DOI: http://dx.doi.org/10.7551/978-0-262-31709-2-ch022

# An artificial lizard regrows its tail (and more): regeneration of 3-dimensional structures with hundreds of thousands of artificial cells

Alessandro Fontana[1] and Borys Wróbel[1,2]

[1]Evolving Systems Laboratory, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland, and
[2] Systems Modelling Laboratory, IO PAS, Sopot, Poland
fontana@evosys.org, wrobel@evosys.org

**Abstract**

Biological multicellular structures can not only self-generate from a single cell but also self-regenerate after damage. In this paper we investigate self-regeneration in a model of artificial development, Epigenetic Tracking. 3-dimensional cellular structures grown using our model reach a size and a level of complexity unmatched by other models in the field, thanks to several features of Epigenetic Tracking. One of these features is that only a small fraction of cells in the body, called drivers, orchestrate development. In this paper we use the mechanim for the generation of drivers based on the diffusion of morphogens as a foundation of several new mechanisms in Epigenetic Tracking, and show that these mechanisms allow for self-regeneration after removal of arbitrarily large portions of the multicellular body.

## Introduction

Models of development can be divided into two broad classes: grammatical and chemical. The first class includes L-systems, introduced by Lindenmayer (10) to model plant growth, and models based on context-free or context-sensitive grammars, instruction trees or directed graphs (e.g., 1; 7; 8). Grammatical models can generate surprisingly complex, life-like shapes, even though they do not include mechanisms corresponding to the biological processes working at the molecular level. In contrast, chemical models (e.g., 2; 9; 11) do include such mechanisms, inspired by information processing inside cells (gene regulatory networks) and by communication between cells (diffusion of chemical substances; considered already by Alan Turing, 1952). Although some models of development can be considered either grammatical or chemical, the division between these classes is fuzzy, and much can be achieved by combining features of both to bring computational efficiency on one hand, and biological plausibility on the other. Such combination of features stands behind Epigenetic Tracking (3).

In Epigenetic Tracking, self-generation of 3-dimensional multicellular structures starts from a single cell containing a genome that encodes all the information necessary to direct development. The genome can be evolved using a genetic algorithm with a fitness function measuring the proximity of the final structure (the body) to a target shape. The complexity of the bodies obtained with Epigenetic Tracking (number of fine morphological features and patterning) and their size (reaching millions of cells) has not been matched yet by any other model. Self-generation of such large and complex bodies is possible thanks to the division of cells into normal cells and drivers: a small fraction of cells that orchestrate development.

We have recently introduced a new mechanism into Epigenetic Tracking, a mechanism for the generation of drivers, based on the diffusion of chemical substances (morphogens; 6). In the present paper we introduce several additional new mechanisms, guided by the assumption that regeneration replays the events that occurred during development (13, see also 2). We show that these additional mechanisms allow to add self-repair to the list of phenomena we previously investigated in our system (ageing and cancer, 4; and the hypothetical transfer of information between somatic cells and the germline, 5; Fontana and Wróbel).

## Epigenetic Tracking: a model of evolving, self-generating multicellular structures

In Epigenetic Tracking multicellular bodies consist of cube-shaped cells on a 3-dimensional grid. The growth starts from a single cell and continues through a pre-specified number of developmental stages. Cells belong to two categories: *normal* and *drivers*. Each driver has an associated array of digits, called *mobile code*. All the cells carry the same genome, an array of characters (from a 4-letter alphabet). The mobile code can be considered as an abstraction for the set of regulatory factors present in a cell: it allows drivers to behave differently despite sharing the same genome.

The genome consists of developmental genes, which all have a left part and a right part. The left part contains three fields: *switch*, which specifies if a gene is active or inactive, *timer*, and *mobile sequence*. At each developmental stage, the mobile sequence of each developmental gene is compared with the mobile code value of each driver, and the timer is compared with the value of the current developmental stage. If both match, the right part of the gene fires. Each

