an original attempt to integrate the three mechanisms of SA, PF and GR discussed above. Very few previous theoretical models of biological development or bio-inspired artificial life systems have combined them in various ways.

Naturally, beyond the proof-of-concept simulations presented here, a more systematic exploration is needed. Next steps should involve the mass-production of virtual organisms to support (a) statistical analysis of shape and (b) evolutionary search based on module variation and function. First, future work should investigate *functional* meta-design: once a self-developing infrastructure is mature, what computing capabilities can it support? What do its cell-agents and organ-regions actually represent in practice? In fact, the problem is most often reverse: there is a need for precise self-formation capabilities in many systems made of otherwise functionally computing agents, whether processor-carrying micro-units, software agents, robot parts, mini-robots, etc. Second, after growth and function, one must also define how the system *evolves*, i.e., how it *varies* (randomly) and how it is *selected* (nonrandomly). Different selection strategies are possible, either focusing on pre-specified forms, pre-specified

functions, or allowing unspecified outcomes. Reconciling the antagonistic poles of “planning” and “autonomy” ultimately hinges on two complementary aspects: (a) fine-grain variation-by-mutation mechanisms yielding a large number of search paths and (b) loose selection criteria yielding a large number of fitness maxima. With more search paths covering more fit regions, evolution is more likely to find good matches.