The Self-Made Puzzle:

Integrating Self-Assembly and Pattern Formation Under Non-Random Genetic Regulation

René Doursat

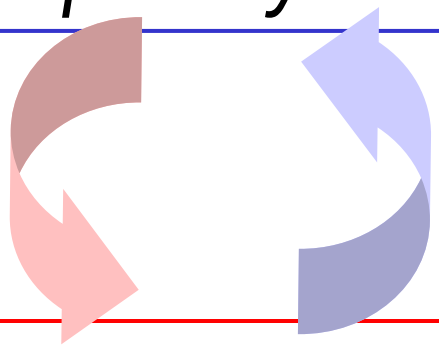
Institut des Systèmes Complexes, CREA
CNRS & Ecole Polytechnique
57-59, rue Lhomond, 75005 Paris, France

Designing Complexity

- Complex systems engineering

understanding natural complex systems

exporting:
- decentralization
- autonomy
- evolution



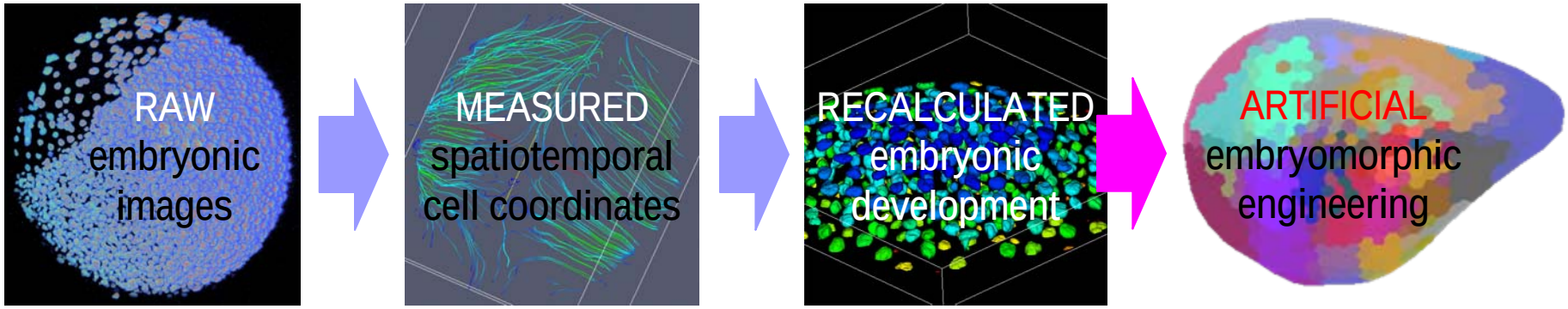
importing:
- modeling
- simulation

designing a new generation of artificial systems

Designing Complexity

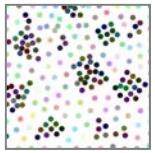
- **Toward a new discipline: "Embryomorphic Engineering"**
  - ✓ observing, modeling & transferring biological development
    - automating the **observation** and description of developing organisms with image processing, statistical and machine learning techniques
    - designing mathematical/computational **models** of embryonic growth
    - **implementing** biological development in engineering systems: distributed architectures as a prerequisite for evolutionary innovation

European projects "Embryomics" & "BioEmergences"

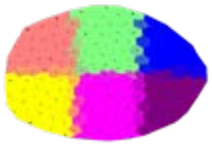


The Self-Made Puzzle

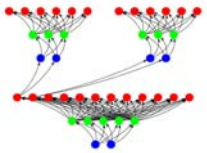
- Integrating *self-assembly* and *pattern formation* under non-random *genetic regulation*
  - ✓ self-assembly (SA)



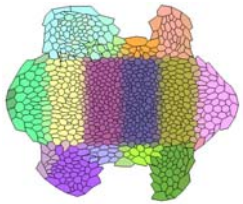
    - usually focuses on pre-existing components endowed with fixed shapes
    - . . . but cells *dynamically divide and differentiate* toward selective adhesion
  - ✓ pattern formation (PF)



    - generally orderly states of activity on top of continuous 2-D or 3-D substrate
    - . . . but gene expression patterning arises in *perpetually reshaping* organism
  - ✓ non-random genetic regulation (GRN)



    - both phenomena often thought stochastic: mixed components that randomly collide in SA; spots and stripes that pop up from instabilities in PF
    - . . . but cells are *pre-positioned* where they divide, and genetic identity domains are *highly regulated* in number and position



→ *integrate these 3 aspects in artificial "embryomorphic" systems*

The Self-Made Puzzle

1. Self-Assembly of Pre-Patterned Components

2. Pattern Formation in Pre-Assembled Media

3. Integrating Self-Assembly and Pattern Formation Under Genetic Regulation

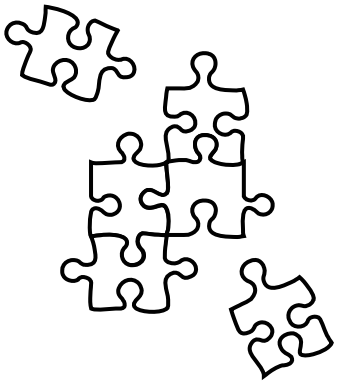
4. Toward Evolutionary Meta-Design

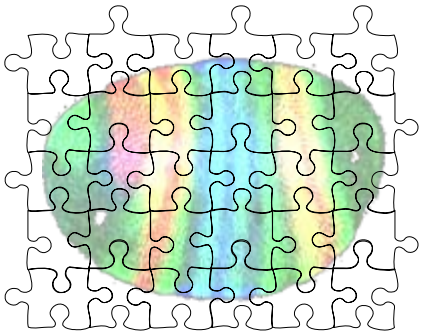
1. Self-Assembly

- The "jigsaw puzzle" metaphor of self-assembly
  - ✓ "piece" of the puzzle
    - an elementary component of the system—molecule, cell
  - ✓ "shape" of a piece
    - its binding affinities with other components—electric field, differential adhesion
  - ✓ "state" of the puzzle
    - a particular spatial arrangement of its pieces
  - ✓ "solutions" of the puzzle
    - energy minima, i.e., states where all pieces best satisfy each other's constraints

1. Self-Assembly

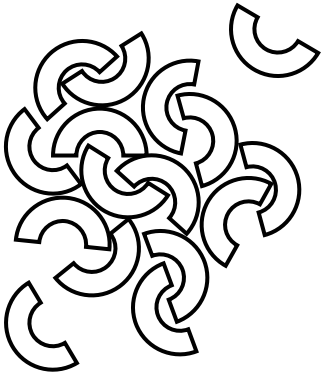
- Jigsaw puzzles
  - ✓ affinities
    - fit between components is all-or-none (inflexible)
    - system has a unique solution, the absolute energy minimum
  - ✓ component types
    - shape constraints are unique to each piece, by geometry and/or markings
    - compatibility with other pieces is a unique event
    - unique solution requires a long time to find

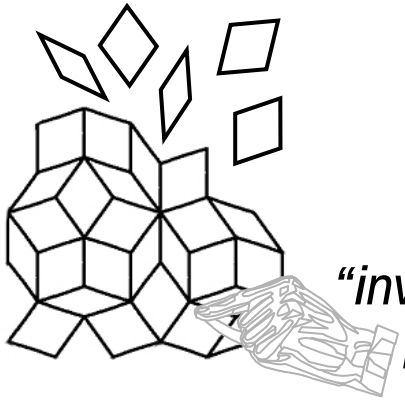


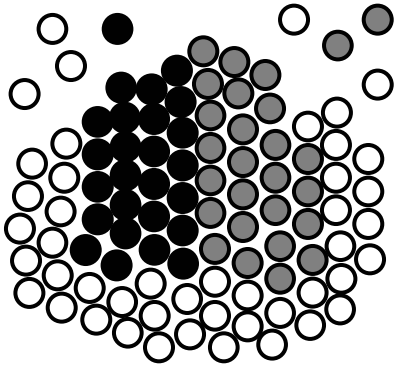


1. Self-Assembly

- Self-assembling systems (molecular or multicellular)
  - ✓ affinities
    - fit between components is approximate (flexible)
    - degrees of well-formedness, associated with degrees of energy costs
    - system has multiple “solutions” that are low-energy states
  - ✓ component types
    - few distinct components types, shared by multitude of clones
    - many equivalent states, invariant by permutation of components
    - much easier convergence toward one of many low-energy solutions







1. Self-Assembly

➢ *Molecular*-style self-assembly

- ✓ existence of components
  - molecules generally pre-exist in the solution before they self-assemble

- ✓ binding fate
  - molecules initially form a homogeneous mixture (the puzzle box)
  - molecules bind to each other through stochastic collisions
    (possibly with help from enzymes, but the original encounter remains stochastic)

- ✓ shape determination
  - molecules settle on a relatively fixed/passive geometrical shape
    (possibly after folding)
  - molecules admit only a limited amount of deformation when coming into contact with other molecules