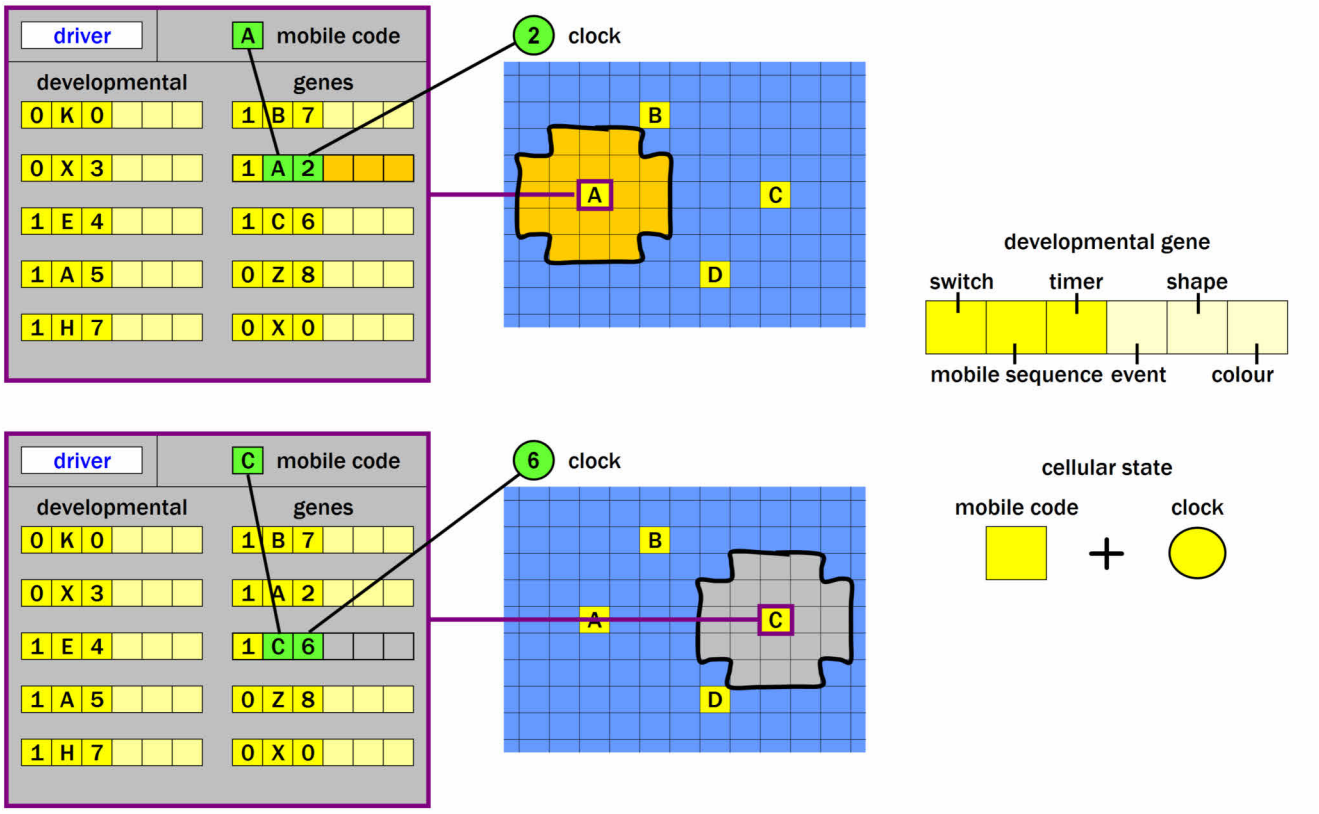


Figure 1: Orchestration of developmental events by drivers in Epigenetic Tracking. On the top, a driver (one of the cells in yellow) with mobile code A triggers a proliferation event at developmental stage 2. On the bottom, a driver with mobile code C triggers a cell death event at stage 6. Black lines indicate the match of the fields in the left part of the genes with the mobile codes of the drivers and the clock. The shape of the created or deleted cell masses is encoded in the right parts of genes. The schematic genome has 10 genes.

right part has three fields. One field determines the type of event – local proliferation or death (removal of cells from the grid; Fig. 1). Another field specifies the shape of the local structure created by proliferation or removed by cell death. The third field determines the phenotype of the normal cells produced in case of proliferation (which is represented by their colour).

