Design Patterns of Pattern Formation and Morphogenesis in a Declarative Programming Language

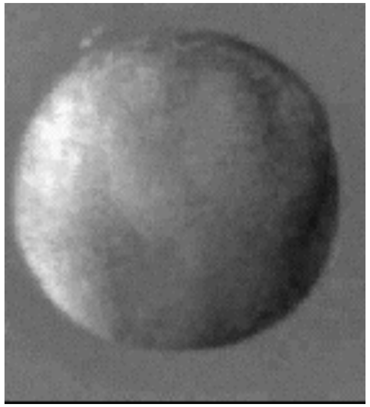
Jean-Louis Giavitto

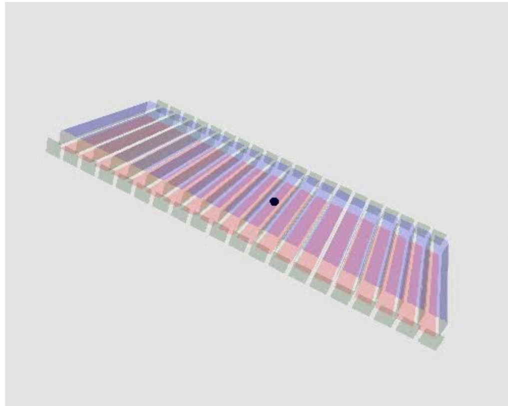
CNRS

IBISC, Université d'Evry

LIS team, MGS project

http://mgs.spatial-computing.org





IBISC
Informatique, Biologie Intégrative & Systèmes Complexes
CNRS – Université d'Evry - Genopole

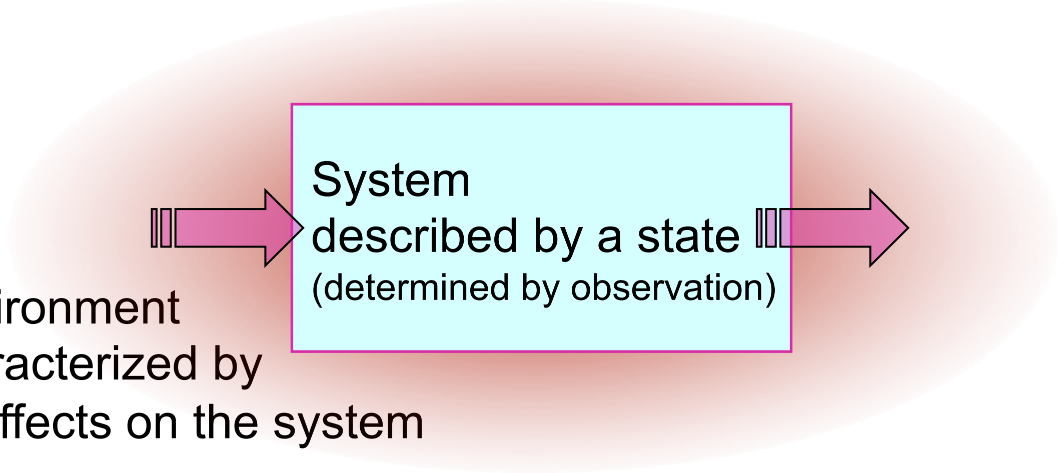
Modelling morphogenesis: the approach of A. Turing

THE CHEMICAL BASIS OF MORPHOGENESIS

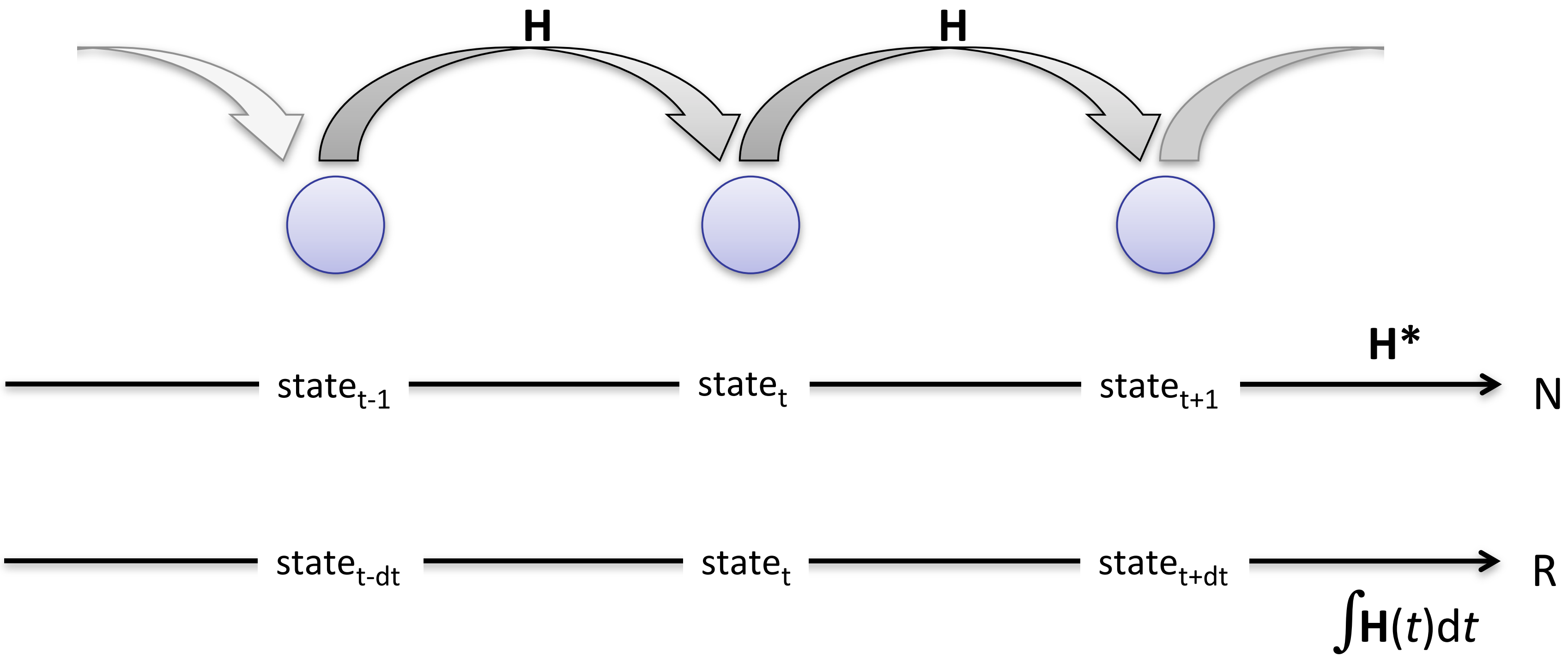
By A. M. TURING, F.R.S. *University of Manchester*

(*Received* 9 *November* 1951—*Revised* 15 *March* 1952)

With either of the models one proceeds as with a physical theory and defines an entity called ‘the state of the system’. One then describes how that state is to be determined from the state at a moment very shortly before. With either model the description of the state consists of two parts, the mechanical and the chemical.

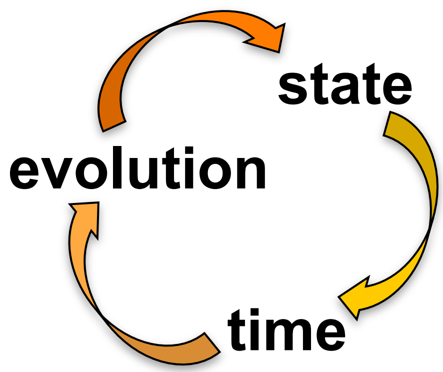


Specifying a dynamical system (for simulation)



Specification of

- **structure of state**
- **structure of time**
- **evolution function**



Morphogenesis as a **Dynamical System**

Modelling a dynamical system

- state, including space (e.g. fields)
- time
- evolution function

C : continuous, D: discrete	**PDE**	**Coupled ODE**	**Iteration of functions**	**Cellular automata**	**…**
state	C	C	C	D	…
time	C	C	D	D	…
space	C	D	D	D	…

Modelling morphogenesis: the approach of A. Turing

The model takes two slightly different forms. In one of them the cell theory is recognized but the cells are idealized into geometrical points. In the other the matter of the organism is imagined as continuously distributed. The cells are not, however, completely ignored, for various physical and physico-chemical characteristics of the matter as a whole are assumed to have values appropriate to the cellular matter.

- **Uniform matter, continuous-oriented system description**
  One choice is to ignore cells completely, e.g., Physiome models tissues as continua with bulk mechanical properties and detailed molecular reaction networks, which is computationally efficient for describing dense tissues and non-cellular materials like bone, extracellular matrix , fluids, and diffusing chemicals, *but not for situations where cells reorganize or migrate.*

versus

- **Cell-oriented discrete system description**
  Multi-cell simulations are useful to interpolate between single-cell and continuum-tissue extremes because cells provide a natural level of abstraction for simulation of tissues, organs and organisms.

  Treating cells phenomenologically reduces the millions of interactions of gene products to several behaviors: most cells can move, divide, die, differentiate, change shape, exert forces, secrete and absorb chemicals and electrical charges, and change their distribution of surface properties.

(CompuCell3D manual)

Aggregate- *vs.* **Entity-based** models

Modelling morphogenesis: the predefined medium

The interdependence of the chemical and mechanical data adds enormously to the difficulty, and attention will therefore be confined, so far as is possible, to cases where these can be separated.

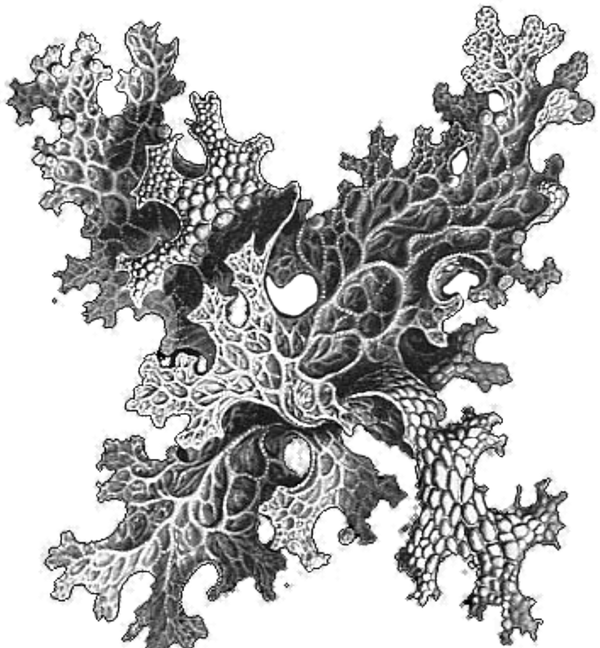
Suppose, for instance, that a 'leg-evocator' morphogen were being produced in a certain region of an embryo, or perhaps diffusing into it, and that an attempt was being made to explain the mechanism by which the leg was formed in the presence of the evocator. It would then be reasonable to take the distribution of the evocator in space and time as given in advance and to consider the chemical reactions set in train by it.

Compatible with

- the notion of morphogenetic field
- cell fate

But

- there is evidence for
  feedback loops between the shape and the process inhabiting the shape



from E. Haenkel (cited by C. Goodman-Strauss): example of a negative curvature surface. Curvature can be controlled while the surface is growing along a 'front'

The medium/process problem

In determining the changes of state one should take into account

(i) The changes of position and velocity as given by Newton's laws of motion.

(ii) The stresses as given by the elasticities and motions, also taking into account the osmotic pressures as given from the chemical data.

(iii) The chemical reactions.

(iv) The diffusion of the chemical substances. The region in which this diffusion is possible is given from the mechanical data.

The medium/process problem

In determining the changes of state one should take into account

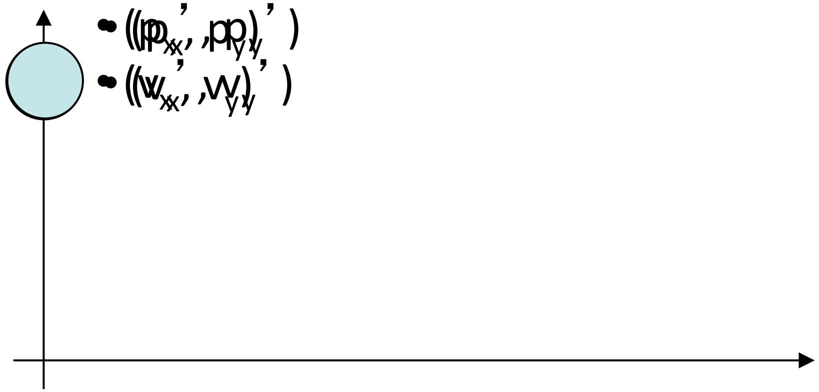
(i) The changes of position and velocity as given by Newton's laws of motion.

(ii) The stresses as given by the elasticities and motions, also taking into account the osmotic pressures as given from the chemical data.

(iii) The chemical reactions.

(iv) The diffusion of the chemical substances. The region in which this diffusion is possible is given from the mechanical data.

a falling ball



at any time a state is a position and a speed

A dynamical system (DS)

The medium/process problem

In determining the changes of state one should take into account

(i) The changes of position and velocity as given by Newton's laws of motion.

(ii) The stresses as given by the elasticities and motions, also taking into account the osmotic pressures as given from the chemical data.

(iii) The chemical reactions.

(iv) The diffusion of the chemical substances. The region in which this diffusion is possible is given from the mechanical data.