1. Self-Assembly

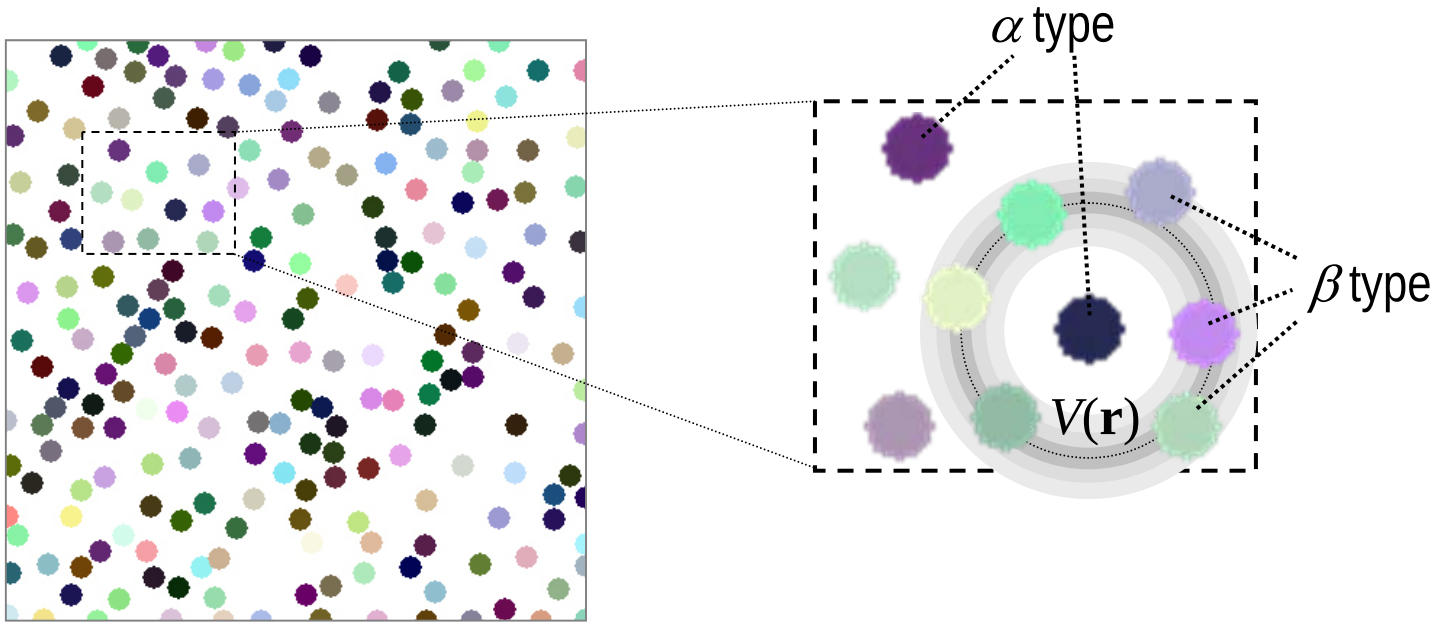
➢ *Multicellular*-style self-assembly

- ✓ existence of components
  - cells are dynamically created *during* self-assembly by division of other cells
    (of course, not *ex nihilo*, but by self-assembly of pre-existing molecules at lower level)
- ✓ binding fate
  - cells appear *on the spot*, again by cellular division
  - cells generally bind to their immediate *neigborhood*
    (possibly after targeted migration, which is also a highly nonrandom process)
- ✓ shape determination
  - cells dynamically and actively *change* their shape
  - cells differentiate under the influence of molecular signalling from other cells (e.g., induction)
  - cell vary their adhesion properties *depending on their neighbors*

1. Self-Assembly

➢ A simple model of swarm behavior

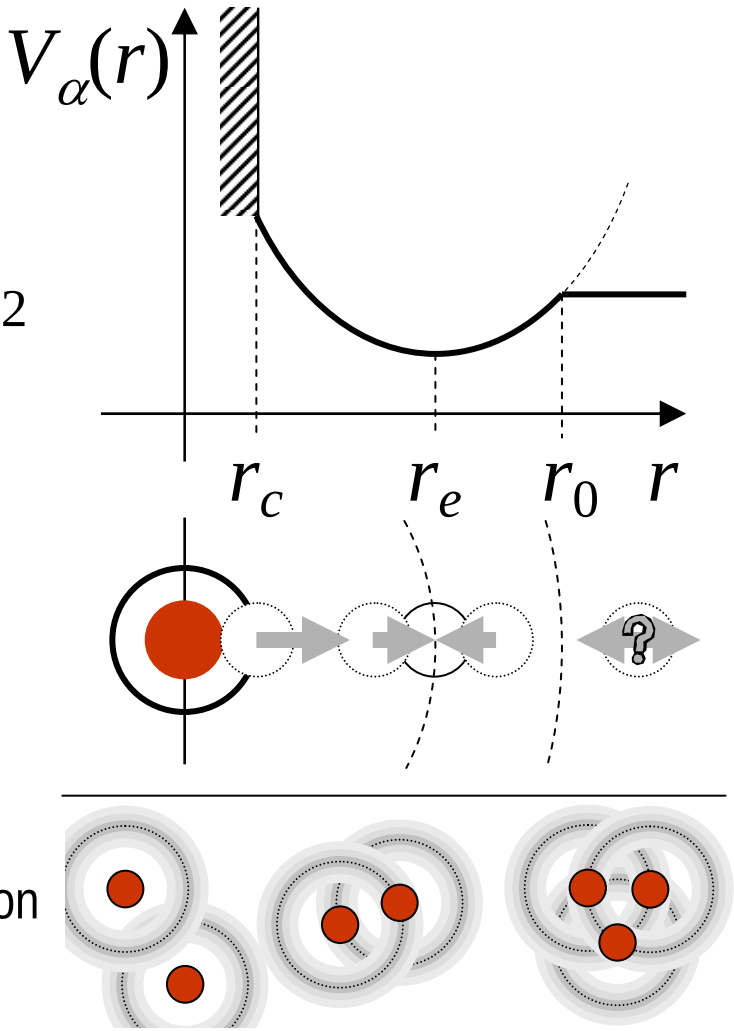
- ✓ illustrating “existence of components” and “binding fate”
  - in 2-D space, two types of particles (α and β)
  - attractive and repulsive interactions, modeled as potentials $V(\mathbf{r})$ around each particle
  - V is the equivalent of a geometrical “shape”, i.e., specific binding affinities



1. Self-Assembly

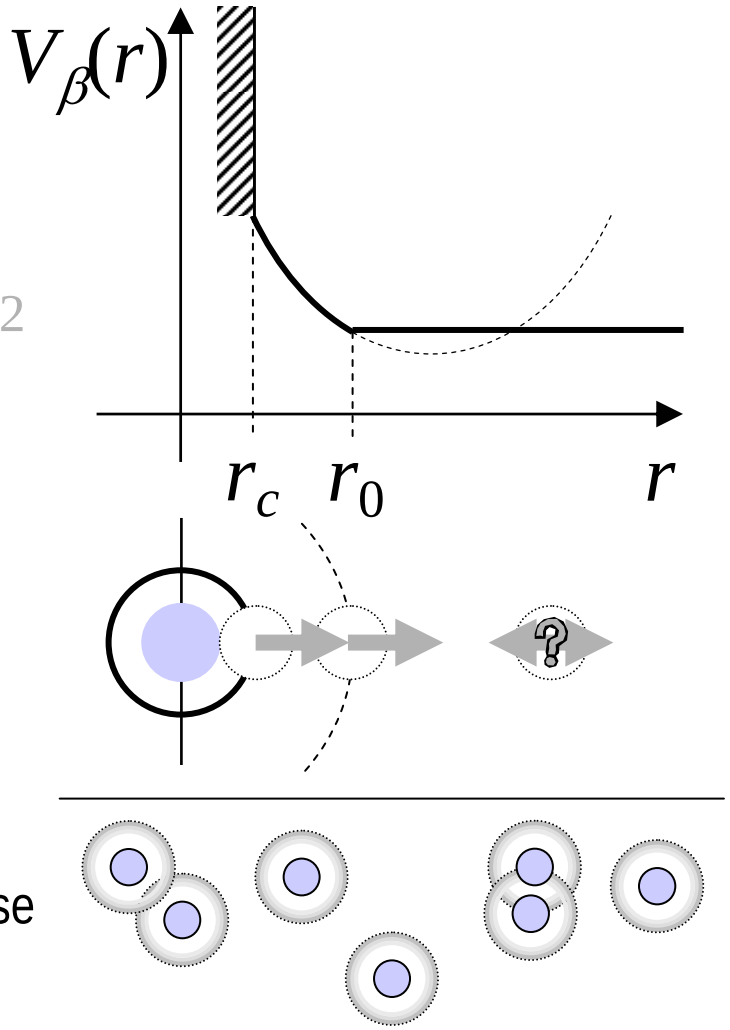
➢ Type-α potential

- ✓ isotropic
  - $V_\alpha(\mathbf{r}) = V_\alpha(r)$, with $r = \| \mathbf{r} \|$
- ✓ hard core below r_c
  - non-deformable particles of radius $r_c / 2$
  - $V_\alpha(r) = +\infty$ for $r < r_c$
- ✓ horizont beyond r_0
  - $V_\alpha(r) = \text{cst}$ for $r > r_0$
  - particles do not see one another
- ✓ equilibrium at r_e
  - $V_\alpha(r) \sim (r - r_e)^2$ for $r_c < r < r_0$
  - ring-shaped quadratic basin of attraction



1. Self-Assembly

- Type-β potential
  - ✓ isotropic
    - $V_\alpha(\mathbf{r}) = V_\alpha(r)$, with $r = \| \mathbf{r} \|$
  - ✓ hard core below r_c
    - non-deformable particles of radius $r_c / 2$
    - $V_\alpha(r) = +\infty$ for $r < r_c$
  - ✓ horizont beyond r_0
    - $V_\alpha(r) = \text{cst}$ for $r > r_0$
    - particles do not see one another
  - ✓ no attraction
    - corresponds to $r_0 < r_e$
    - only repels particles that come too close