A proliferation produces normal cells and drivers. Drivers are placed among normal cells after proliferation, and are much fewer than normal cells in number. Each new driver obtains a new and unique mobile code. The creation of drivers in the experiments described in this paper relies on the diffusion of morphogens, belonging to a finite number of types (we recently implemented this mechanism in Epigenetic Tracking, 6). After a driver orchestrates a developmental event, it persists in the structure and becomes a source of the morphogen which had the lowest consertation at the driver's position when this cell was activated; the diffusion of this morphogen will contribute to the chemical landscape in the body and influence the creation of future drivers (Fig. 2).

The concentration of a particular morphogen ($C$) in the body follows the equation $C = 9 - round(D/G)$, where $D$ is the Euclidean distance between a given position on the grid and the closest driver that produces the morphogen, and $G$ is a system parameter. If the formula gives a negative number, the concentration is set to 0. Because of the rounding, the concentrations can take an integer value between 0 and 9, and positions close-by can have the same concentrations of all morphogens. The cell in the centre of each such region (determined by averaging the coordinates of all the cells there) is a candidate for becoming a new driver. The drivers are created after sorting the regions by size, in the largest region first, provided that each new driver is sufficiently far away from the closest existing driver (this distance is also a system parameter).

The mobile code of the new drivers is derived from the code of the driver that created it through proliferation, but also includes information about the concentration of morphogens in the region. To do so, the mobile code is sep-

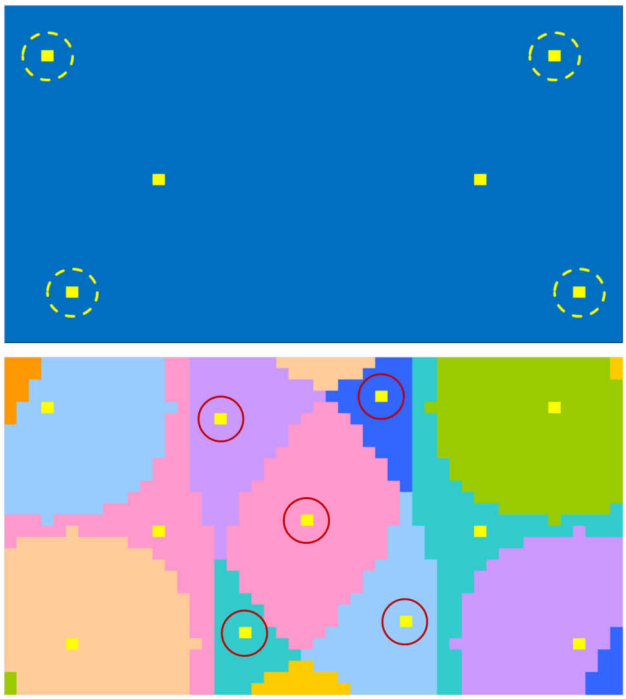

Figure 2: Generation of new drivers based on diffusion of morphogens. Driver cells which have orchestrated an event (yellow cells circled in top panel) persist after doing so in the structure and produce morphogens. There is a finite number of morphogens and their concentrations are rounded, so the structure divides up into multicellular regions (each marked with a different colour in the bottom panel) having the same concentrations of morphogens. New drivers (indicated with red circles) form from the central cells in such regions, if these cells are sufficiently distant from any other driver.