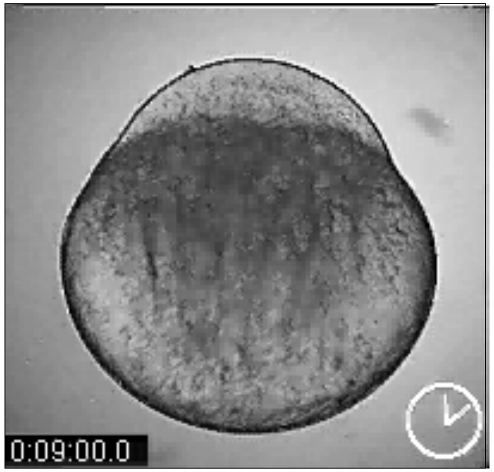
a falling ball

- (p_x', p_y')
- (v_x', v_y')

at any time a state is a position and a speed

A dynamical system (DS)

a developing embryo

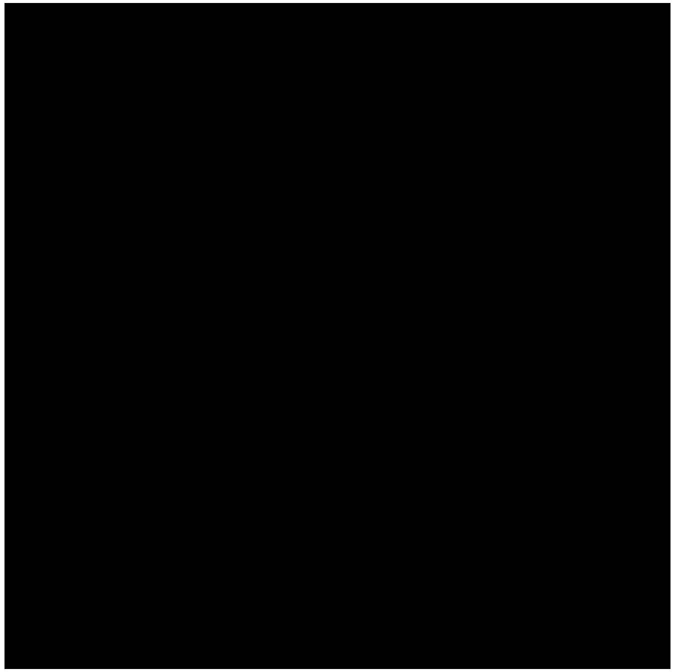


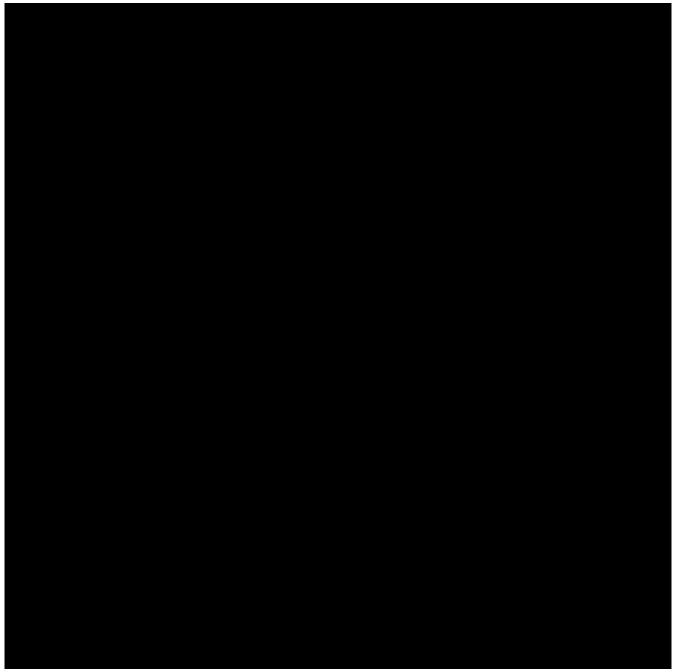
the structure of the state
(chemical and mechanical state of each cell)
is changing in time

A dynamical system with a dynamical structure $(DS)^2$

What has changed since Turing's time

It might be possible, however, to treat a few particular cases in detail with the aid of a digital computer. This method has the advantage that it is not so necessary to make simplifying assumptions as it is when doing a more theoretical type of analysis. It might even be possible to take the mechanical aspects of the problem into account as well as the chemical, when applying this type of method. The essential disadvantage of the method is that one only gets results for particular cases. But this disadvantage is probably of comparatively little importance.





P. Prusinkiewicz, c.2003
Diffusion and reaction in a deformable surface (E. Coen's *expanding canvas* metaphor). Spring-mass system. No topological change.

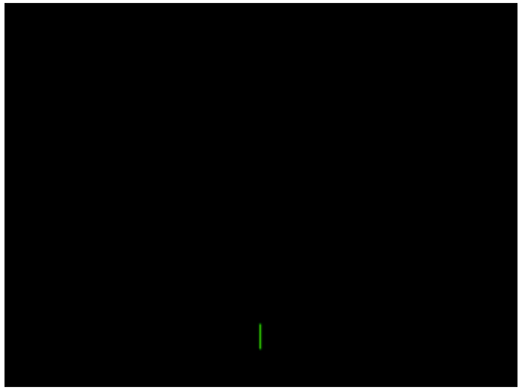
Software as Science ?

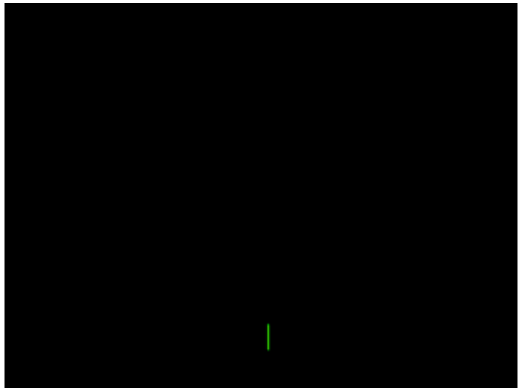
M.-P. Cani & F. Bertails

- Intelligibility
  The entire process should be accessible for analysis into a finite, not very large number of stages, each stage being represented as a monotonic function of some definite initial conditions and a single variable such as time, or distance, etc. (Gurwitsch, 1944)
  → compress behavior or shape in few rules
- Simulation is only a first step: models must enable reasoning
  → stay close to mathematical formalism

A program is a formal object (and some form of reasoning on it is possible) but a 10^6 lines of codes is not an explanation !

A good example of **declarative** formalism: Lindenmayer systems





P. Prusinkiewicz

- The structure of a tree can be coded by a string of parenthetised symbols
- A symbol is an elementary part of the plant
- The symbol between [and] represents a sub-tree
- Additional conventions are used to represent main axis, orientation, depth, etc.
- A rule

  $$s_0 \longrightarrow s_1\ s_2\ s_3\ \ldots$$

  represents the evolution of s_0

Diffusion and reaction in a linear *growing medium*

M. Hammel and P. Prusinkiewicz (1996)

The following rules state that a differentiated cell (heterocyst) returns to a vegetative state if the concentration of the activator is too low. In addition, if the cell is large enough, it continues to grow.

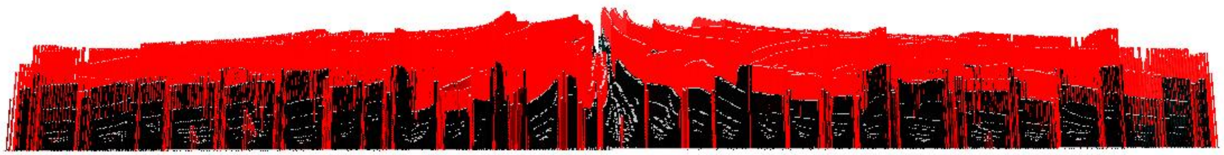
```
e / (D(e) & (e.a < thr) | (e.x >= shorter*gr))
=> {type ="C", a=e.a/gr, h=e.h/gr, x=e.x*gr, p=e.p};
```

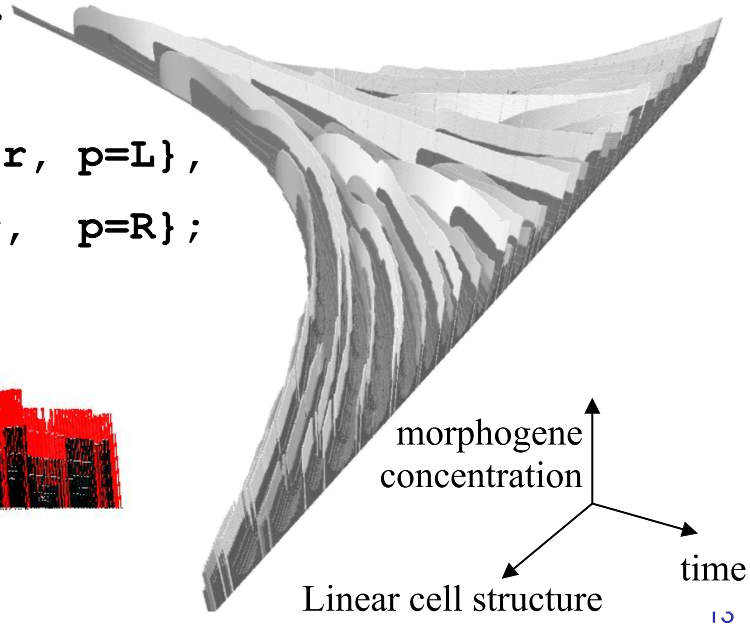
The following rule specifies when a cell with a left polarity divides.
Only vegetative cells can divide (hence the predicate C in the rule guard) and it must be large enough. The volume of the two daughter cells remains the same, so there is no variation in the concentration.

```
e / (C(e) & (e.x >= lm) & (e.p == L))

=> {type="C", a=e.a, h=e.h, x=e.x*shorter, p=L},

   {type="C", a=e.a, h=e.h, x=e.x*longer,  p=R};
```

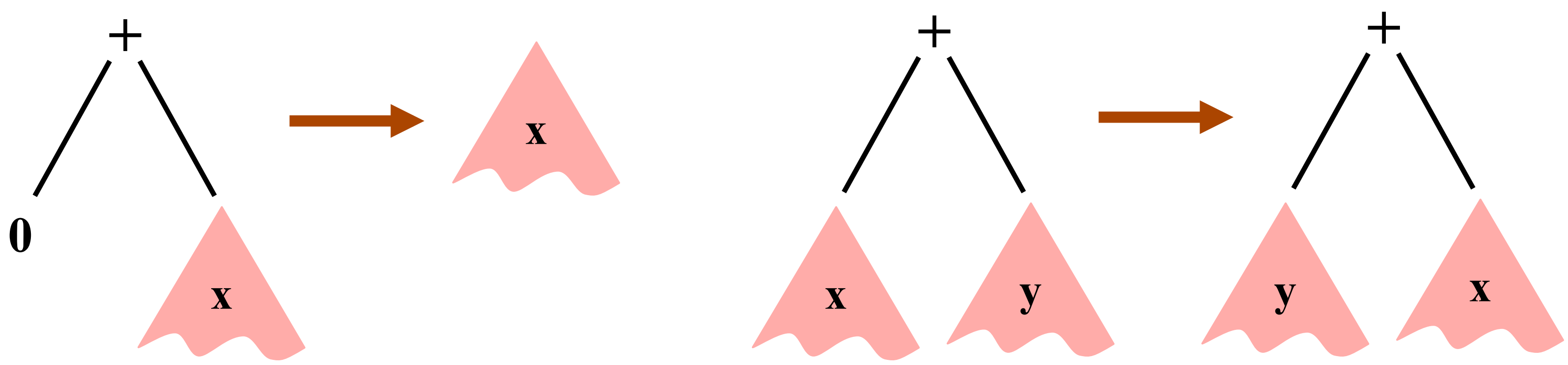




Rewriting systems (and abstract transition systems)

- Rewriting system
  - Used to formalize equationnal reasoning
  - A generative device (grammar)
  - Replace a sub-part of an entity by an other
  - Set of rewriting rules $\alpha \rightarrow \beta$
    - α: pattern specifying a sub-part
    - β: expression evaluating a new sub-part

- Example: arithmetic expressions simplification



A non-standard interpretation

$$e_1 + e_2 \rightarrow \ldots$$

e_1 can be a cell and e_2 a signal

e_1 and e_2 can be interacting cell

$+$ is the possibility of *interaction* between entities (or some other relationships)

$\rightarrow$ is the passing of time, a local evolution, a transition, the concretization of the interaction

Examples: if *e* is a cell and *i* a biochemical signal

$e + i \rightarrow e'$ growth (evolution of *e* on signal *i*)

$e + i \rightarrow e + i'$ quorum sensing

$e + i \rightarrow e' + e''$ division

$e + i \rightarrow .$ apoptose

Complex systems ←→ Rewriting techniques

Modelling	Specification
State (space)	**Data structure**
hierarchical and tree organizations	formal trees (or terms)
arbitrary complex organizations	***?***
Evolution function	**Set of rules**
interactions → evolution	α : pattern → β : expression
local **evolution laws**	rewriting rules

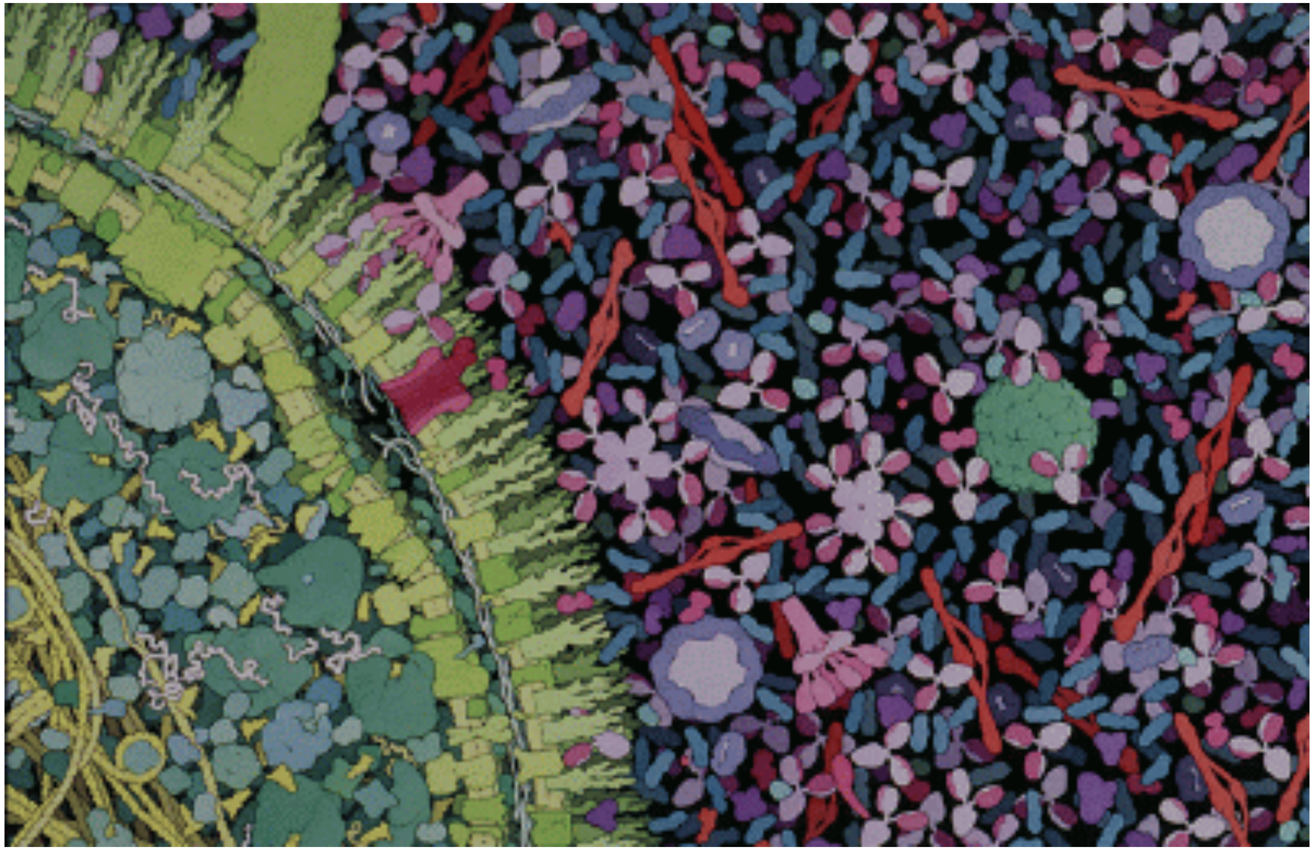
Simulation	Application
Trajectories	Derivations
Time management	Rule application strategy
discrete, event-based, synchronous vs. asynchronous	maximal parallel, sequential, deterministic, stochastic

Properties

- *local evolution rules*
  mandatory when you cannot express a global function/relation because the domain of the function/relation is changing in time
- interaction based approch
  the l.h.s. of a rule specifies a set of elements in *interaction*, the r.h.s. the result of the interaction
- *the phase space is well defined but not well known*
  a generative process enumerates the elements but membership-test can be very hard
- *various kind of time evolution* (for the same set of rules)
- *demonstration by induction*
  on the rules or on the derivation (e.g. growth function in L system)

How to extend to arbitrary spatial structure?