1. Self-Assembly

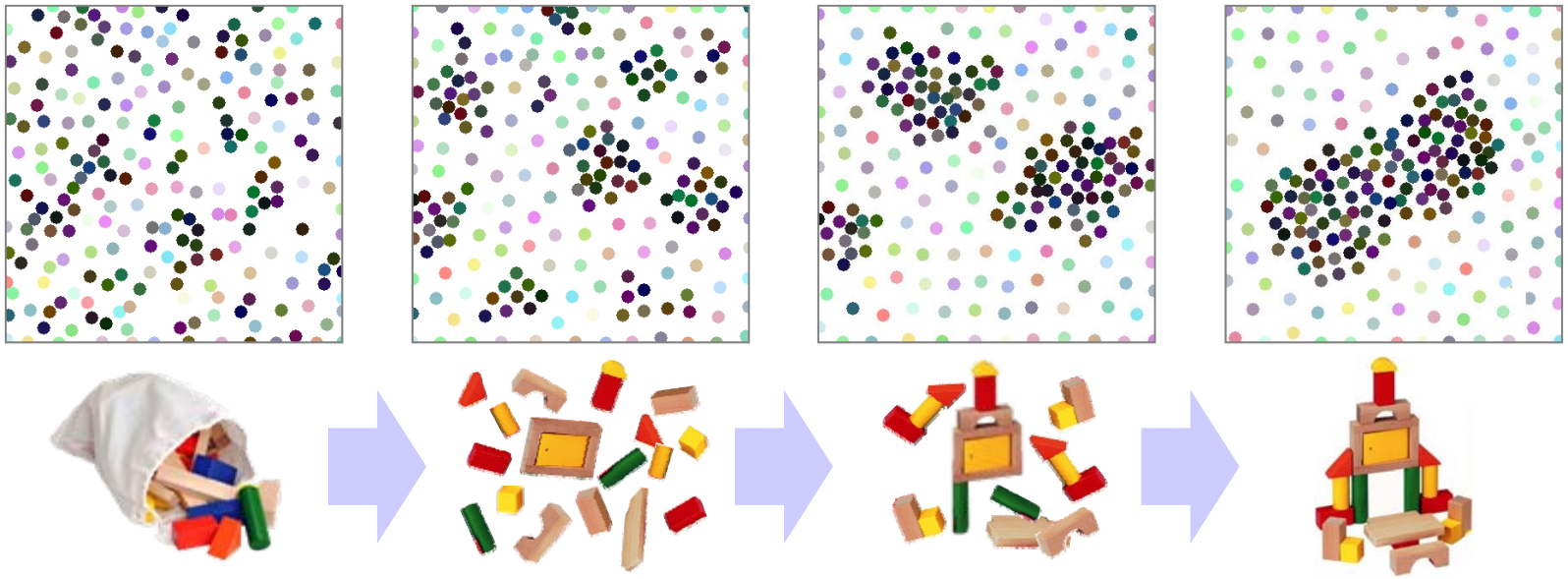
➢ Particle dynamics

- ✓ similar to collective motion models (Reynolds, Vicsek)
  - except particles are not self-propelled; no constant speed enforced
  - velocity may vary in both norm and direction
- ✓ simple equations of motion
  - with inertia:
    $$m\ddot{\mathbf{x}}_i = -\lambda\dot{\mathbf{x}}_i - \sum_j \nabla_i V(\mathbf{x}_j, \mathbf{x}_i) + \eta$$
  - without inertia:
    $$\lambda\dot{\mathbf{x}}_i = -\sum_j \nabla_i V(\mathbf{x}_j, \mathbf{x}_i) + \eta$$
  - $\mathbf{x}_i$ is the position of particle i, λ is viscosity, η is noise
  - $V(\mathbf{x}_j, \mathbf{x}_i) = V(r_{ij} = \| \mathbf{x}_i - \mathbf{x}_j \|)$ is the potential created by particle j in $\mathbf{x}_i$
  - V actually depends on *both* types of particles i and j: $V = V_\alpha$ for α–α interactions, and $V = V_\beta$ for α–β and β–β interactions

1. Self-Assembly

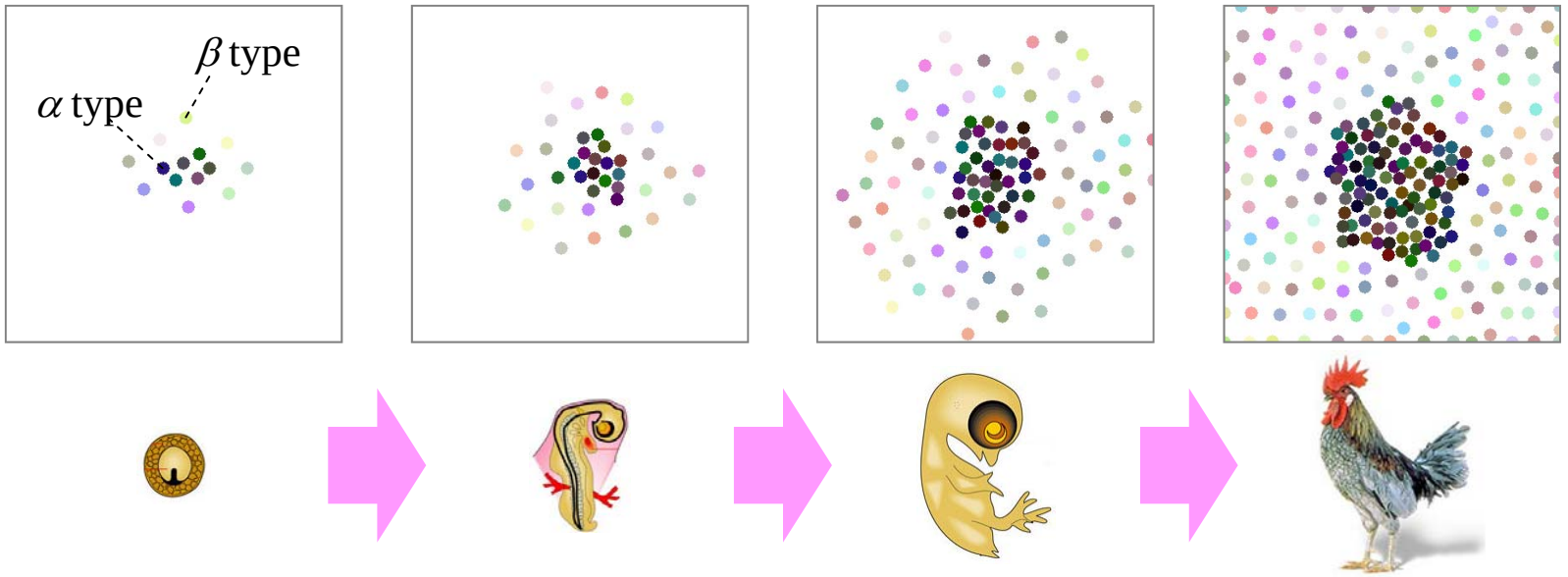
- *Molecular*-style SA: structuration from a random mix
  - ✓ "shaking the puzzle box"
    - α particles randomly collide and cluster together within a sea of β particles
    - like molecules, dissociated cells can also spontaneously sort again
    - however, mostly in artificial experiments; not a major natural mechanism

    → *the complex architecture of an organism does not emerge out of a giant swarm of trillions of disaggregated cells reassembling in parallel*



1. Self-Assembly

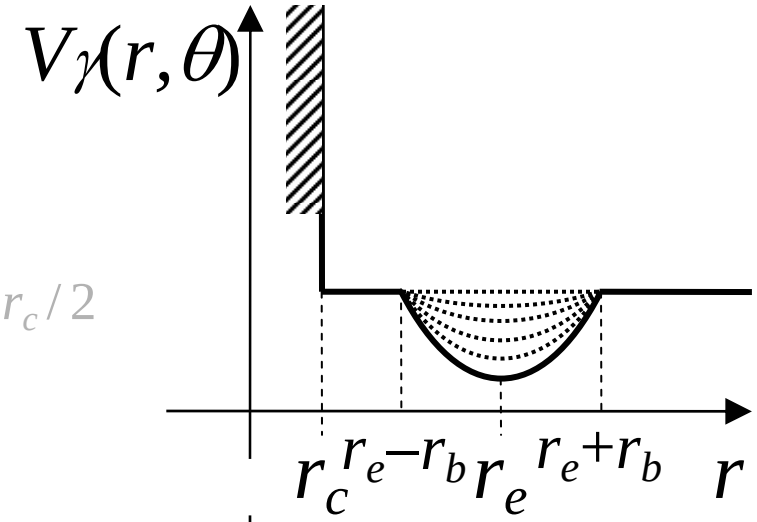
- *Multicellular*-style SA: structuration from development
  - ✓ "growing the embryo"
    - starting with only a few particles of each type
    - particles *divide* into same-type particles, under uniform probability
    - new particles pop up *pre-positioned* near the type that produced them
    - particles only briefly rearrange within their local neighborhood

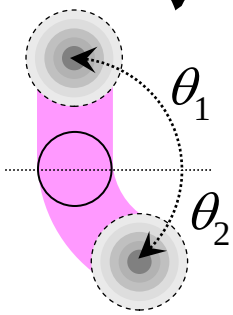


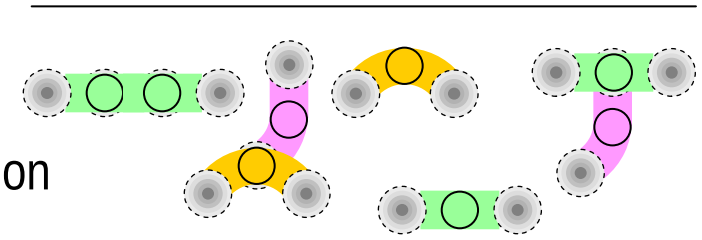
1. Self-Assembly

➢ Illustrating "shape determination": type-γ potential

- ✓ anisotropic
  - $V_\alpha(\mathbf{r}) = V_\alpha(r, \theta)$
- ✓ hard core below r_c
  - non-deformable particles of radius $r_c / 2$
  - $V_\alpha(r) = +\infty$ for $r < r_c$
- ✓ horizont beyond r_0
  - $V_\alpha(r) = \text{cst}$ for $r > r_0$
  - particles do not see one another
- ✓ bipolar attraction
  - $V_\alpha(\mathbf{r}) \sim (\mathbf{r} - \mathbf{r}_1)^2 + (\mathbf{r} - \mathbf{r}_2)^2$, with $\mathbf{r}_1 = (r_e, \theta_1)$ and $\mathbf{r}_2 = (r_e, \theta_2)$
  - two localized quadratic basins of attraction