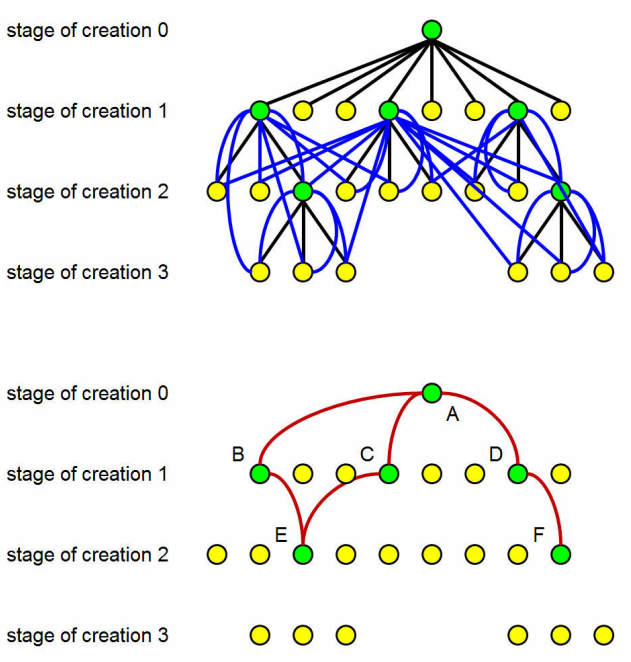


Figure 3: Dependencies among drivers. The cells turned into drivers are originally produced in proliferations orchestrated by other drivers (these dependencies are represented by black lines on the top panel), but their formation is influenced by morphogens produced by other drivers (blue lines), with the exception of drivers created in the first stage (which are not created using morphogens; see text for details). This information can be used to outline a dependency graph (bottom panel) between drivers which were activated during development (shown in green), and can be reactivated for regeneration. For example driver E depends on driver cells B, and A.

arated into as many sub-fields as there are developmental stages, and the concentrations are encoded in the sub-field corresponding to the current stage. For example, if the concentration values are [4,7,8,2], the number encoded in this sub-field (using a 4-digit positional code) will be 4782. The parameters of the system can be varied to ensure a sufficiently high driver density as to preserve evolvability (6).

Because the first proliferation originates from one initial driver cell (the zygote), morphogen regions are absent, and another mechanism is needed at this point to form new drivers: they are placed using a pre-specified pattern. This initial placement of drivers can be seen as deriving from morphogen gradients in the egg itself (maternal factors, 13).

## Four new mechanisms in Epigenetic Tracking to allow for self-regeneration

Because of the central assumption of our model of regeneration (that the state of the body at the start of regeneration has to be similar to some state during embryonic development), the drivers that orchestrated development need to persist in the body – they will be needed to orchestrate the events during regeneration. These drivers, once they trigger a developmental event, are kept in a deactivated state (so that this event is not triggered again). This is the first new mechanism introduced in our system.

The second mechanism permits detection of damage. All cells created in a proliferation event send chemical signals to the deactivated driver which created them. If many of these cells are destroyed, this driver can be reactivated because it receives less of these signals. A large damage can result in the reactivation of many drivers.

Before such reactivation can occur, however, the debris left by the damage needs to be removed, so that the structure contains no cells created during development by the driver that is reactivated, or indeed any cells whose creation depended on such a driver (using a dependency graph like