- Anabaena was « easy » because of the linear uniform structure
- How to handle the complex spatial structure of a cell?



David S. Goodsell

The MGS project

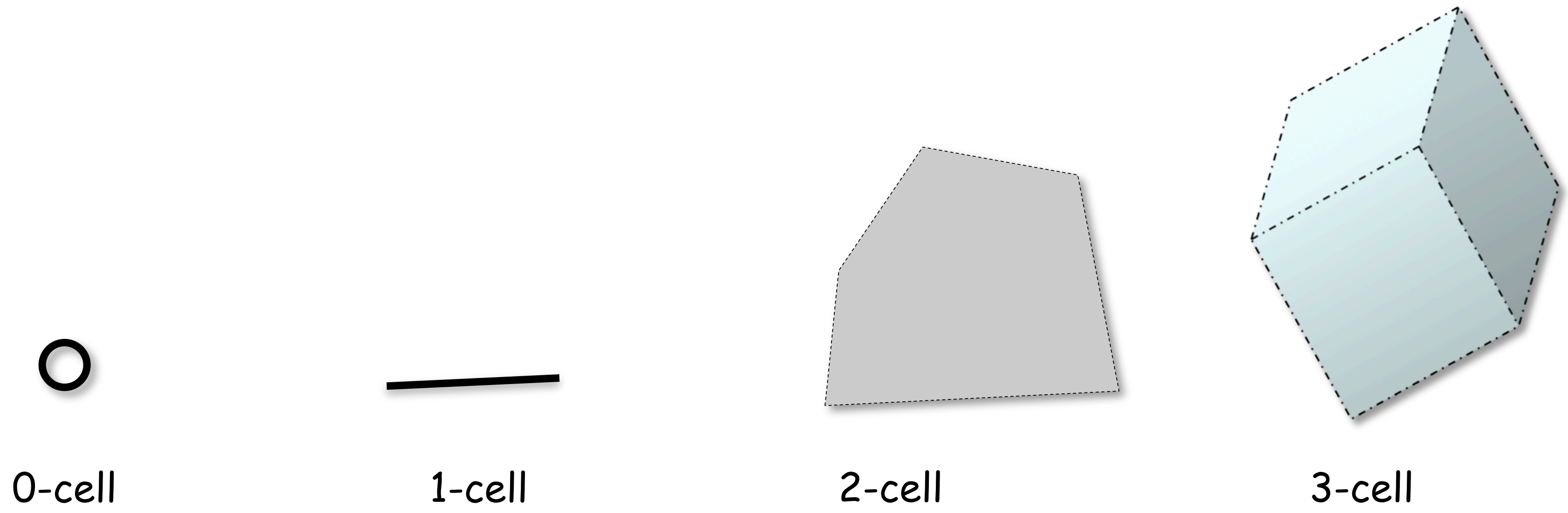
- Language dedicated to the simulation of $(DS)^2$
- Declarative (declarative simulation *vs* procedural)
- Abstract rewriting of complex spatial structures:
  - Data structure = topological collections

    sequence, generalized array, (multi-)set, arbitrary graph, Delaunay triangulation, g-map, …, <u>cell complexes</u>
  - Control structure = transformation
    - two powerful languages to specify sub-collections (elements in interaction)
    - Various rule application strategies: maximal parallel, asynchronuous, stochastic, Gillespie-like, …

Topological collection: representing the underlying space

Representation of space and structure

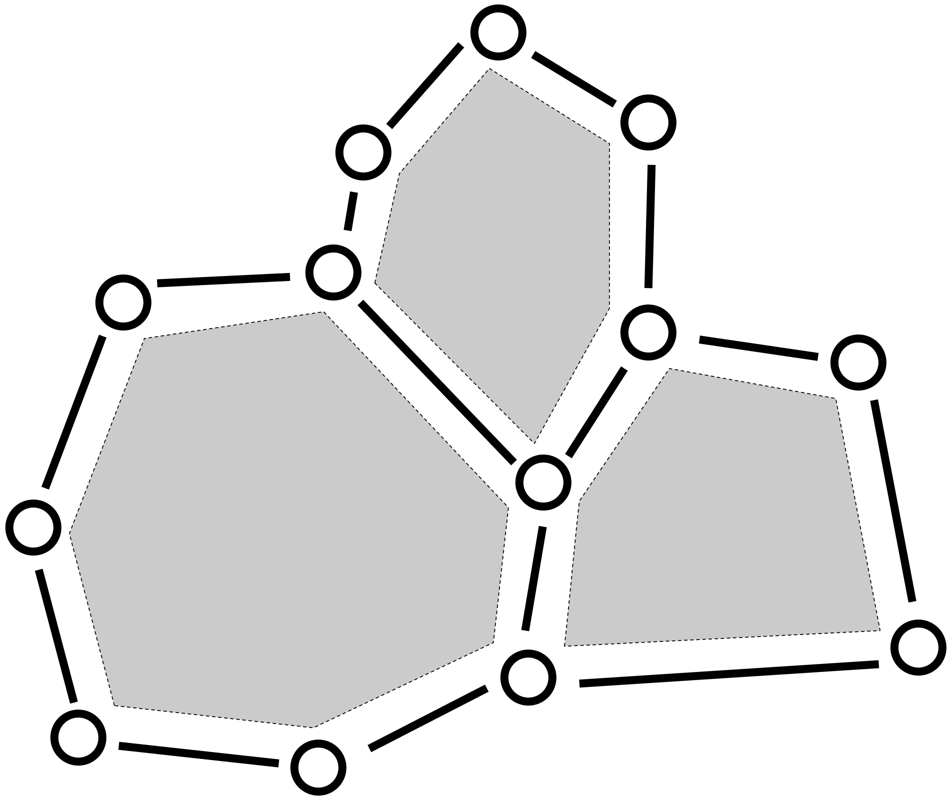
- Structure:
  - Collection of *topological cells*



Topological collection: representing the underlying space

Representation of space and structure

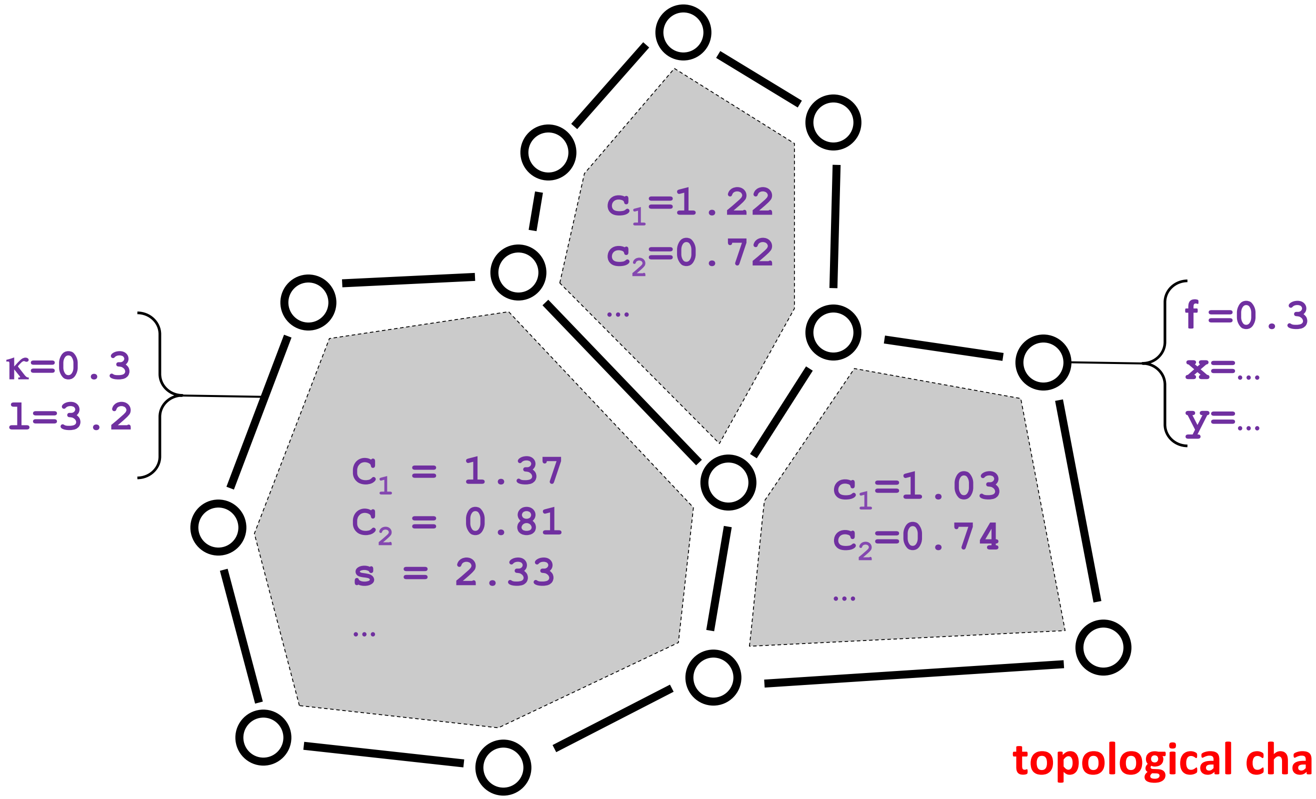
- Structure:
  - Collection of topological cells
  - *Incidence relationships*



Topological collection: a data-field over topological cells

Representation of space and structure

- Structure:
  - Collection of topological cells
  - Incidence relationship
- Data : *associating values* with topological cells ≈ field in physics



topological chain/cochain

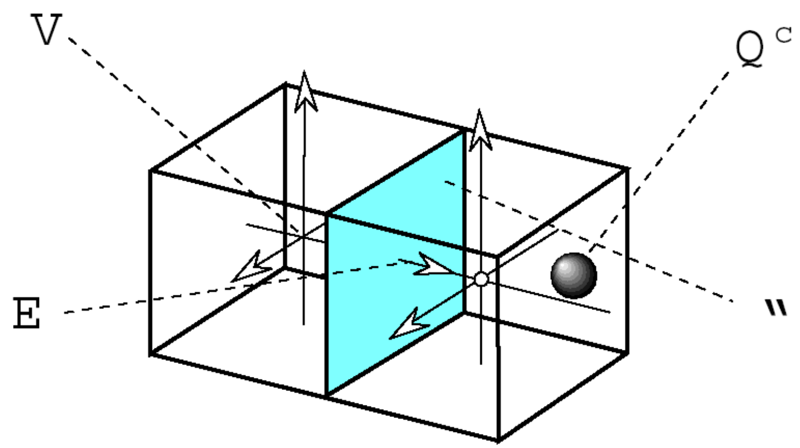
Higher dimensional objects for complex simulations

Example of electrostatic Gauss law [**Tonti** 74]

- Electric charge content ρ : dimension 3
- Electric flux $\boldsymbol{\Phi}$: dimension 2
- Law available on a arbitrary complex domain

$$\phi = \oiint w \cdot dS = \frac{Q^c}{\varepsilon_0} = \iiint_{(V)} \frac{\rho}{\varepsilon_0} d\tau$$

electric field in space:
- V: electric potential (dim 0)
- E: voltage (dim 1)
- w: electric flux (dim 2)
- Qc: electric charge (dim 3)



A Direct Discrete Formulation of Field Laws: The Cell Method

Topological rewriting = transformation

$1 + 2 \rightarrow \ldots$ (arithmetic) term rewriting

arithmetic operation

$a \,.\, b \rightarrow \ldots$ string rewriting (~ L systems)

string concatenation

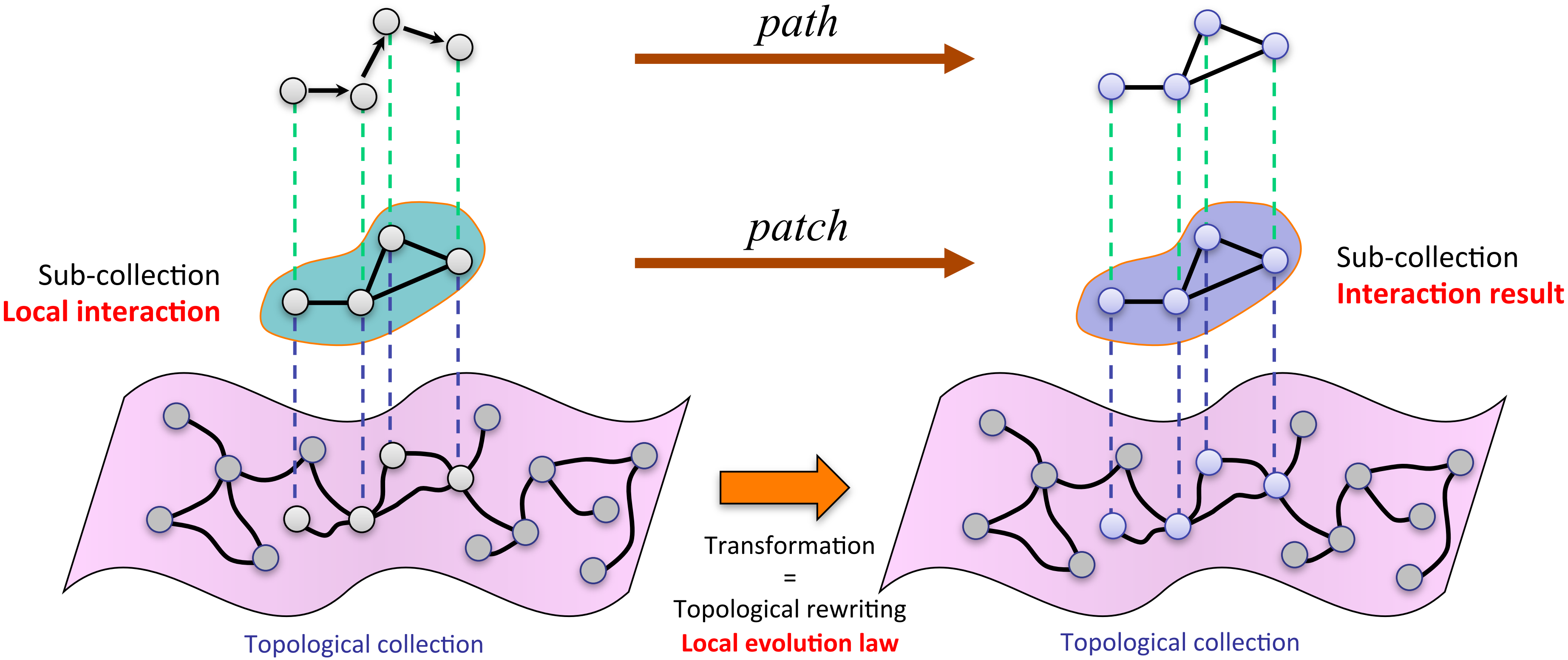
$2H + O \rightarrow H_2O$ multiset rewriting (~ chemistry)

multiset concatenation (= the chemical soup)