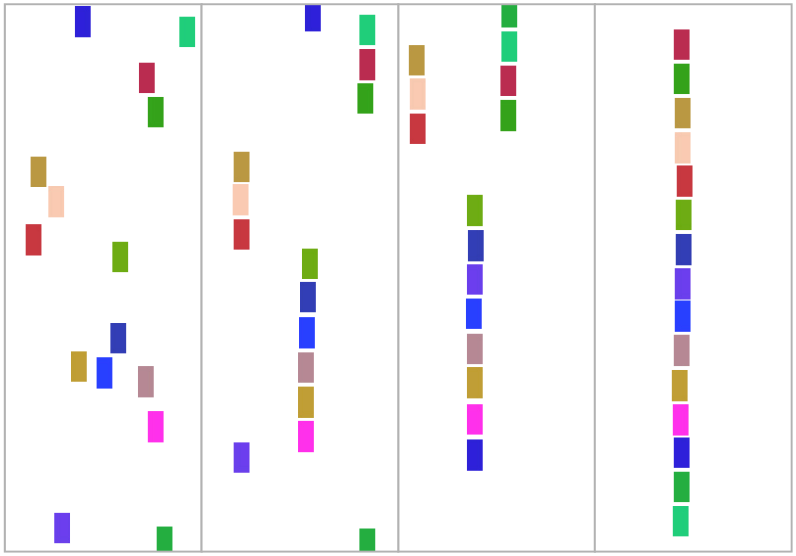


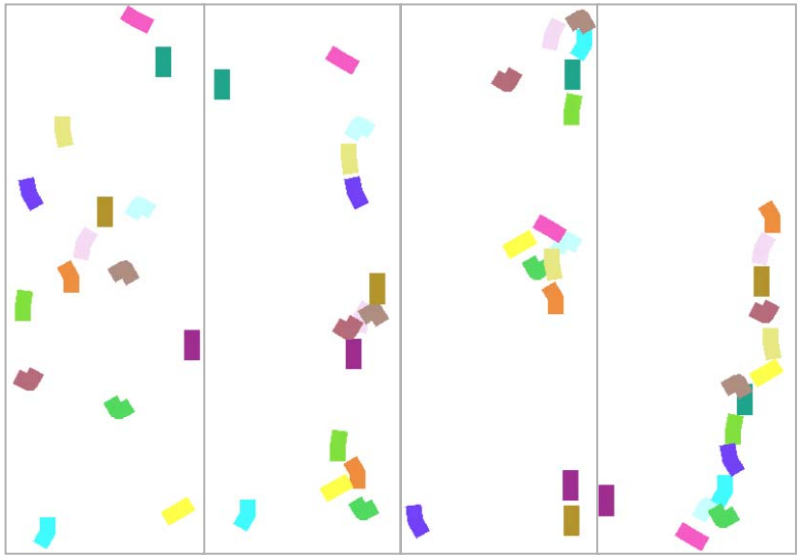


1. Self-Assembly

➢ *Molecular*-style SA: colliding pre-shaped particles

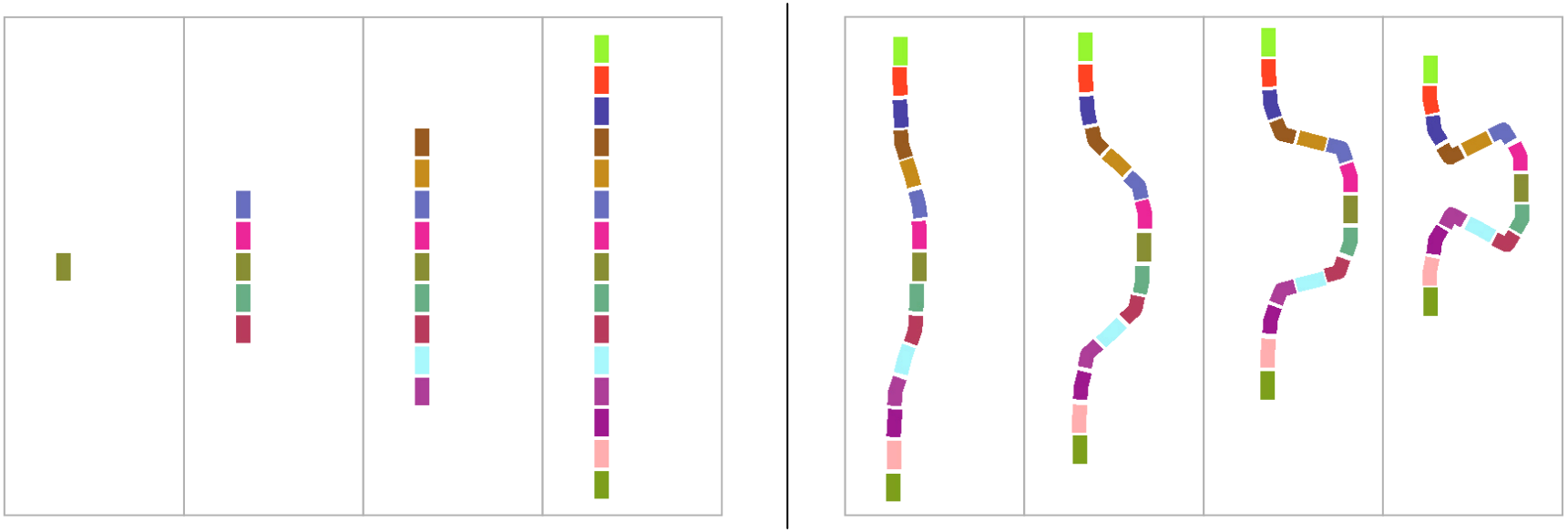
- ✓ 15 particles of type γ interacting via polar potential $V_\gamma(\mathbf{r})$
  - drawn as small rectangles (straight or bent) instead of discs
  - **colliding SA:** identical particles with vertical poles $(\theta_1, \theta_2) = (\pi/2, -\pi/2)$ snap into place, forming a straight chain
  - **pre-shaped SA:** uniquely shaped particles, with various (θ_1, θ_2), are unable to coordinate: they only explore suboptimal and unstable states





1. Self-Assembly

- ***Multicellular*-style SA: growing and reshaping particles**
  - ✓ 15 particles of type γ interacting via polar potential $V_\gamma(\mathbf{r})$
    - drawn as small rectangles (straight or bent) instead of discs
    - **growing SA:** the same string can be formed by dividing vertical particles
    - **reshaping SA: then**, each particle dynamically bends its shape in specific ways, making the string invaginate (final angles same as pre-shaped particles)



1. Self-Assembly

- Biological cells use mechanisms that greatly facilitate SA
  - ✓ future artificial systems design could follow a similar approach
    - instead of letting components haphazardly try to match each other's pre-existing constraints, like molecules in a solution. . .
    - . . . let components dynamically create and reshape themselves "on the spot," as cells do
  - ✓ from *stochastic* (molecular-style) self-assembly to *programmable* (multicellular-style) self-assembly
    - components must be able to dynamically modify their behavior (divide, differentiate, migrate) through *communication*
    - cells do not just snap into place; they send molecular signals to each other
  - → *cells form **patterns of differentiation** at the same time that they are self-assembling*

The Self-Made Puzzle

1. Self-Assembly of Pre-Patterned Components

2. Pattern Formation in Pre-Assembled Media

3. Integrating Self-Assembly and Pattern Formation Under Genetic Regulation

4. Toward Evolutionary Meta-Design

2. Pattern Formation

➢ Pattern formation vs. morphogenesis

- ✓ since Turing (1952), “morphogenesis” is often confused with “pattern formation”
  - yet they do not emphasize the same aspect of emerging order
- ✓ pattern formation = emergence of statistically regular *motifs*
  - in quasi-continuous and initially homogeneous 2-D or 3-D media
  - shimmering landscapes of *activity* on a more or less fixed backdrop

  → *pattern formation “paints” a pre-existing space*
- ✓ morphogenesis = generation of complex, heterogeneous *form*
  - originally, biological development of organs and structures of an organism
  - by extension: physical (geomorphogenesis), social (urban morphogenesis)...
  - creation of intricate *architectures* and structures

  → *morphogenesis “sculpts” its own space*

2. Pattern Formation

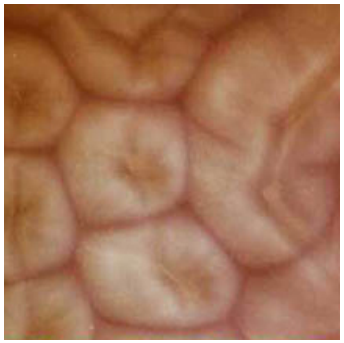
- **Diversity of pattern formation behaviors**
  - ✓ different substrates and scales
    - fluid, electromagnetic, mechanical, chemical, biochemical
  - ✓ different classes of mechanisms and models
    - convection cells, reaction-diffusion, activator-inhibitors, synchronization of oscillators
  - ✓ different types of patterns
    - static, steady-state or dynamically changing (traveling waves)
    - classical geometrical families: spots, stripes, spirals, branches

2. Pattern Formation

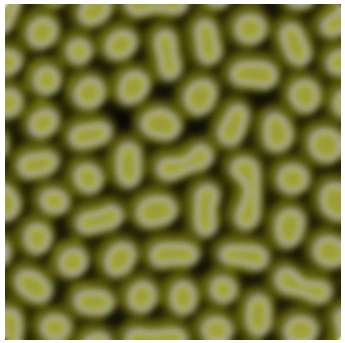
➢ Traditional PF is stochastic, biological PF is not

- randomness at micro-level (elts) and meso-level (patterns)
- PF research focuses on *instabilities* and amplification of fluctuations
- outcome generally *unpredictable* in number and position of domains
- conversely, macroscopic formation fairly regular: repeated motifs, statistical *uniformity* like textures

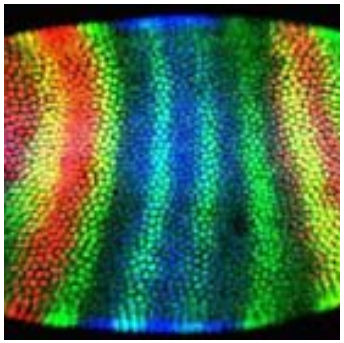
- mesoscopic organs and limbs have intricate, non-random morphologies
- reaction-diffusion based(?) animal coats are only a marginal aspect
- development is *reproducible* in number and position of body parts
- most of organism development is under deterministic genetic control: *heterogeneous*, rich in information