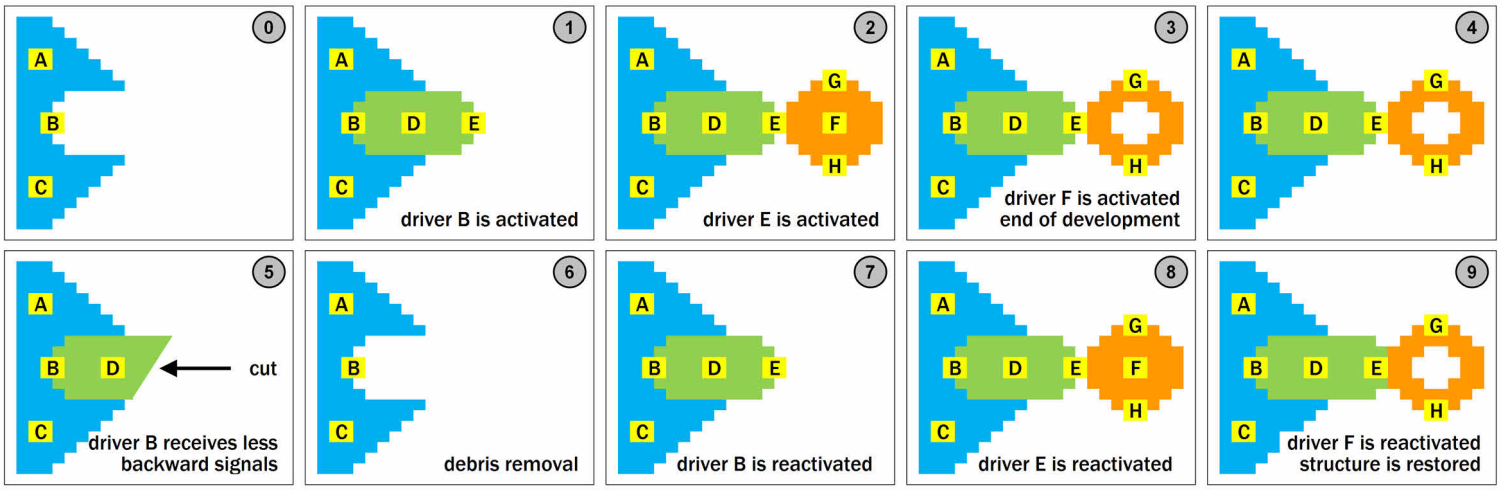


Figure 4: Regeneration of a body part. Panels (and stages) 0-4: the developmental trajectory for a hypothetical body part. After the end of development (stage 4), a portion of the structure is cut at stage 5. After debris removal (stage 6), driver B is reactivated, leading to the recreation of drivers E, which is also reactivated, and so on, until the structure is completely regrown.

the one in Fig. 3). Should such cells remain, some parts of the structure could be duplicated or regeneration would not work correctly. This debris removal is the third new mechanism.

The fourth mechanism recreates exactly the same landscape of morphogens as during development, by exluding from the sources of morphogens all the drivers activated originally (during development) after the driver that is now reactivated during regeneration. For example, if the reactivated driver was originally activated at stage 5, a driver activated at stage 12 would be excluded.

Finally, we needed to deactivate one of the mechanisms that were present in previous versions of Epigenetic Tracking: in the version presented here, proliferation does not cause the cells present at this point in the structure to be pushed away, neither during development nor during regeneration (otherwise the morphogen landscape would not be recreated). When a profileration is triggered by a driver embedded in the existing structure, the new cells are placed in the grid only if the relevant positions are free, so that no old cell is deleted. The proposed mechanisms ensure that the drivers activated to produce a given body part during development are reactivated during regeneration, leading to the same sequence of events (Fig. 4).

## Results and Discussion: Self-generation and self-regeneration of multicellular structures in Epigenetic Tracking

We have run 10 independent simulations of evolution using a genetic algorithm, with constant pupulation size (124 individuals). At the first generation, all genomes were random. At each subsequent generation the genomes of the individuals in the new population were created as follows: (i) the genomes for the 16 best invidual in the previous population were copied from the new population without change (elitism); (ii) 96 genomes were inherited from the previous population with selection probability proportional to each individual's fitness, crossover (one point, with 50% probability) and mutation (the rate is 0.005 per character in the genome); (iii) 12 genomes were created competely randomly. This influx of random genomes introduces new genes into the population to increase evolvability. Another measure to increase evolvability in Epigenetic Tracking, called germline penetration, creates genes with mobile sequences that match mobile codes in the driver cells which have not been activated during development. and inserts them into the genome of next generation's individuals (see Fontana and Wróbel, for the discussion of the biological motivation for this mechanism).

The fitness function rewarded the proximity of the adult structure (the structure after 12 developmental stages) to the shape of a lizard, requiring about 500 000 cells. Each simulation was run for 40 000 generations, and in all simulations the best individual was very close to the target (Fig. 5; only two champions representative for 10 are shown; all the champions had gemomes with 80-100 developmental genes). Then, portions of the final structure were removed and allowed to regenerate (Fig. 5) following the debris removal. We performed tens of such experiments, with different damages for different champions, and in all simulations the regeneration was perfect.

Our model of regeneration draws the inspiration from the amazing properties displayed by many biological organisms. The regenerative capabilities in the living world occupy a