$v_1.\sigma_1 + v_2.\sigma_2 \rightarrow \ldots$ topological rewriting (MGS)

gluing cell in a cell complex

Transformation



Pattern matching : specifying a sub-collection of elements in interaction

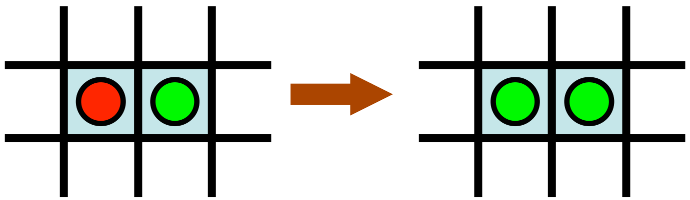
- *Path transformation* (path = sequence of neighbor elements)
  - Concise but limited expressiveness
- *Patch transformation* (arbitrary shape)
  - Longer but higher expressiveness

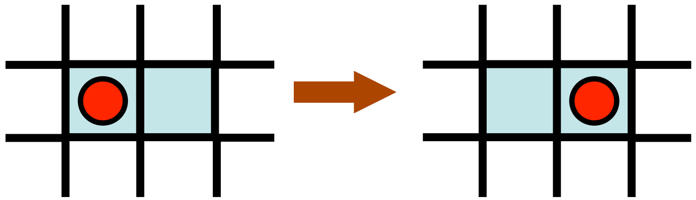
Example: Diffusion Limited Aggregation (DLA)

- Diffusion: some particles are randomly diffusing; others are fixed
- Aggregation: if a mobile particle meets a fixed one, it stays fixed

```
trans dla = {
      `mobile , `fixed  => `fixed, `fixed ;
      `mobile , <undef> => <undef>, `mobile
}
```

NEIGHBOR OF





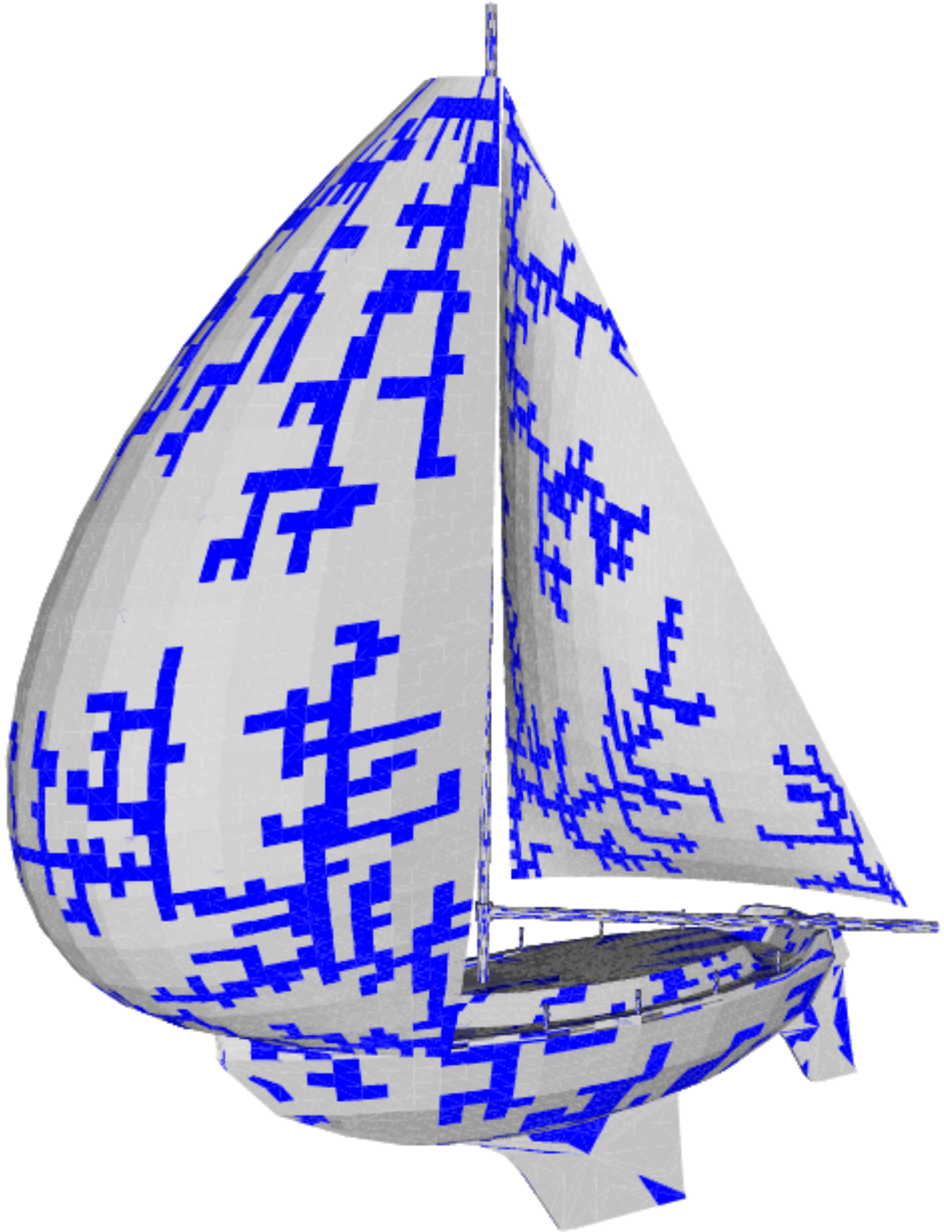
Example: Diffusion Limited Aggregation (DLA)

- Diffusion: some particles are randomly diffusing; others are fixed
- Aggregation: if a mobile particle meets a fixed one, it stays fixed

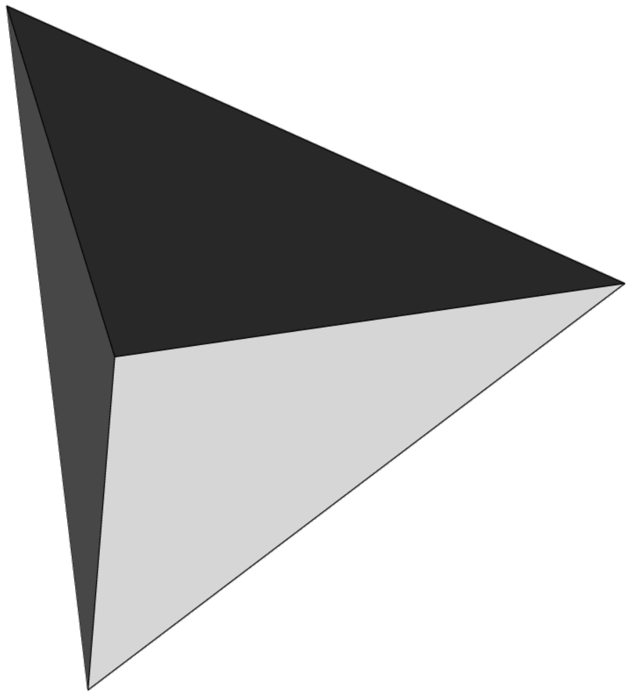
```
trans dla = {
    `mobile , `fixed   => `fixed, `fixed ;
    `mobile , <undef> => <undef>, `mobile
}
```

this transformation is an abstract process that can be applied to any kind of space

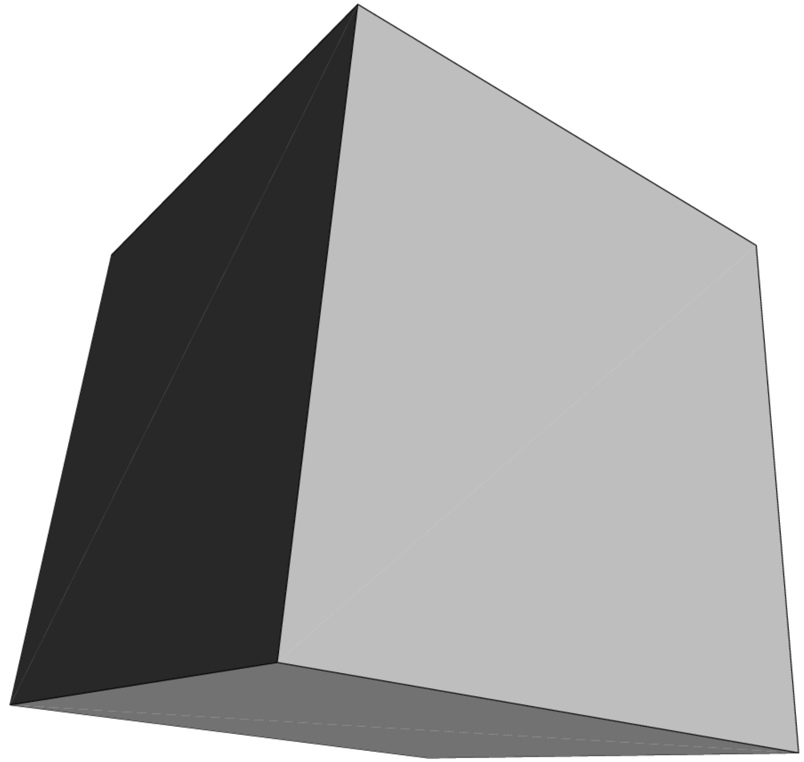
Polytypisme



Fractal construction by carving



Sierpinsky sponge (4 steps)



Menger sponge (2 steps)

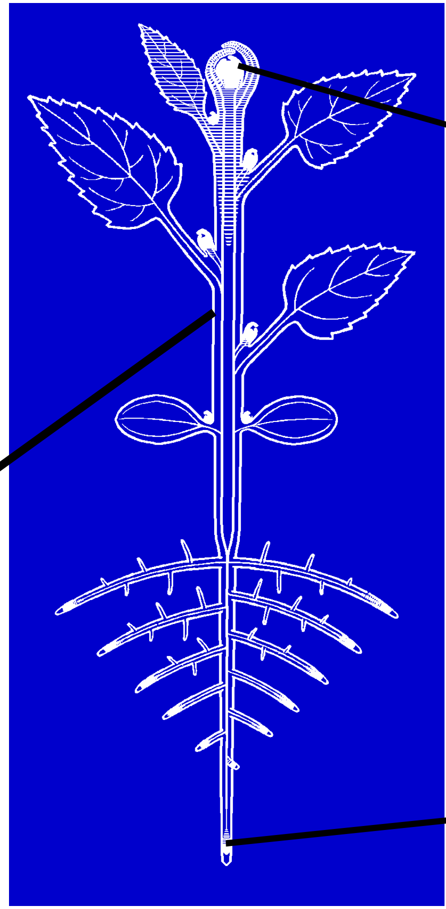
The Growth of a Meristem

[PNAS 103(5), 1627-1632, 2006]

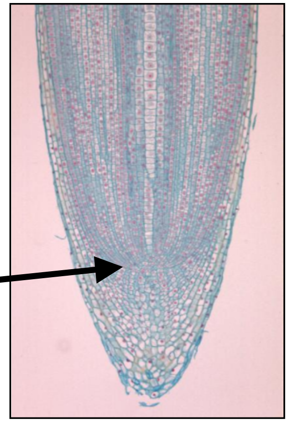
Pierre Barbier de Reuille
Mikaël Lucas
Jan Traas
Christophe Godin
CIRAD/INRA/INRIA

Organs positionning at the shoot apex

Cambium

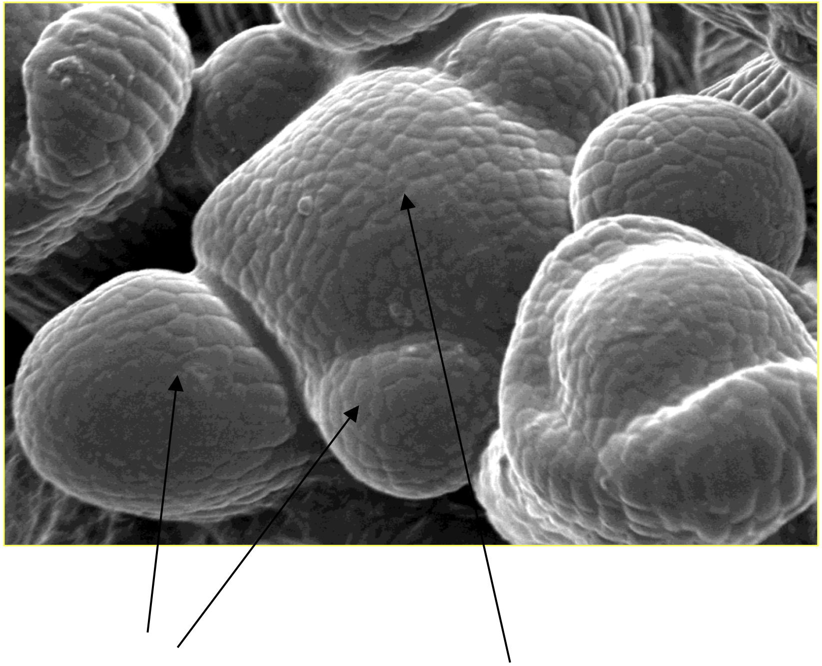


Shoot apical meristem



Root apical meristem

A shoot apical meristem



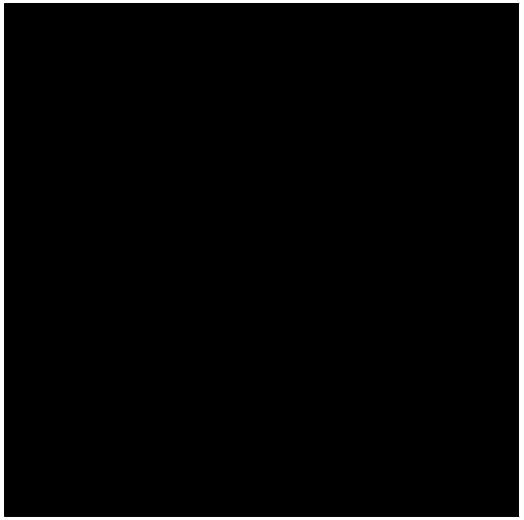
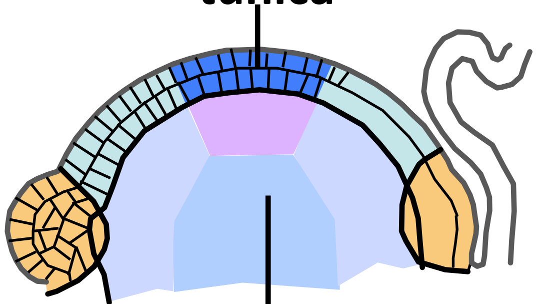