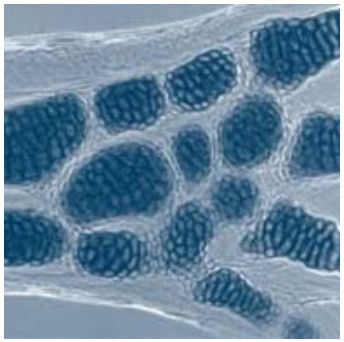
convection cells
www.chabotspace.org



reaction-diffusion
texturegarden.com/java/rd



fruit fly embryo
Sean Caroll, U of Wisconsin



larval axolotl limb
Gerd B. Müller

2. Pattern Formation

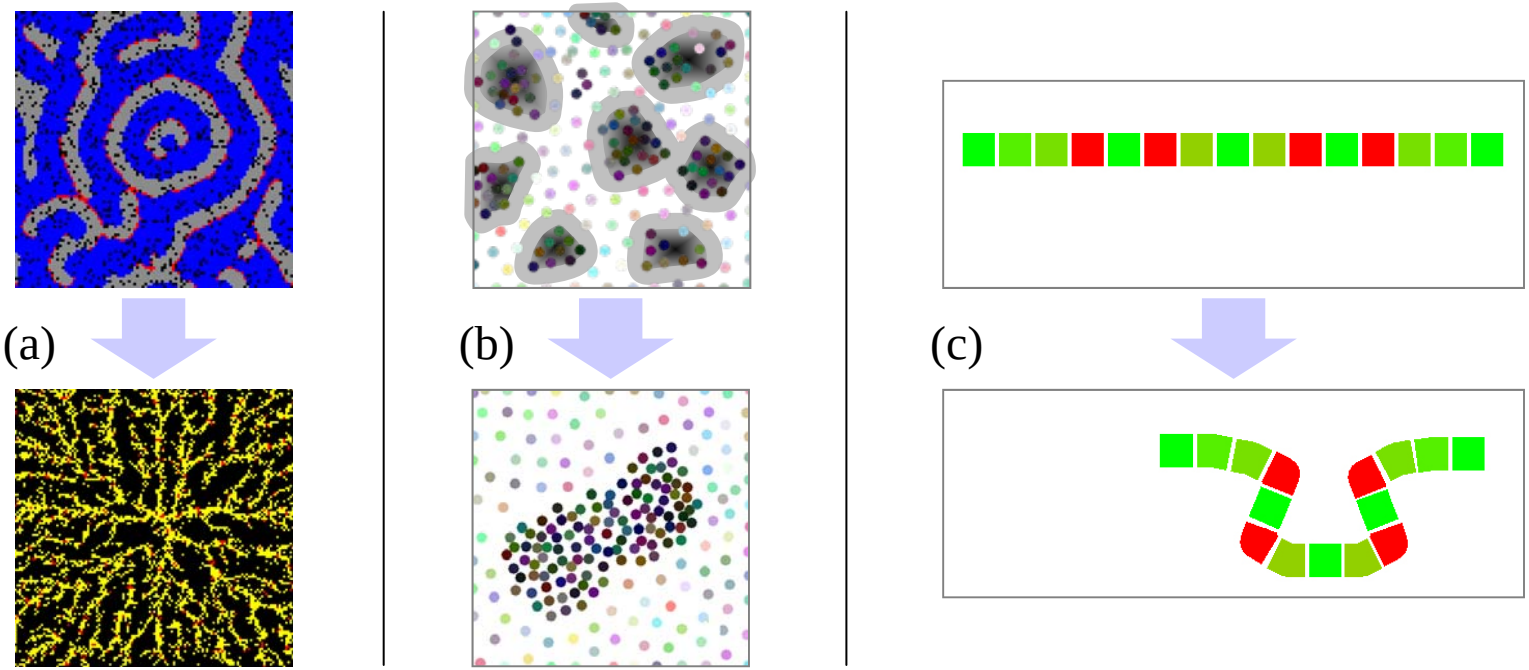
- Biological morphogenesis relies on *informed* agents
  - ✓ non-biological, physical-chemical pattern formation
    - elements are molecules, simple bodies or elementary volumes of homogeneous solution
    - each element contains very little information, making simple constraints (activation vs. inhibition)
  - ✓ biological, multicellular morphogenesis
    - unique characteristic: each one of its self-organizing elements, the cell, contains a rich source of information stored in the DNA
    - this information endows it with a vast repertoire of highly non-trivial behaviors
    - even admitting that DNA is less than a “program,” it is still at least, a *repository of stimuli-response rules*, vastly superior in quantity of functional information to purely physical-chemical elements

2. Pattern Formation

- Embryogenesis combines PF and morphogenetic SA
  - ✓ shapes from patterning; patterns from shaping
    - structures are “sculpted” from the self-assembly of elements, prompted by the “painting” of their genetic identity
    - conversely, newly formed shapes are able to support, and trigger, new domains of genetic expression
  - ✓ tightly integrated loop under non-random genetic regulation
    - DNA is “consulted” at every step of this exchange, in every cell
    - it produces the proteins that guide the cell’s highly specific biomechanic behavior (shaping) and signalling behavior (patterning)

2. Pattern Formation

- "Shape from patterning" examples
  - ✓ deriving morphogenetic SA (bottom frames) from PF (top frames)
    a) slime mold amoebae first generate waves of chemical signalling (top), then follow concentration gradients and aggregate (bottom)
    b) type-α particles differentiating from a prepattern before assembling
    c) bending angle of each γ particle also determined by a prepattern of identity



http://zool33.uni-graz.at/schmickl

The Self-Made Puzzle

1. Self-Assembly of Pre-Patterned Components

2. Pattern Formation in Pre-Assembled Media

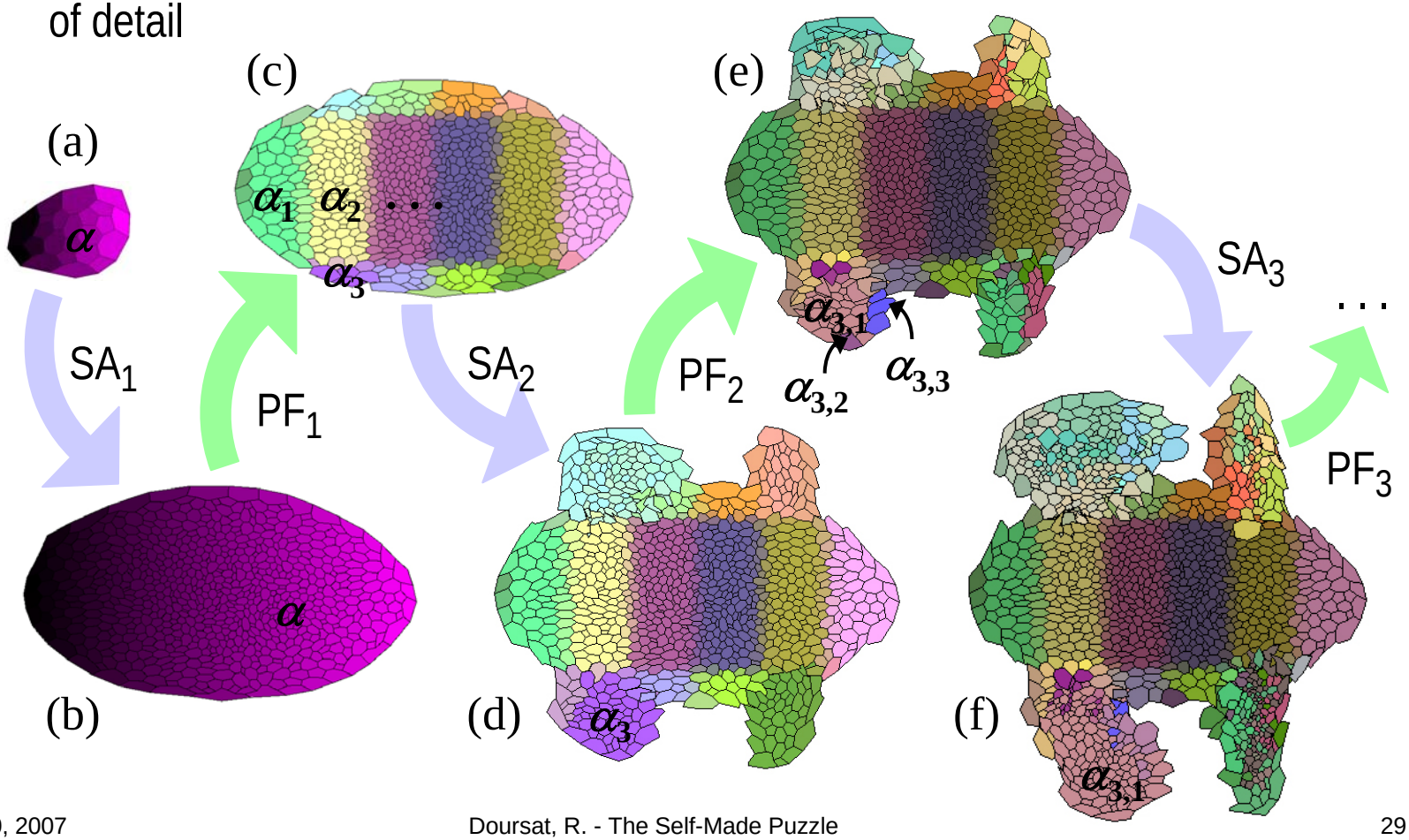
3. **Integrating Self-Assembly and Pattern Formation Under Genetic Regulation**