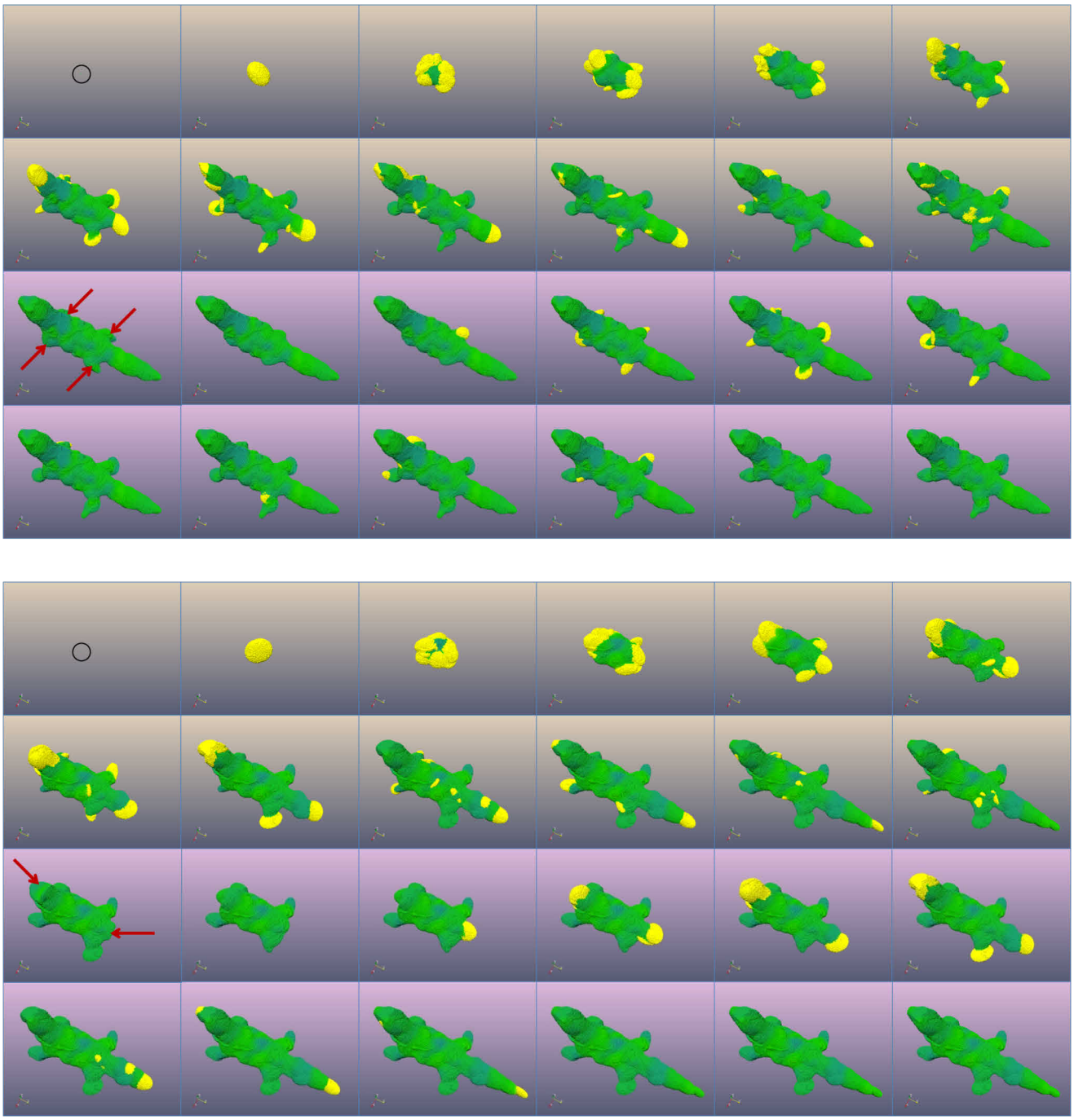

Figure 5: Self-generation and self-regeneration of multicellular structures consisting of hundreds of thousands of cells: the main contribution of this paper. Two champions of independent evolutionary runs are shown in two separate panels, each with 24 frames. In both cases, development unfolds from a single cell (circles) in 12 developmental stages (frames 1-12). After stage 12, the four limbs (top panel) or the tail and the head (bottom panel) are cut off (arrows). After the debris is removed, the structures are completely regenerated.

wide spectrum, ranging from limited cell renewal within tissues, to full regeneration of entire multicellular organisms starting from small fragments. The results obtained in this paper correspond to the latter, most complex extreme of this spectrum, perhaps raising doubts about the biological plausibility of the mechanisms we proposed.

We have built our model of regeneration on two foundations: (i) drivers who were active during development persist, deactivated, in their original positions at later stages of life; (ii) morphogens released from drivers created during later development stages than the damaged part are excluded during regeneration. Both foundations are required to recreate the same conditions during regeneration as the conditions during development in order to guarantee perfect regrowth after damage. A less perfect regrowth – more common in nature – will result, we suspect, if the mechanisms we introduced in Epigenetic Tracking do not work perfectly. We predict that then the regeneration will work better for more elongated parts of the structure (such as limbs or a tail), corresponding to more "isolated" driver subsets, i.e. with less dependencies on subsets in other body parts. More central body parts (e.g., the trunk) correspond to driver subsets with more dependencies, so they require a profound debris removal. We plan to investigate in our future work if these intuitions agree with simulation results for imperfect mechanisms in Epigenetic Tracking.