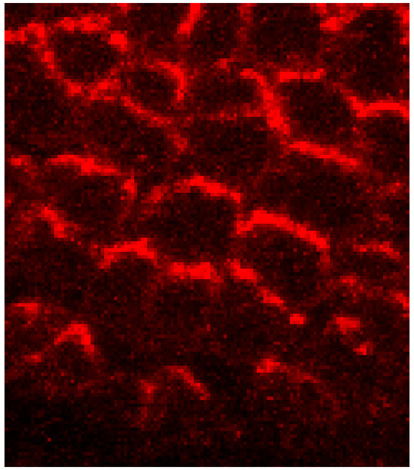
Image sequence showing cell division patterns via membrane-bound PIN1, in Shoot Apical Meristem (SAM), nearby floral meristems, and the boundaries between them (M. Heisler). http://computableplant.ics.uci.edu/ (E. Mjolness)



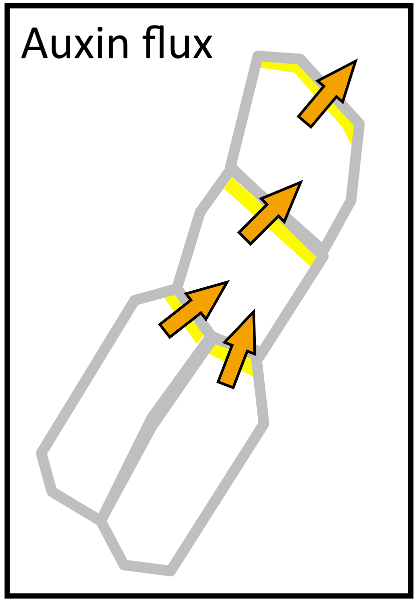
Active transport of auxine

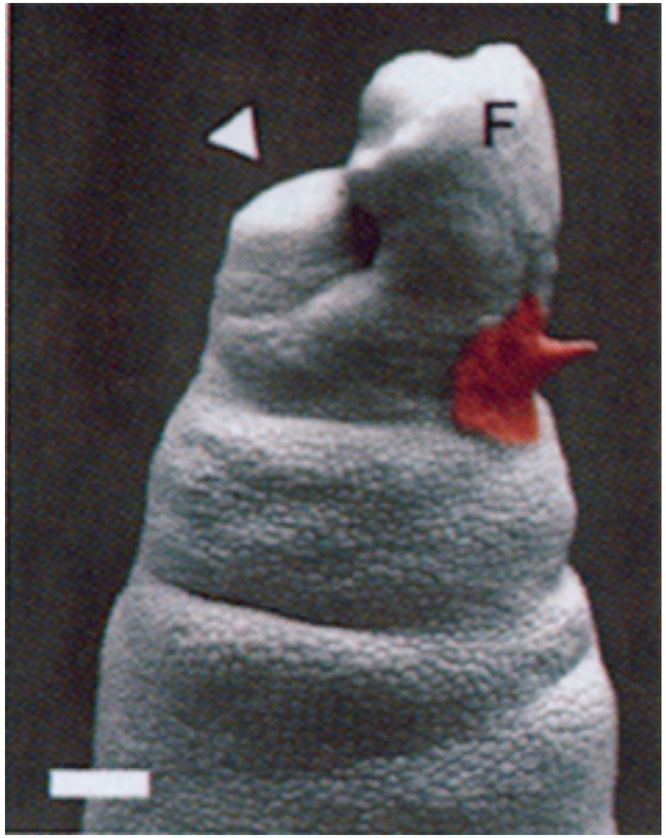
wild type

pin-1 mutant



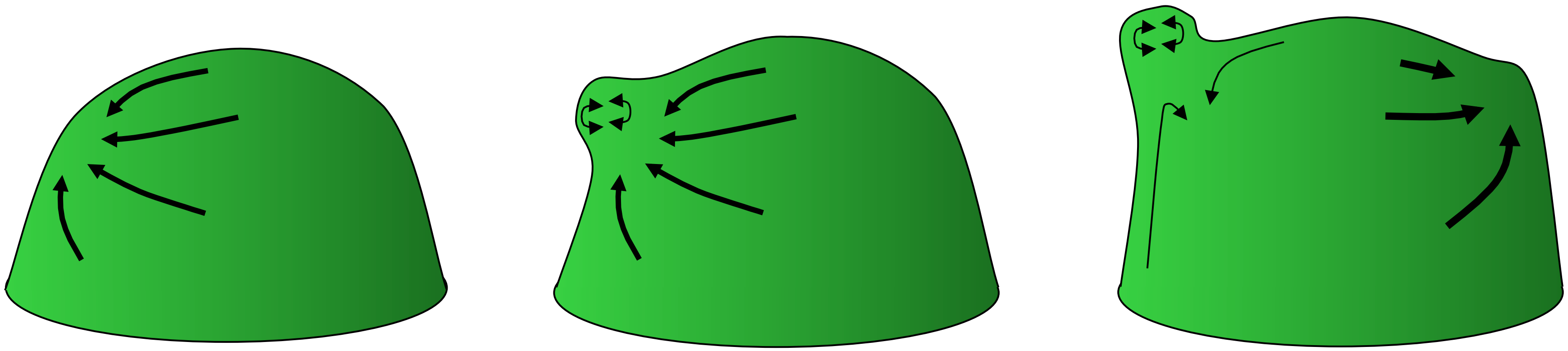
Immunolabelling of PIN-FORMED1 protein



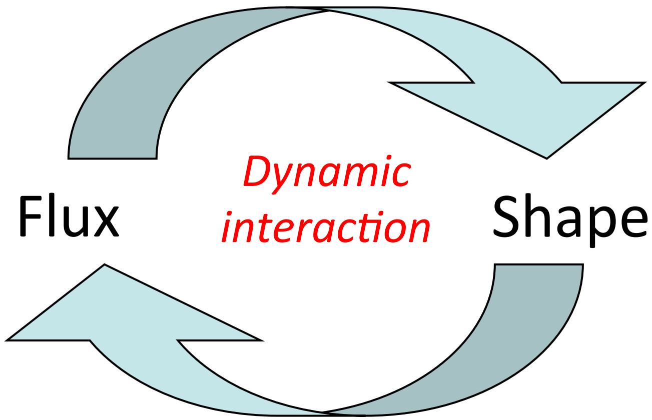


high concentration of auxine induces organ initiation

(DS)^2



flux… changes form… which changes flux…



Model

- ***Cell internal state and processes***
  capacity of division, spring relaxed length, primordium/center,
  concentration of auxin (inhibitor), saturation, auxin degradation / evacuation, promotion to primordium, "pump magnetism"

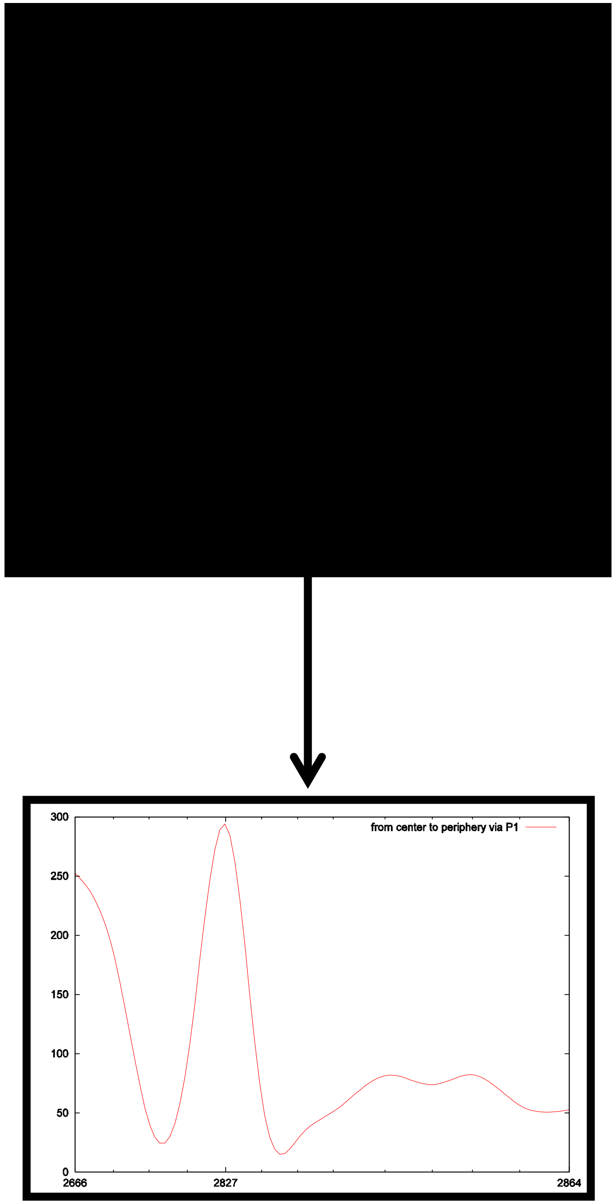
- ***Movement*** (due to cell growth)

- ***Growth***: increase of spring relaxed length

- ***Division***: when size > threshold

- ***Cell interaction***
  Passive diffusion of auxin, active pumping of auxin

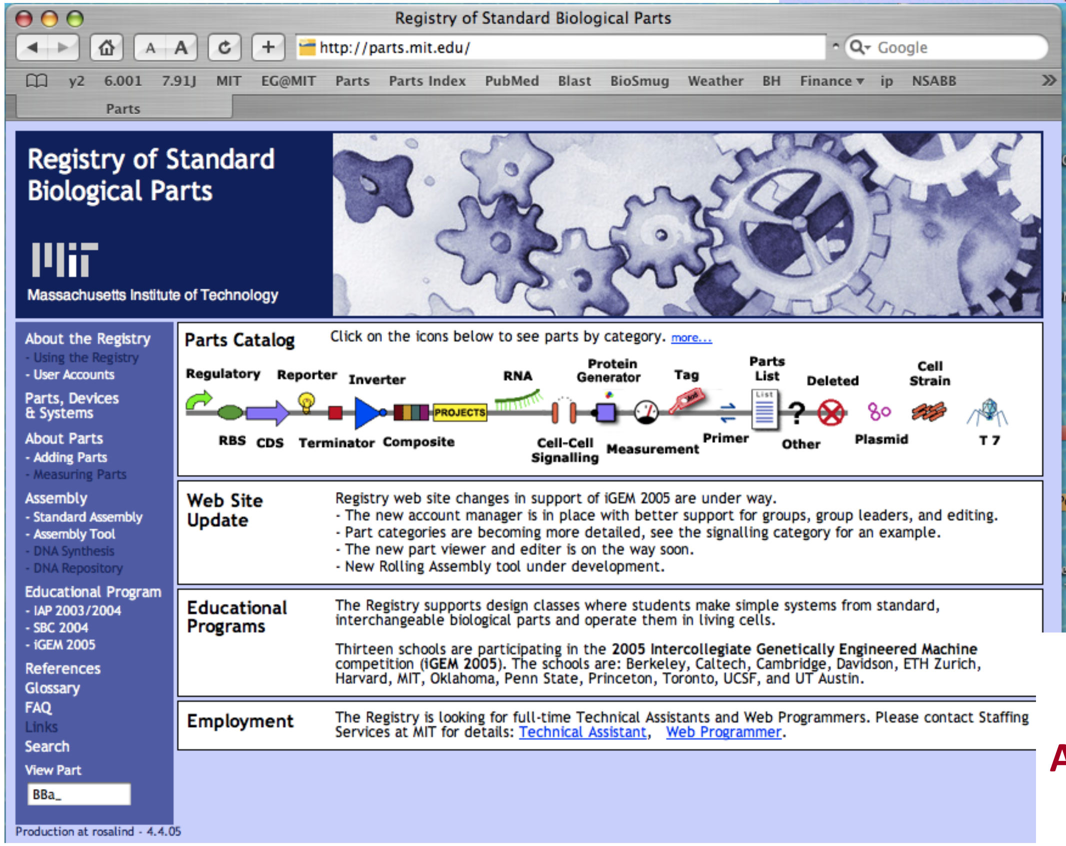
```
trans Auxin = {
        x, y / pump(x,y)
        → {x.auxin -= δ}, {y.auxin += δ}
}
```