4. Toward Evolutionary Meta-Design

3. Integrating SA and PF

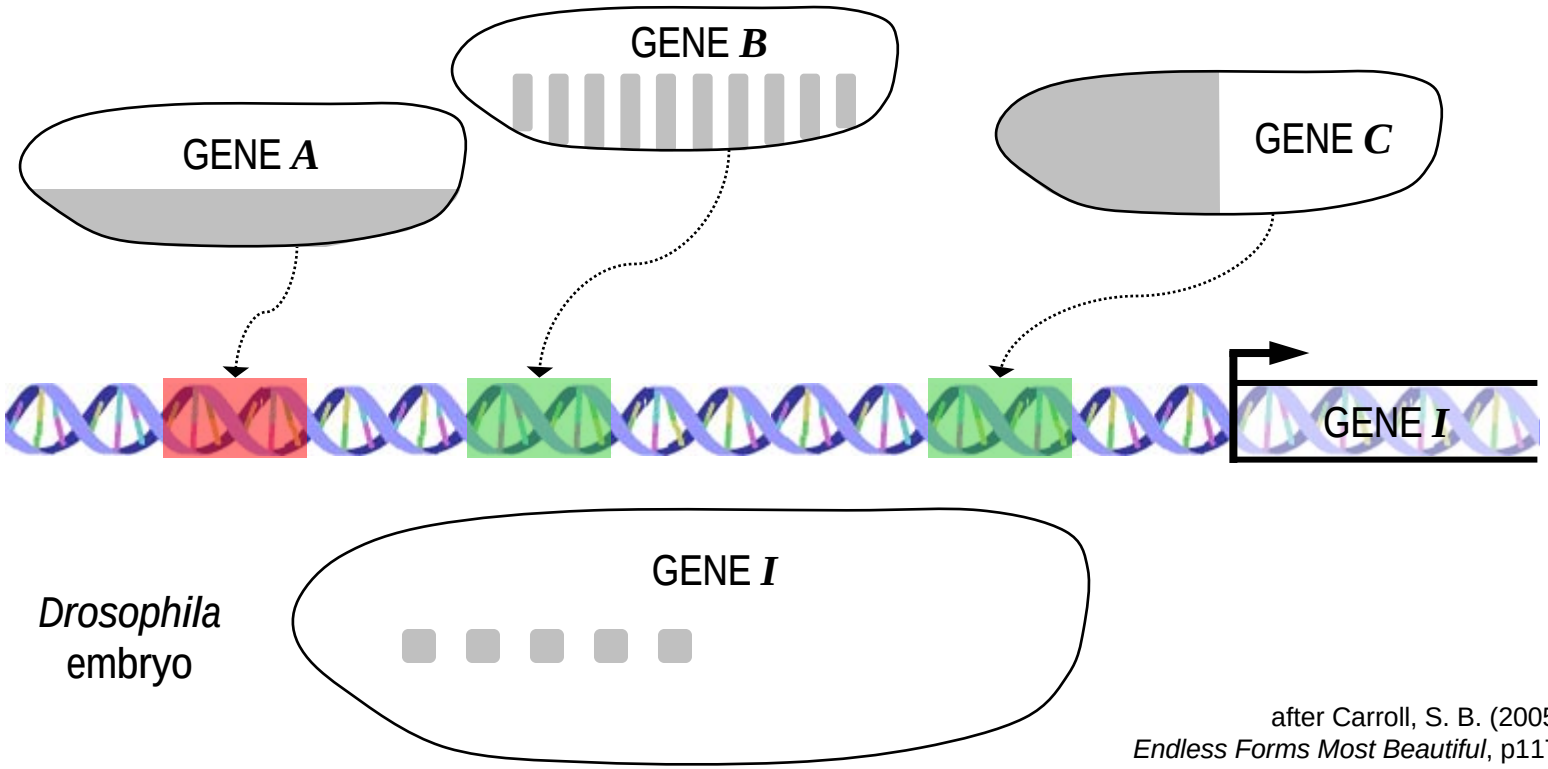
Embryomorphic architectures

- *functional dependency between cell identities and mechanical cell behaviors*
- alternation of PF-induced differentiation and heterogeneous-type SA at all scales of detail



3. Integrating SA and PF

- Developmental genes are expressed in spatial domains
  - ✓ thus combinations of switches can create patterns by union and intersection, for example: $I = (\text{not } A) \text{ and } B \text{ and } C$

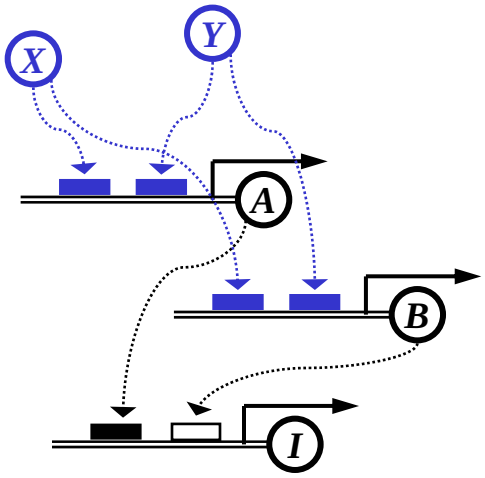


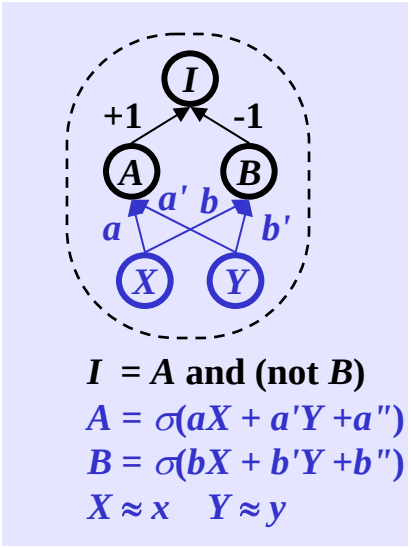
after Carroll, S. B. (2005)
Endless Forms Most Beautiful, p117

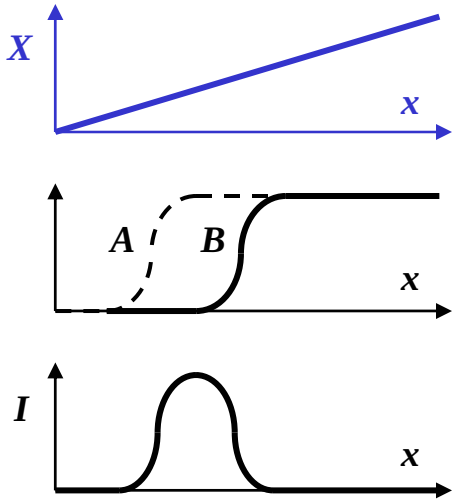
3. Integrating SA and PF

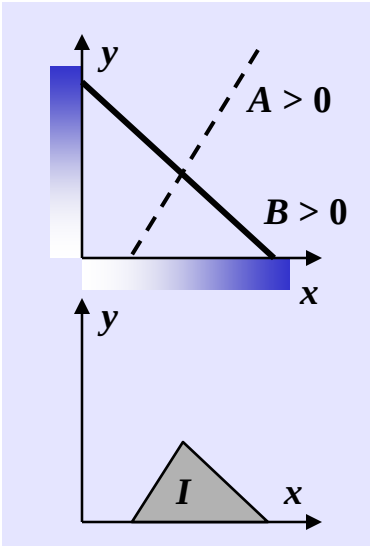
➢ Three-tier GRN model: integrating positional gradients

✓ A and B are themselves triggered by proteins X and Y







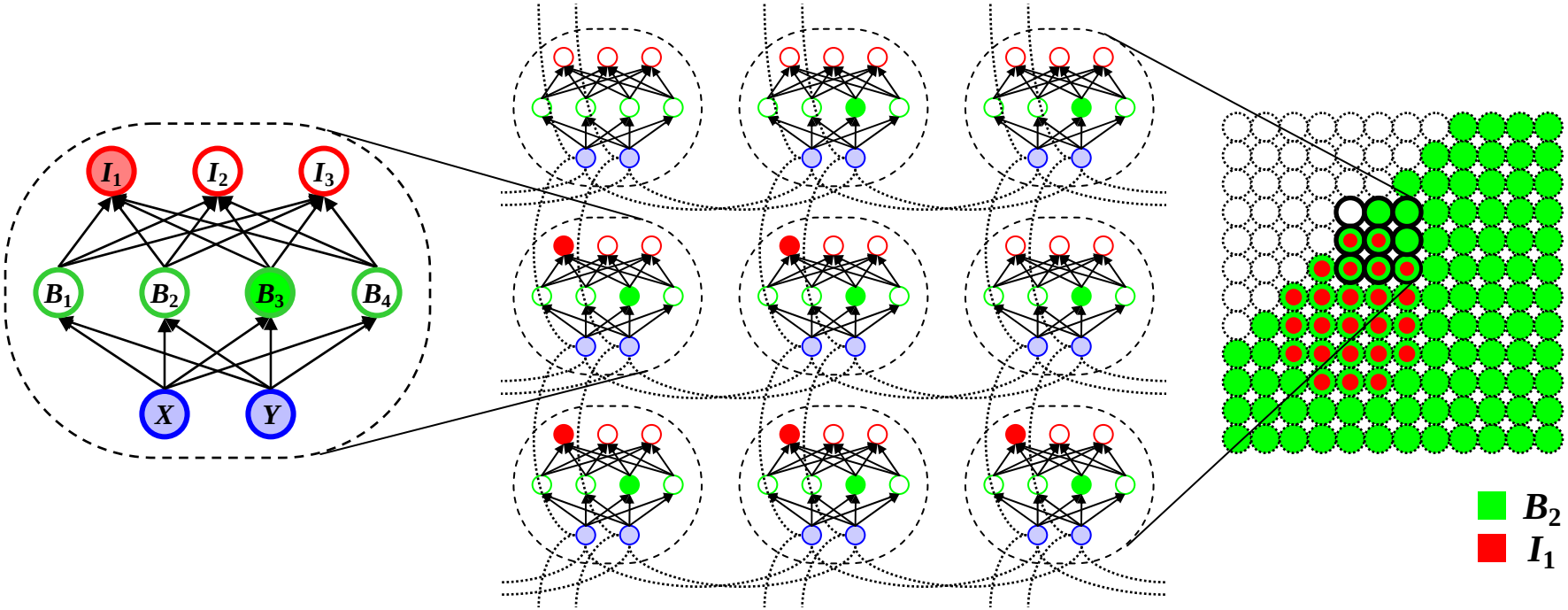


✓ X and Y diffuse along two axes and form concentration gradients

→ *different thresholds of lock-key sensitivity create different territories of gene expression in the geography of the embryo*