Are both foundations described above biologically plausible? Drivers in Epigenetic Tracking are inspired by biological embryonic stem cells (Fontana and Wróbel). Embryonic stem cells are totipotent cells, able to differentiate into all cellular types, while adult stem cells are pluripotent cells persisting throughout life, dividing when there is a need to replenish died cells or to regenerate damaged tissues. So persisting quiescent drivers in Epigenetic Tracking can be compared to adult stem cells. On the other hand, perfect debris removal and recreation of the chemical landscape present during development is not entirely plausible, and the fact that it is not may explain why biological organisms with high complexity have a limited ability to regenerate whole parts of their body. Regardless of the degree of biological plausibility, the mechanisms we introduced for simulated development in Epigenetic Tracking could be implemented in artificial physical systems, provided that physical building-blocks able to store genetic information are available.

## Conclusions

Our model of multicellular development, Epigenetic Tracking, allows to self-generate 3-dimensional structures consisting of millions of cells, structures with shapes that have a level of detail unmatched by other models of artificial development. We presented in this paper a new version of Epigenetic Tracking, in which the drivers – cells that correspond to biological organisers, fewer in number than other cells in the structure – are created using diffusing morphogens, persist through life, and can orchestrate regrowth after damage. The presence of the new mechanisms in the model did not impair evolvability, and allowed for perfect regeneration after damage. We plan to investigate in future work if the regeneration will be less perfect – like in highly complex biological organisms – if these mechanism do not work perfectly. These future work will aim to infer general rules for regeneration in biological and artificial systems.

## References

De Garis, H. (1999). *Artificial embryology and cellular differentiation*. Academic Press.

Eggenberger Hotz, P. (2003). Exploring regenerative mechanisms found in flatworms by artificial evolutionary techniques using genetic regulatory networks. In *Proceeding of the Congress on Evolutionary Computation (CEC '03)*, volume 2, pages 2026–2033.

Fontana, A. (2008). Epigenetic tracking, a method to generate arbitrary shapes by using evo-devo techniques. In *Proceedings of the 8th International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems (EPIROB '08)*.

Fontana, A. (2009). Epigenetic tracking: biological implications. In *Proceedings of 10th European Conference on Artificial Life (ECAL 2009)*, volume 5777 of LNCS, pages 10–17.

Fontana, A. (2010). A hypothesis on the role of transposons. *Biosystems*, 101:187–193.

Fontana, A. and Wróbel, B. A model of evolution of development based on germline penetration of new "no-junk" DNA.

Fontana, A. and Wróbel, B. (2013). Morphogen-based self-generation of evolving artificial multicellular structures with millions of cells. *Proceedings of the Annual Conference on Genetic and Evolutionary Computation (GECCO 2013), in press*.

Gruau, F., Whitley, D., and Pyeatt, L. (1996). A comparison between cellular encoding and direct encoding for genetic neural networks. In *Genetic Programming 1996: Proceedings of the first annual conference*, pages 81–89.

Hornby, G. S. and Pollack, J. B. (2002). Creating high-level components with a generative representation for body-brain evolution. *Artificial Life*, 8:223–246.

Joachimczak, M. and Wrobel, B. (2008). Evo-devo in silico: a model of a gene network regulating multicellular development in 3d space with artificial physics. In

*Proceedings of the 11th international conference on the simulation and synthesis of living systems (ALife XI)*, pages 297–304.

Lindenmayer, A. (1968). Mathematical models for cellular interaction in development. *Journal of Theoretical Biology*, 18:280–289.

Miller, J. F. and Banzhaf, W. (2003). Evolving the program for a cell: from french flags to boolean circuits. In *On growth, form and computers*. Academic Press.

Turing, A. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society*, 237:37–72.

Wolpert, L. and Ticke, C. (2010). *Principles of development*. Oxford University Press.