Auxin level

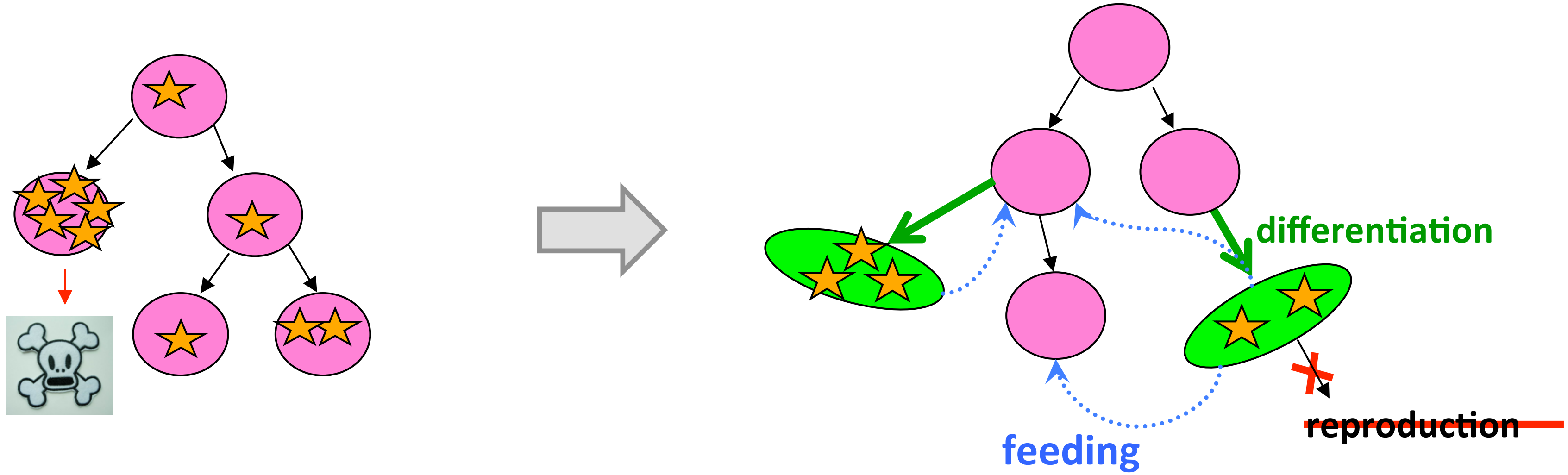
P. Barbier de Reuille

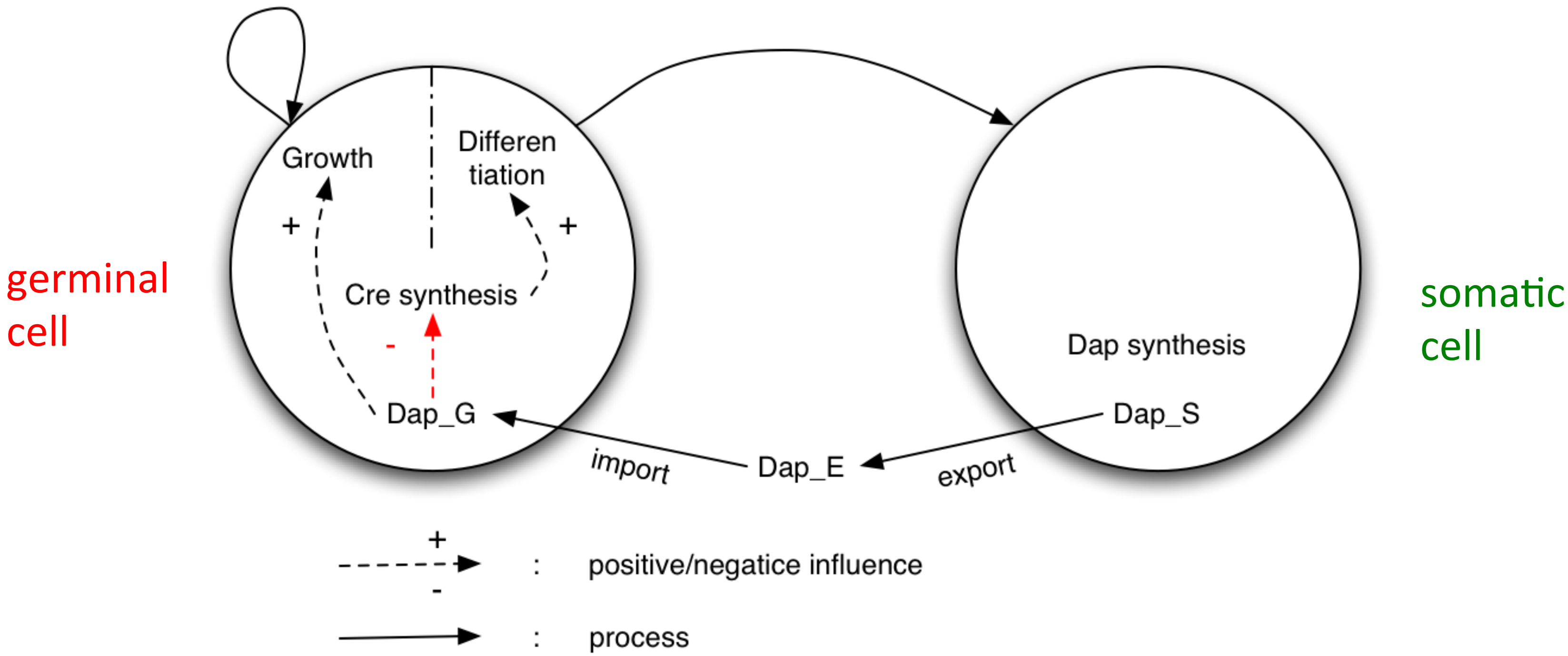
A Synthetic « Multicellular Bacterium »



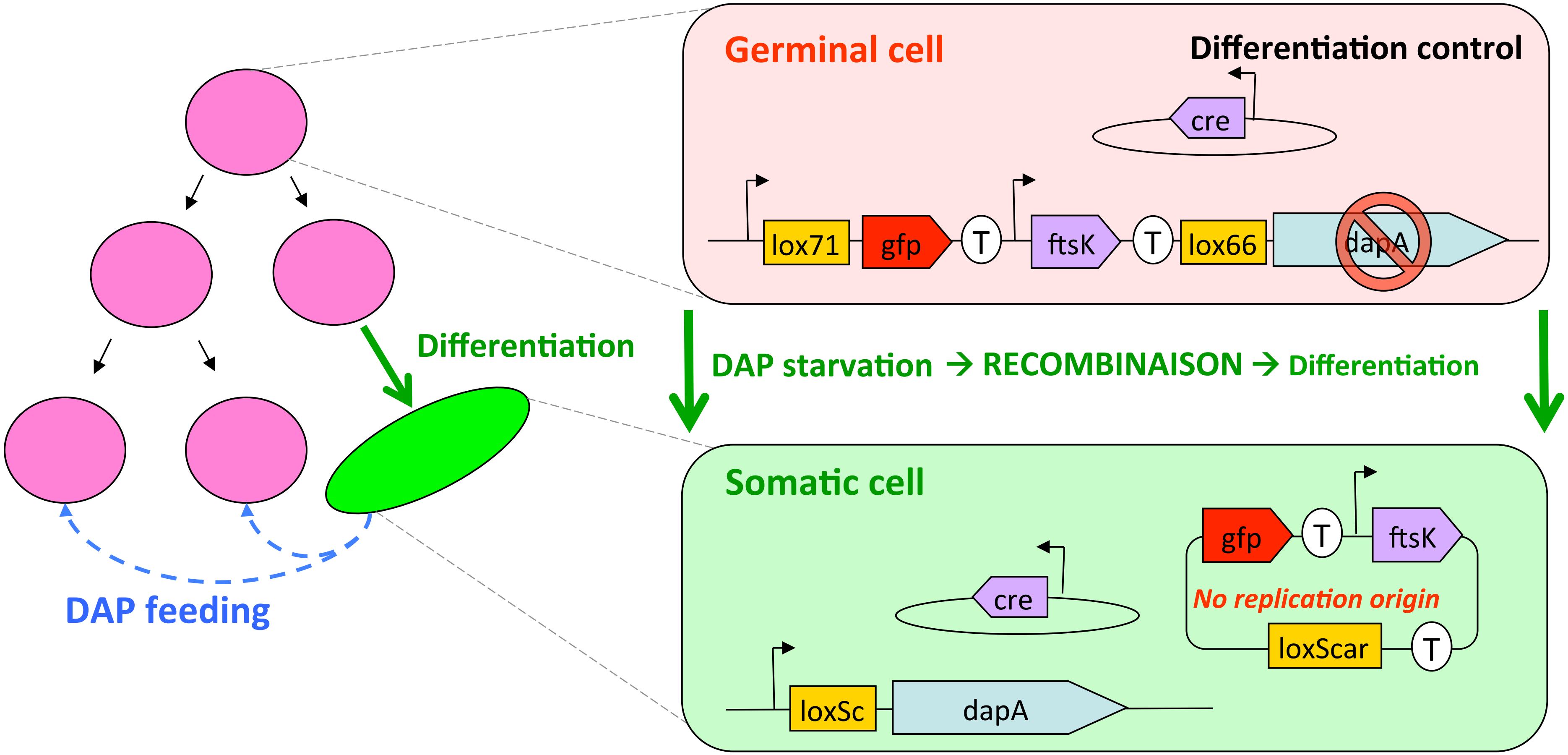
David Bikard, Thomas Landrain, David Puyraimond, Eimad Shotar, Gilles Vieira, Aurélien Rizk, David Guegan, Nicolas Chiaruttini, Thomas Clozel, Thomas Landrain

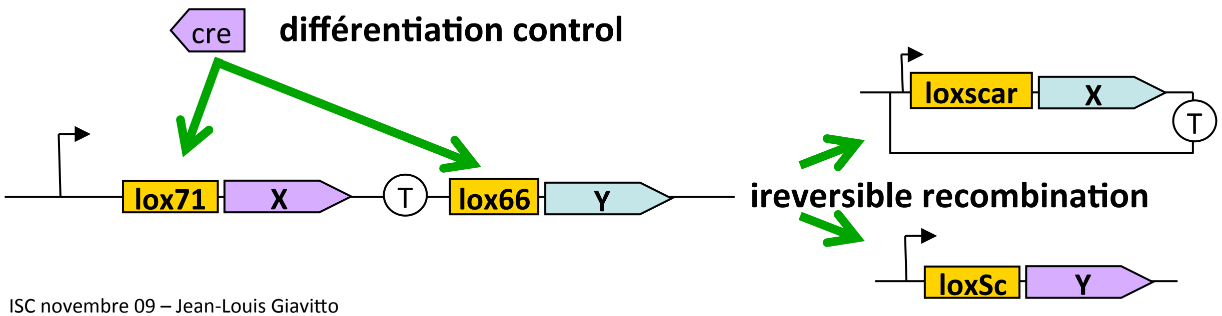
The Paris iGEM project: a « multicellular bacteria » to decouple growth and transgene expression

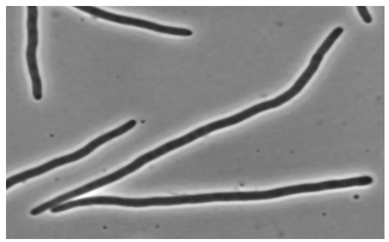




Implementation using BioBricks



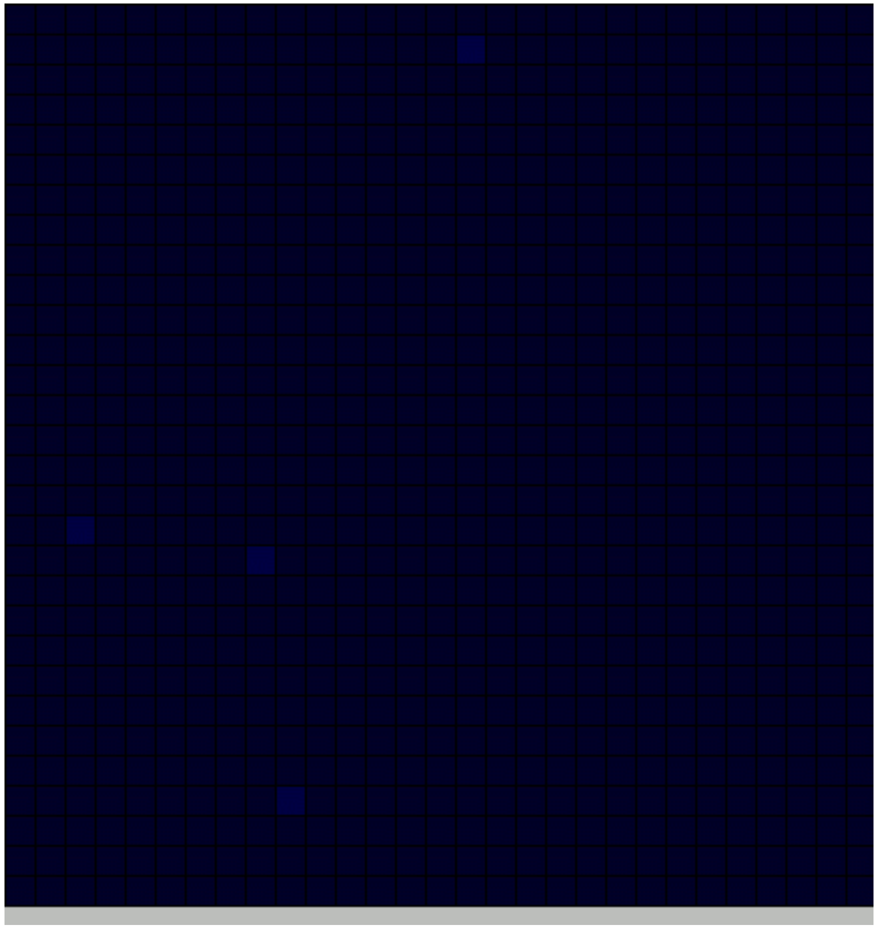




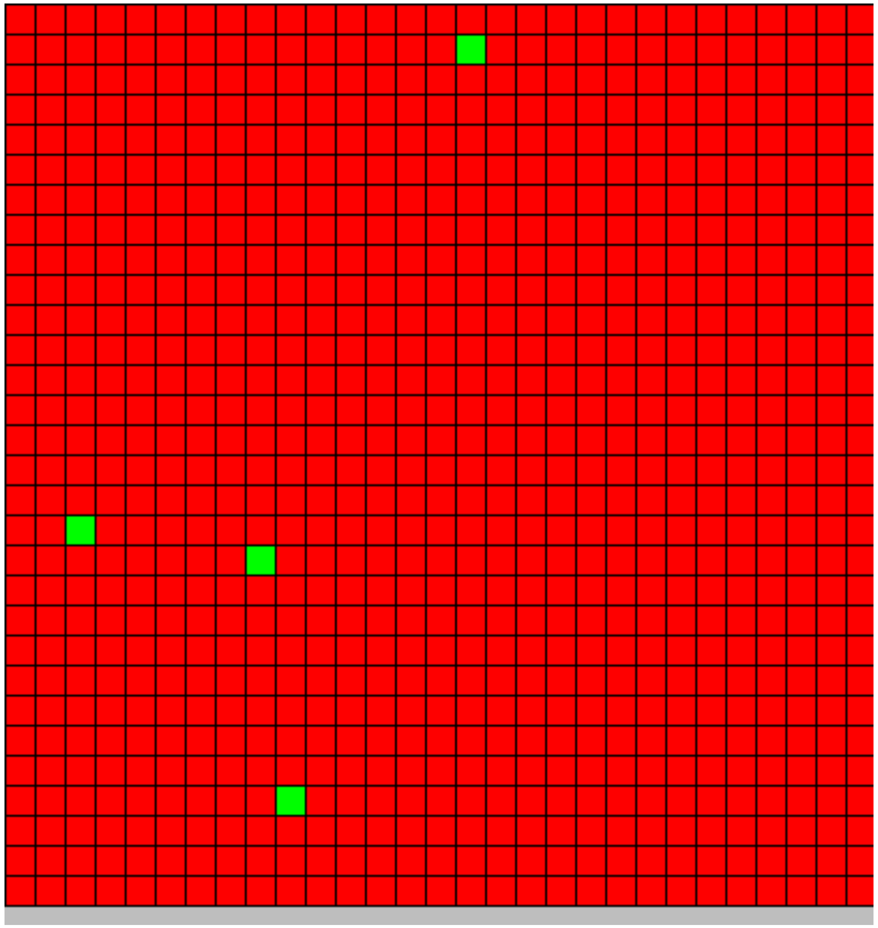
ftsK
needed for
cellular
division

Proof of Concept: Simulation to answer 4 questions

- **How does differentiation induces feeding?** (proof of concept)
  cellular automaton (in MGS)



diffusion of DAP



somatic and germ cell

Proof of Concept: Simulation to answer 4 questions

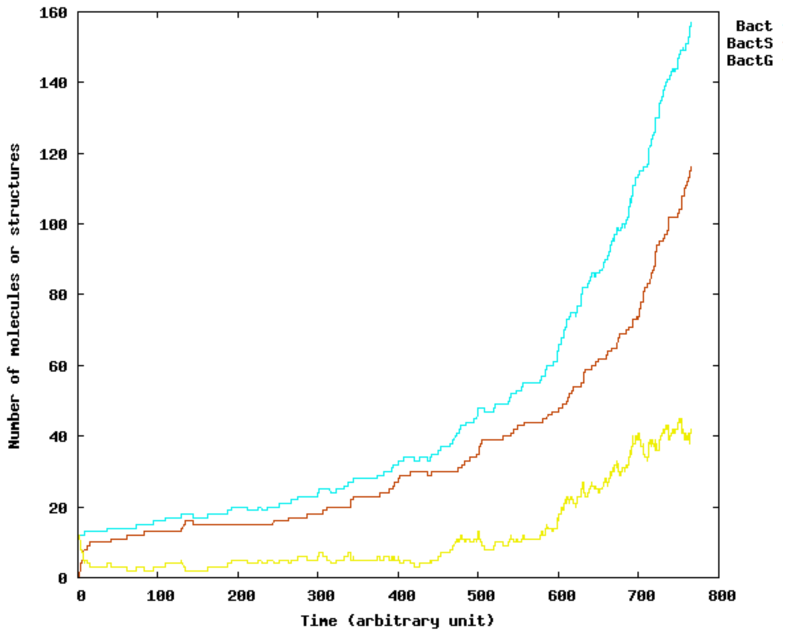
- **How does differentiation induces feeding?** (proof of concept)
  cellular automaton (in MGS)
- **How do spatial organization and distribution evolve?**
  agents based system (in MGS)