3. Integrating SA and PF

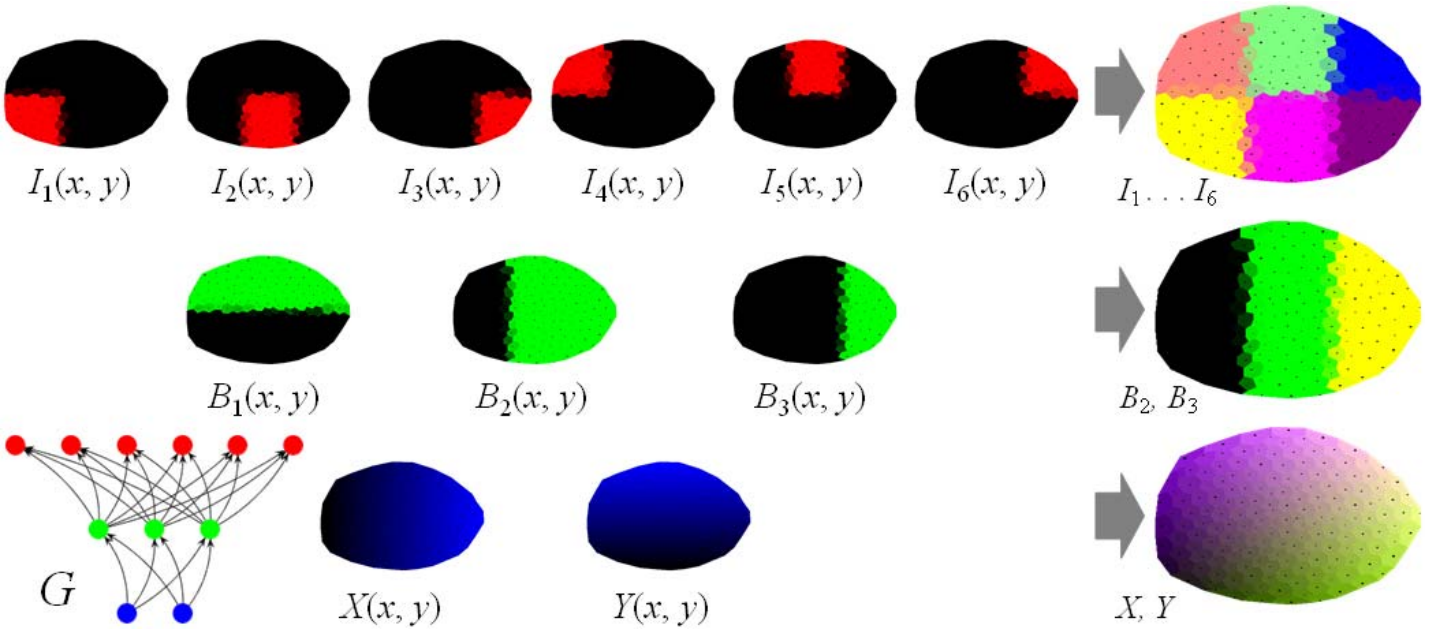
- A lattice of Positional-Boundary-Identity (PBI) GRNs
  - ✓ network of networks: each GRN is contained in a cell, coupled to neighboring cells via the positional nodes (for diffusion)
  - ✓ a pattern of gene expression is created on the lattice



3. Integrating SA and PF

➢ The hidden geography of the embryo

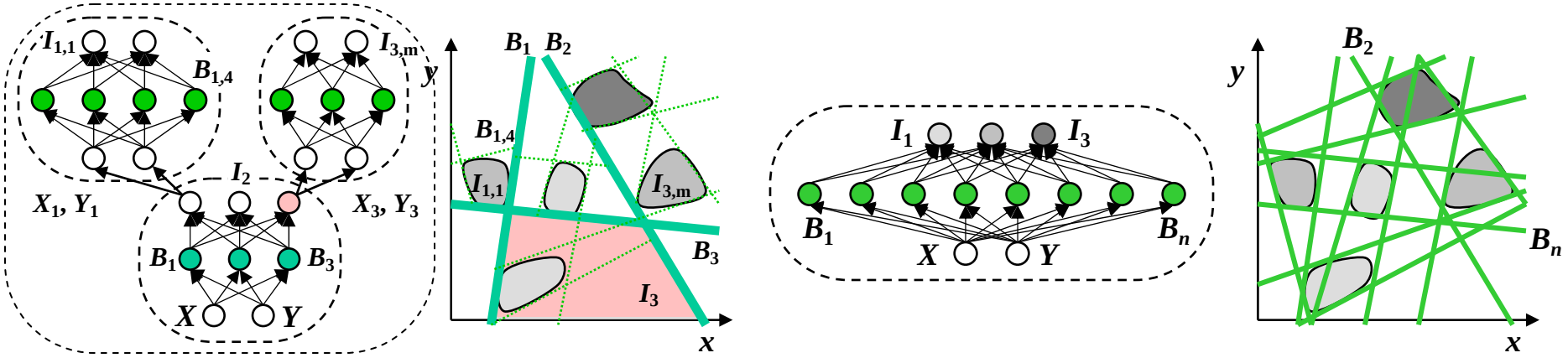
- ✓ self-patterning obtained from a 3B-6I gene regulatory network *G* in a 200-cell oval-shaped embryo
- ✓ each view is “dyed” for the expression map of one of the 11 genes, e.g.: $B_1 = \sigma(Y - 1/2)$, $B_2 = \sigma(X - 1/3)$, $I_6 = B_1 B_3$ ...



3. Integrating SA and PF

➢ Multiscale refinement using a hierarchical GRN

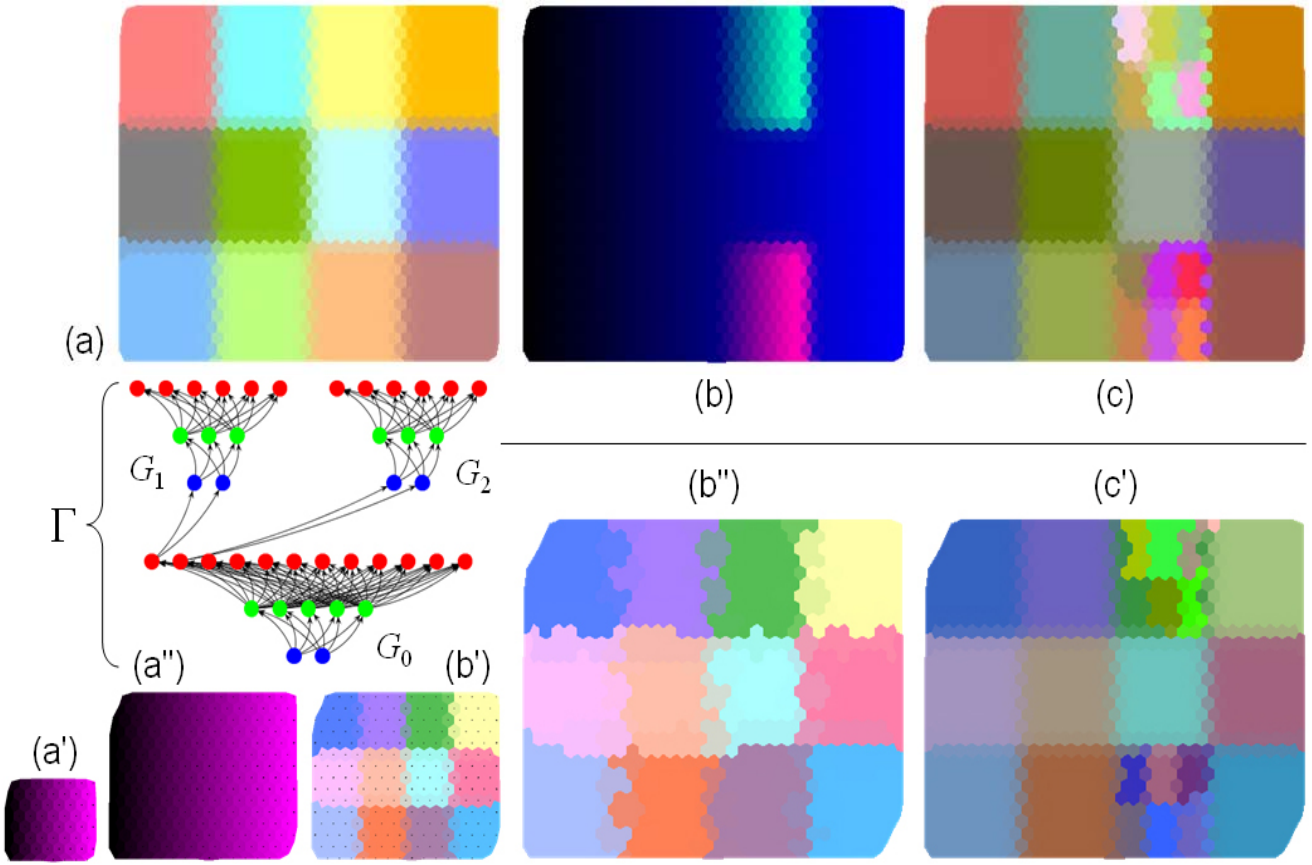
- ✓ instead of one flat tier of B nodes, use a pyramid of PBI modules
- ✓ the activation of an I node controls the onset of a new P layer
- ✓ in the first stage, a base PBI network creates broad domains



- ✓ in the next stage, another set of PBI networks subdivide these domains into compartments at a finer scale, etc.