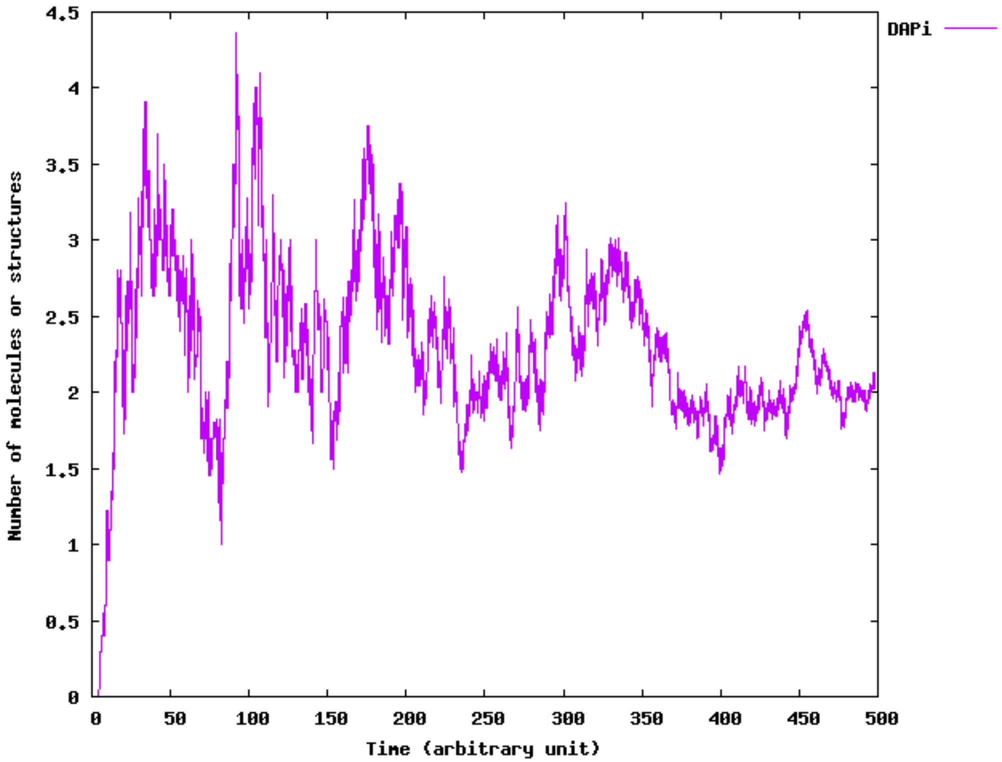
Proof of Concept: Simulation to answer 4 questions

- **How does differentiation induces feeding?** (proof of concept)
  cellular automaton (in MGS)
- **How do spatial organization and distribution evolve?**
  agents based system (in MGS)
- **How robust and tunable is the model?**
  ODE kinetics (matlab)

Proof of Concept: Simulation to answer 4 questions

- **How does differentiation induces feeding?** (proof of concept)
  cellular automaton (in MGS)
- **How do spatial organization and distribution evolve?**
  agents based system (in MGS)
- **How robust and tunable is the model?**
  ODE kinetics
- **How sensitive is the system to noise?**
  Gillespie based simulation (in MGS)





MGS drawbacks and successes

Success

- Polytypisme is good
- Patterns/rules are expressive and usually concise
- Clean semantics

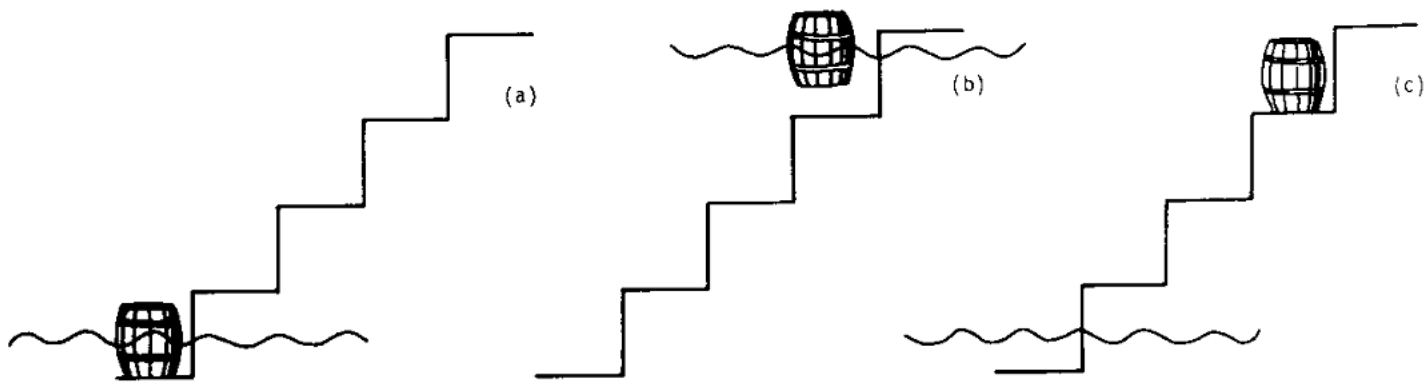
Shortcommings

- Rules may be heavy (e.g. 100 variables for the fractal sponge)
  graphical drawing of rules
  look for better notations (e.g. path pattern)
- Efficiency
  well…
- Implicit methods (solvers) are hairy
  use explicit ones

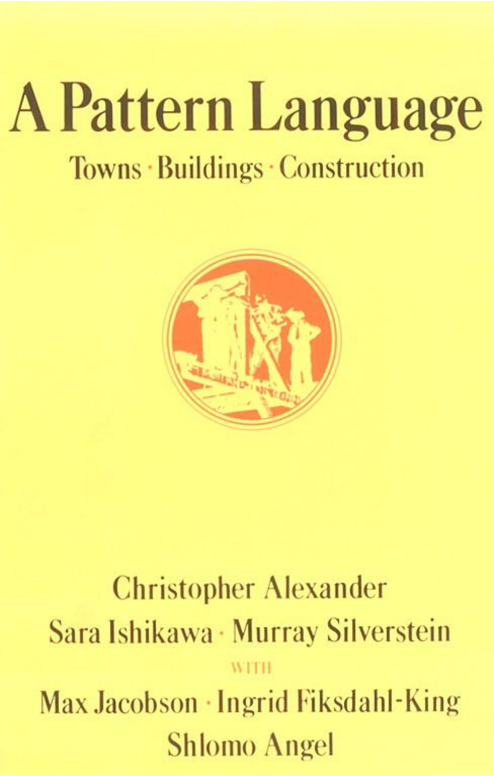
The need of design pattern

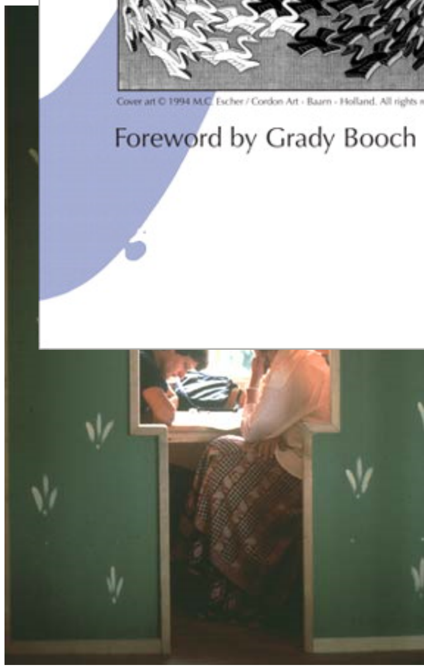
Multiple examples during the workshop

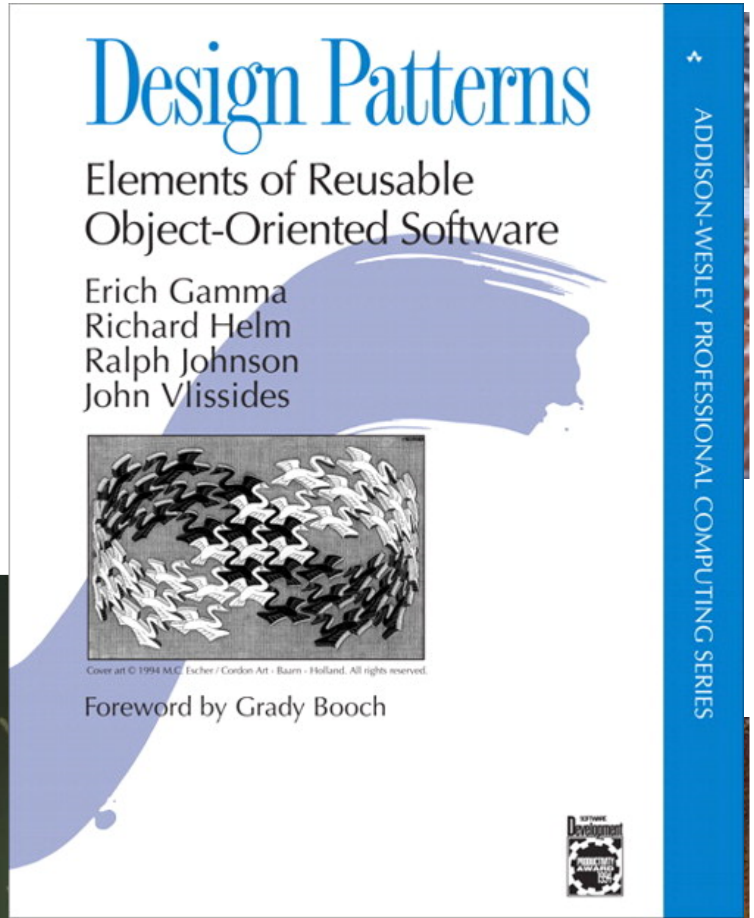
- Cassette, module, function, device, mechanism, gradient, amplification, diffusion, transport, diffusion-reaction, polarization...
  → abstracting biological processes
- Universal Mechanisms of Animal Development
  (basic machinery of development is conserved amongst species, homologous proteins, etc.)
- "Biochemical specification" vs. "causal explication"



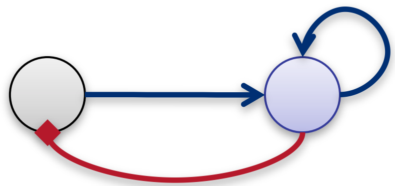
Design patterns







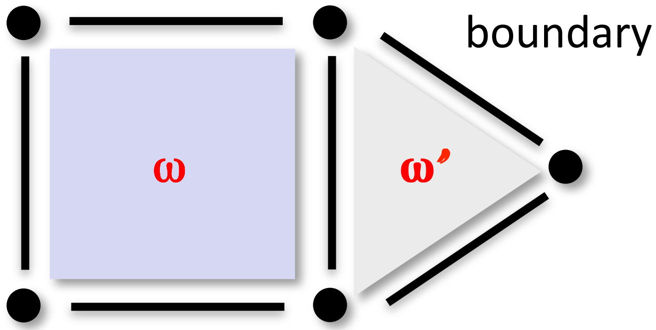
An analogy

« sort »	**« Function »**	« limb construction »
bubble-sort	**algorithms**	• establish positional information PI • differentiate cell wrt. PI • …
iterate over elements	**design patterns**	space-dependant activation of genes
`for (i = 0; i < n; i++)` …	**implementation**	

Purposes:
- **Pedagogical**
- **Heuristic :**
  - **auto-stabilizing systems**
  - **resources consumptions (memory, time, energy)**
  - **…**
- **Technical: compose and reuse models: towards an algebraic (relational) approach to biological processes**

Example: (some special) Transformations as Cochains

- The Boundary Operator ∂
  - Starting point of the elaboration of a discrete diff. calculus
  - *Transport* of data *from cells to their faces*



  - Cochains notation

    The boundary operator is a cochain

$$\partial = \sum_{\sigma \in \mathcal{K}} \partial_\sigma . \sigma \quad \text{with} \quad \forall \sigma \in \mathcal{K},\ \partial_\sigma(g) = \sum_{\tau < \sigma} o_{\sigma\tau}(g).\tau$$

  - MGS notation

```
trans Boundary      = {
  x => CofacesFold(fun y acc -> o_{^y^x}(y) +_G acc,0_G,^x) }
```

Transformation as Cochains

- Derivative Operator $\mathbf{d}$
  - Defined w.r.t. the discrete Stockes' Theorem

$$[\mathbf{d}T, c] = [T, \partial c]$$

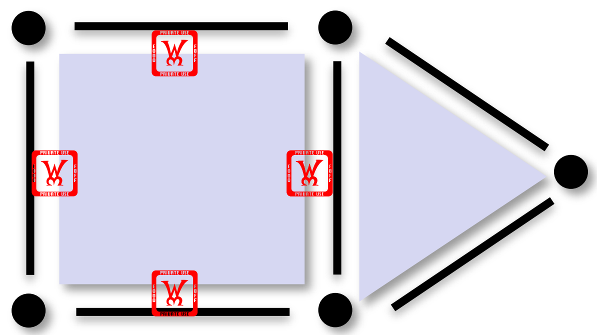
  - Cochains Notation

    One can show that the derivative verifies

$$\mathbf{d} = \sum_{\tau \in \mathcal{K}} \mathbf{d}_\tau.\tau \quad \text{with} \quad \forall \tau \in \mathcal{K},\ \mathbf{d}_\tau(f) = \sum_{\tau < \sigma} (f \circ o_{\sigma\tau}).\sigma$$

  - MGS Notation

    We directly use the Stockes' Theorem



```
let Derivative  T = fun c -> T (Boundary c)
```

Transformation as Cochains

- Illustrative example : the Laplacian Operator Δ

  – The Laplacian in terms of $\star$ and **d** [Desbrun *et al*., 2006]

$$\Delta = \delta \mathbf{d} + \mathbf{d} \delta \qquad \text{where} \quad \delta = (-1)^{n(k-1)+1} \star \mathbf{d} \star$$

  – MGS notation

    Big assumption: the Hodge star $\star$ is replace by the co-derivative $\mathbf{d}^{co}$
    (= uniform geometry)

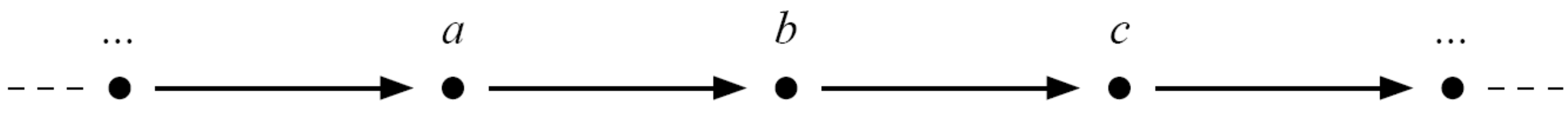
```
let Laplacian T =
  let Sg T' c' =
    T'(trans { x => -1**((dim c')*((dim ^x)-1)+1)*x }(c'))
  in
    fun c ->   Derivative(Sg(Derivative^co(T)))(c)
             + Sg(Derivative^co(Derivative(T)))(c)
```

Transformation as Cochains

- Illustrative example : the Laplacian Operator

  – Corresponding Data Transport (case of dimension 1)

    - Dimension 1: $\Delta = \mathbf{d}^{co}\mathbf{d}$
    - Stockes' Theorem: équivalence with $\partial \circ \partial^{co}$

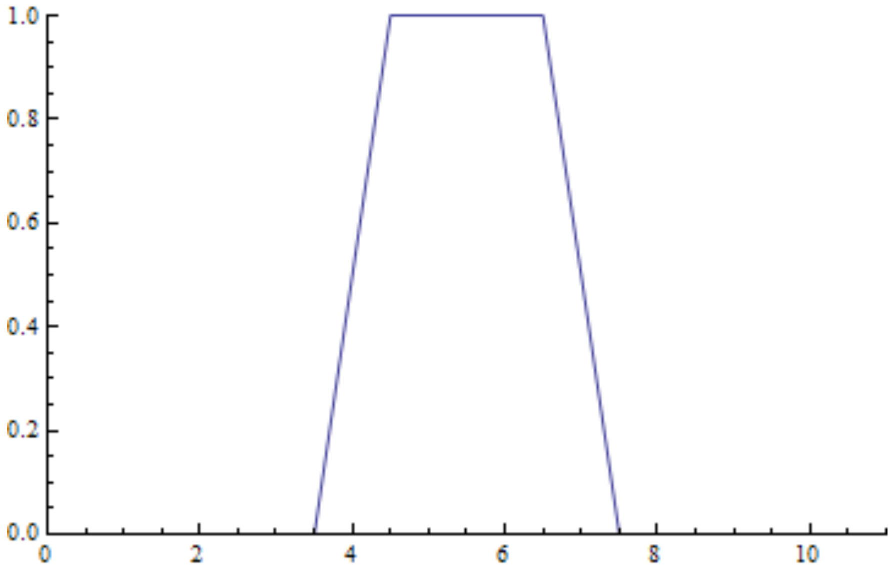


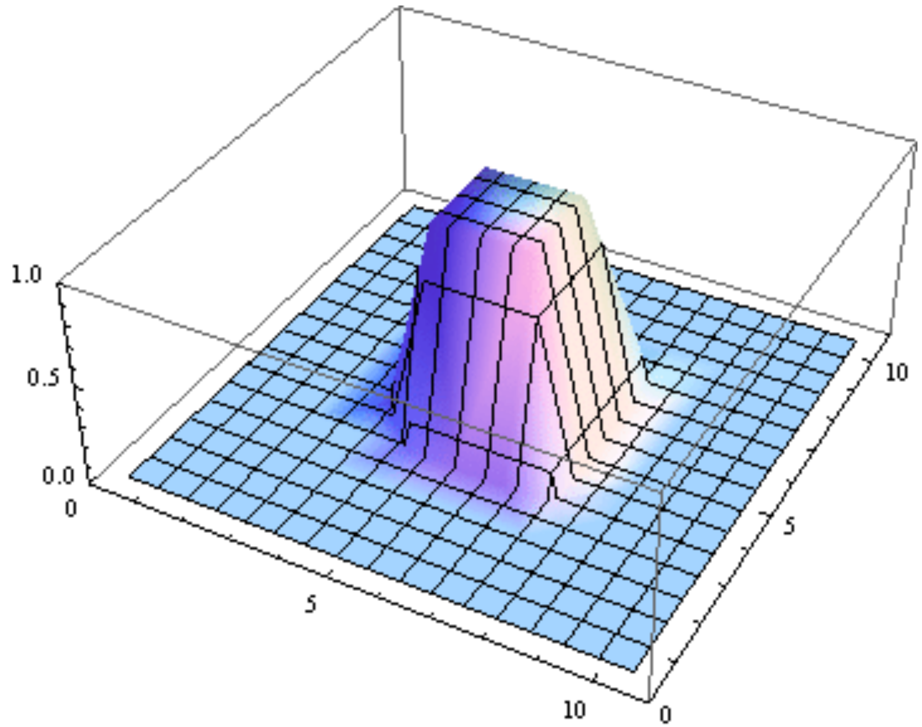
Transformation as Cochains

- Illustrative example : the Laplacian Operator Δ
  - Simulation of diffusion

$$\frac{\partial u}{\partial t} = D\Delta u$$

```
fun diffusion[D,orient](u) =
  u + D*Laplacian[orient=orient](Id)(u) ;;
```





Thanks

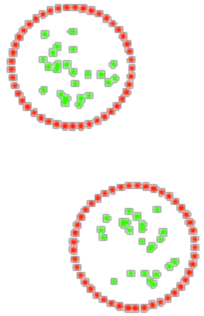
http://mgs.spatial-computing.org

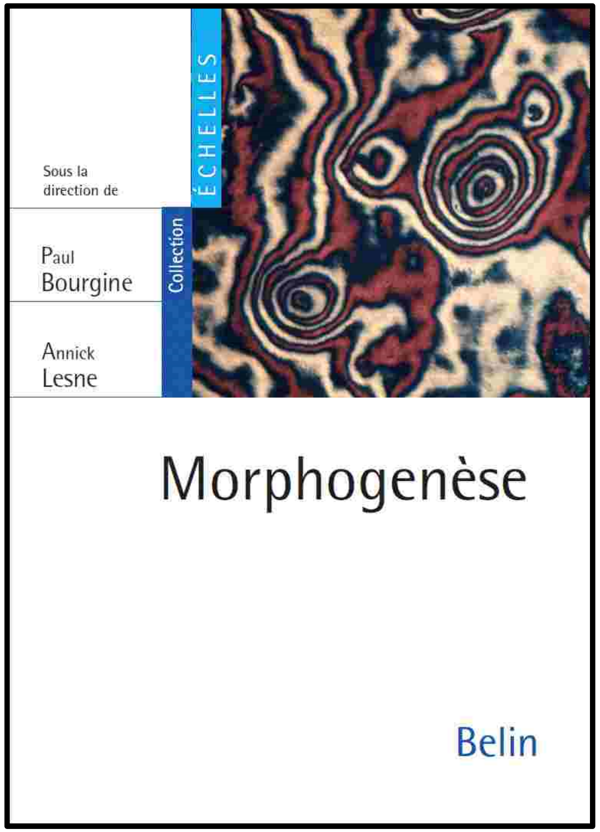
- **Antoine Spicher**
- **Olivier Michel**
- **PhD and other students**

  J. Cohen, P. Barbier de Reuille,

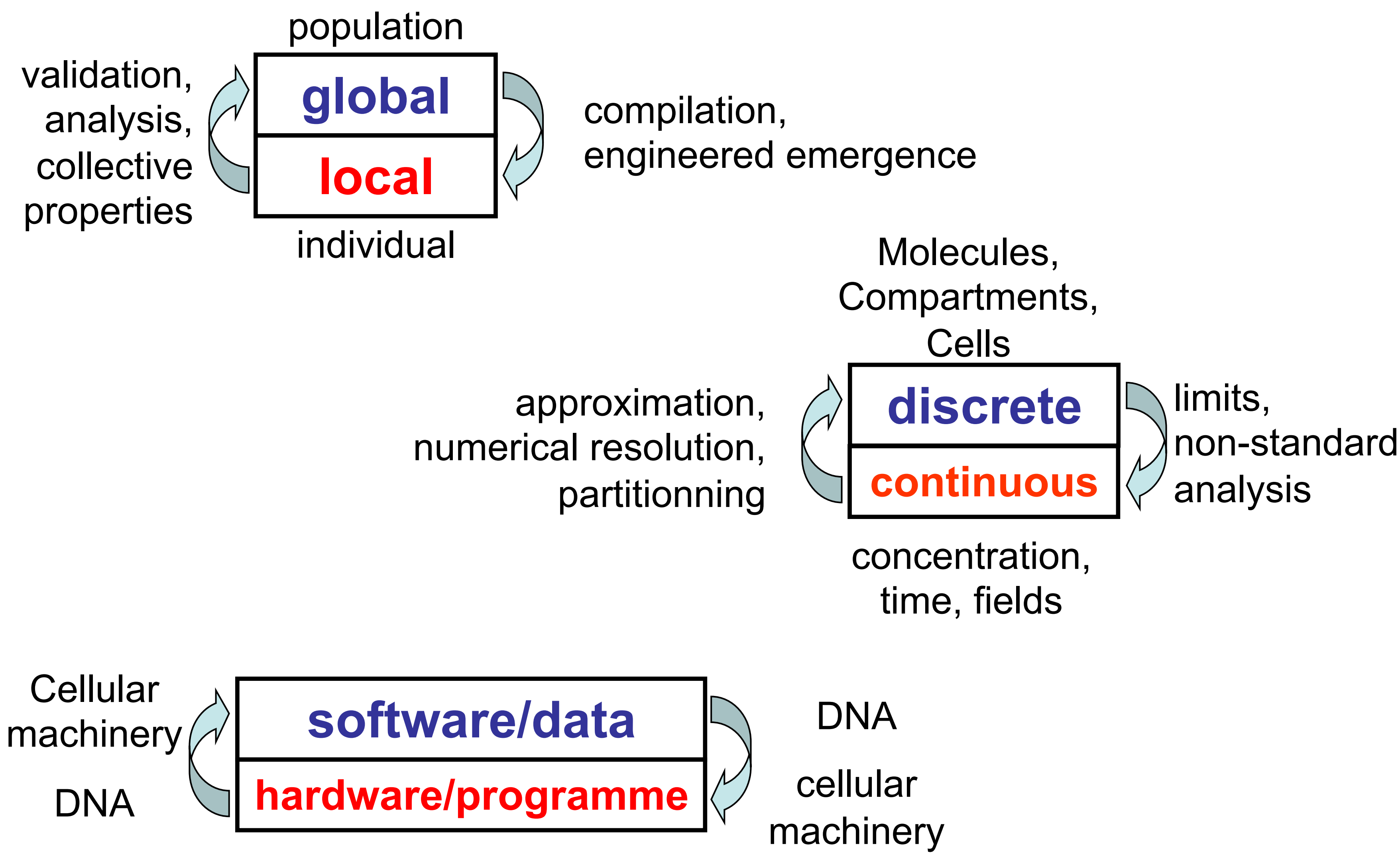
  E. Delsinne, V. Larue, F. Letierce, B. Calvez,

  F. Thonerieux, D. Boussié *and the others…*
- **Collaborations**
  - A. Lesne (IHES, stochastic simulation)
  - P. Prusinkiewicz (Calgary, declarative modeling)
  - P. Barbier de Reuille (meristeme model)
  - C. Godin (CIRAD, biological modeling)
  - H. Berry (LRI, stochastic simulation)
  - G. Malcolm (Liverpool, rewriting)
  - J.-P. Banâtre (IRISA, programming)
  - F. Delaplace (IBISC, synthetic biology)
  - P. Dittrich (Jena, chemical organization)
  - E. Goubault (CEA, topological formalization)
  - F. Gruau (U. PXI, language and hardware)
  - P. Liehnard (Poitier, CAD, Gmap and quasi-manifold)





Three challenges and some tools



(…) An invagination of a germ layer may be explained on a basis of a pressure difference between the two surfaces (or sides) or by cell movements, and so forth. This can be considered as an 'explanation' until we ask about the origin of pressure differences, or the mechanisms involved in cell movement, etc. However, questions of this kind become trivial when a larger process, rather than its individual components becomes the main problem. Suppose for a moment that each element in the succession A, a, C… can be explained separately, e.g. A as a swelling, a as a chemical reaction, etc. Interesting as they may be, these explanations are of subordinate importance when related to the main question: Why indeed is a regular (emphasis added) succession of these obviously quite different processes taking place at all? Most biological problems are of this kind and all of embryogenesis is just such a single problem. Here we require a peculiar or, maybe, original explanatory principle... A process may become accessible to explanation only insofar as one can succeed in substituting [understanding of] a purely phenomenological multiplicity and diversity of events [for understanding] of a less diverse and less arbitrarily created picture correctly reflecting reality. The main aim of such a construction would be as follows. The entire process should be accessible for analysis into a finite, not very large number of stages, each stage being represented as a monotonic function of some definite initial conditions and a single variable such as time, or distance, etc. If this cannot be realized, we consider a given set of events as scientifically inaccessible. On the other hand, even a partial success of such an enterprise is an obvious step forward."

(Gurwitsch, 1944)

cited by Beloussov in "Life of Alexander G. Gurwitsch and his relevant contribution to the theory of morphogenetic fields", Int. J. Dev. Biol. 41, 771-779 (1997)