3. Integrating SA and PF

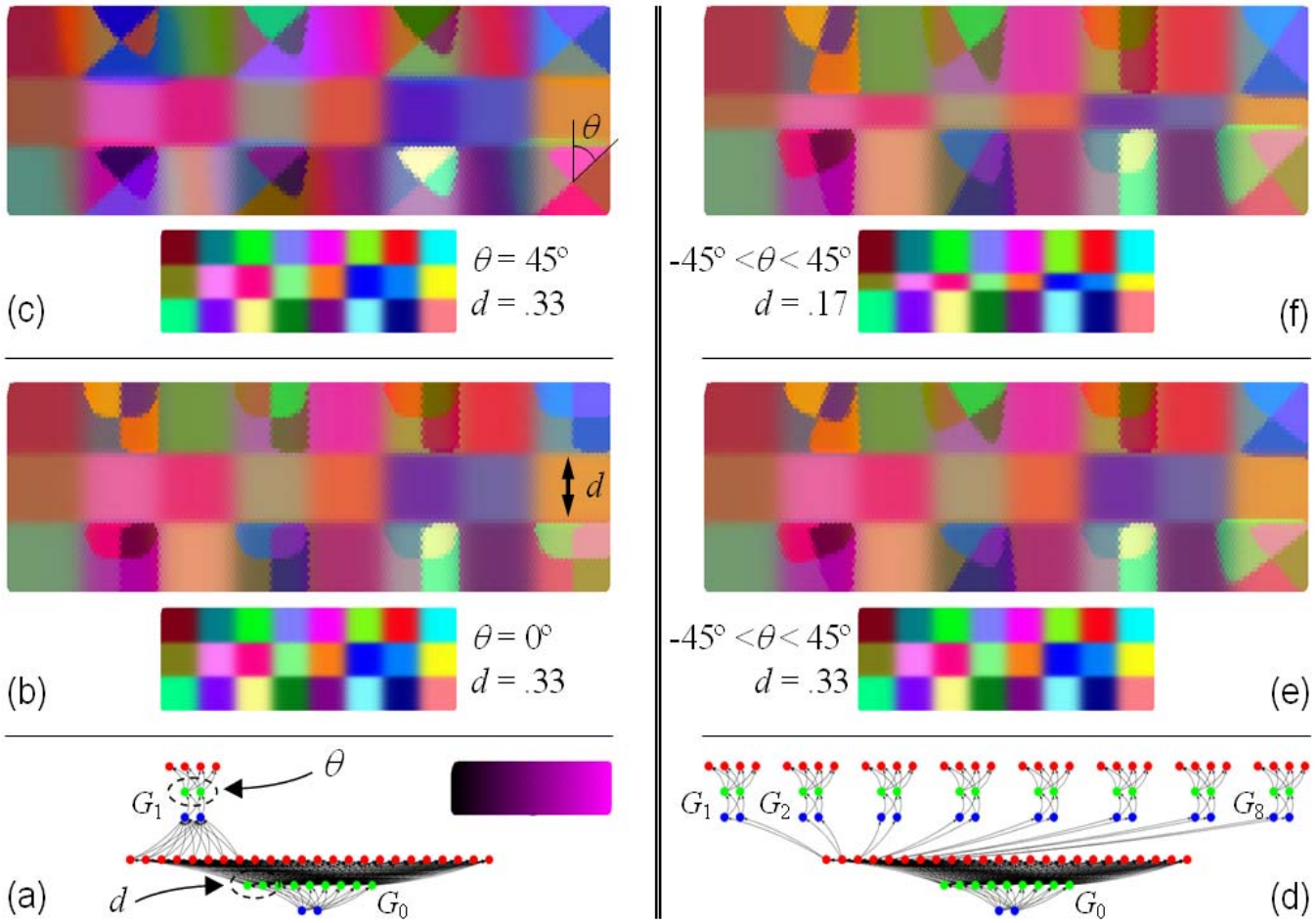
- Static vs. growing multiscale canvas
  - ✓ 32x32 hexagonal lattice of cells, two-level gene network Γ: base subnet G_0, then 2 subnets G_1, G_2 triggered by I_1 and I_2



equivalent pattern obtained by *uniform expansion* from 8x8 cells

3. Integrating SA and PF

➢ The inherent modularity of hierarchical GRNs

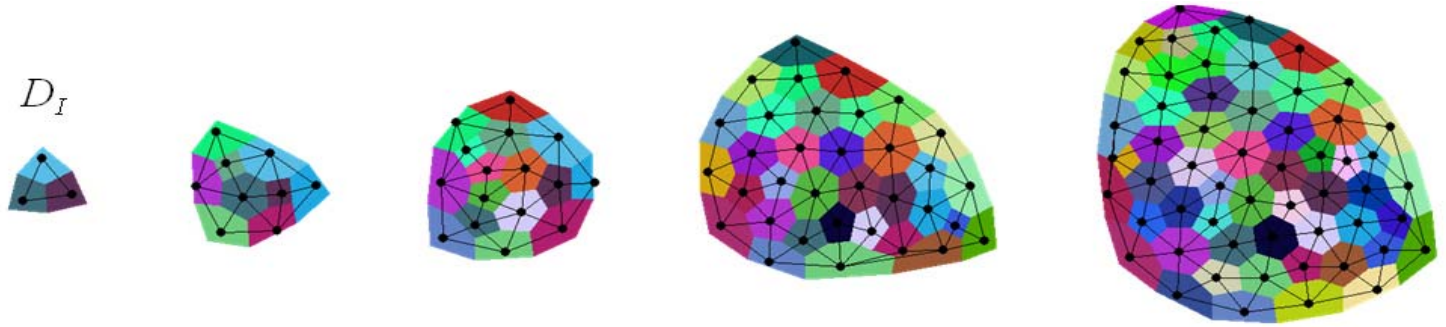


- ✓ organisms contain "homologous" parts (arthropod segments, vertebrate teeth and vertebrae, etc.)
- ✓ homology also exists between species (tetrapod limbs)
- ✓ similarities in DNA sequences reveal that homology is the evolutionary result of *duplication* followed by *divergence*

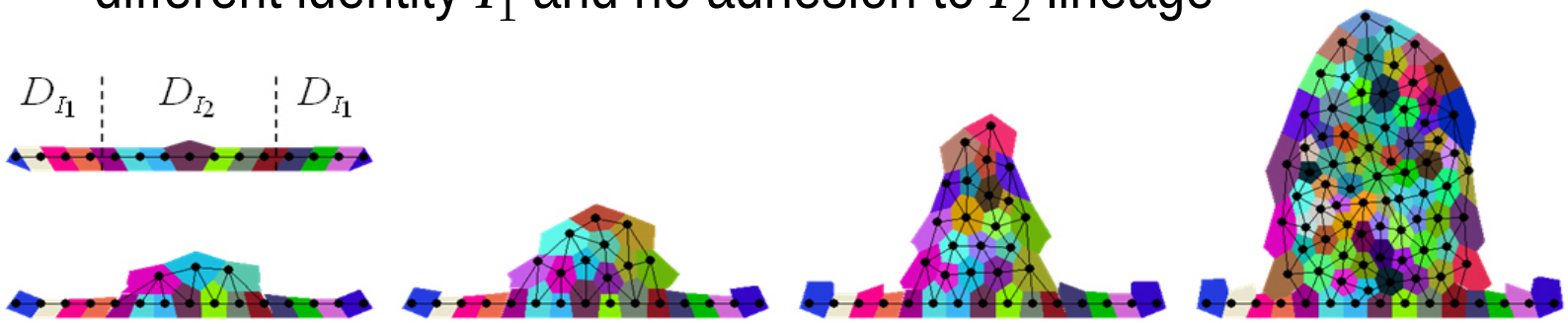
3. Integrating SA and PF

- Simple mesh model of cell adhesion and elasticity

a) isotropic “blob” of identical cells dividing at 1% rate, in which nearby daughter cells rearrange under elastic forces



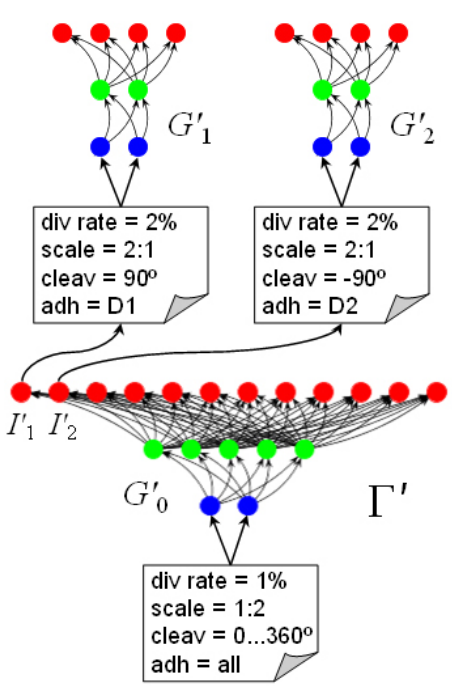
b) anistropic “limb” growth: only center domain I_2 divides (upward stretch due to $2x{:}y$ anisotropic rescaling); lateral cells have different identity I_1 and no adhesion to I_2 lineage

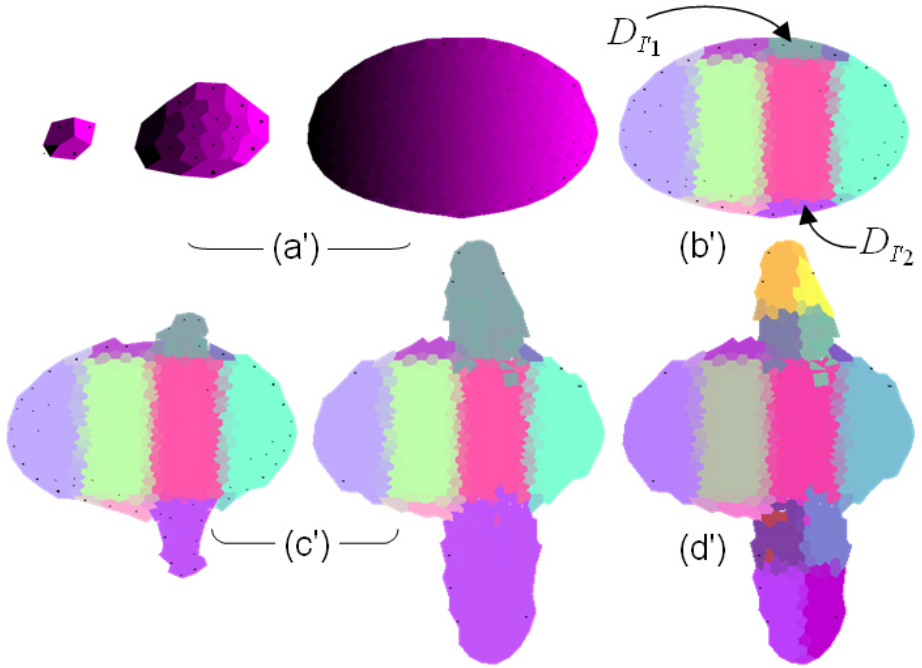


3. Integrating SA and PF

➢ Inhomogeneous cell division (cont'd)

- ✓ using differential adhesion, anisotropic cleavage planes and rescaling, this model can also generate directional offshoot akin to limb development





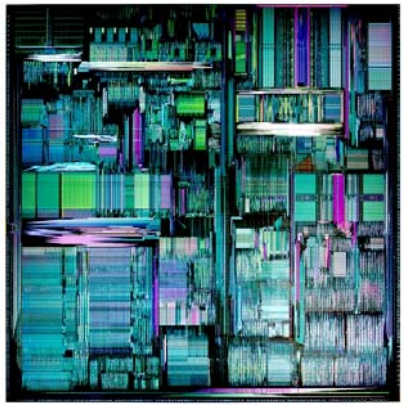
- ✓ here, different weights in base module G'_0 make a thicker central row, and place I'_1 and I'_2 dorsally and ventrally
- ✓ different adhesion coefficients also make I'_1 and I'_2 grow "limbs", sub-patterned by G'_1 and G'_2

The Self-Made Puzzle

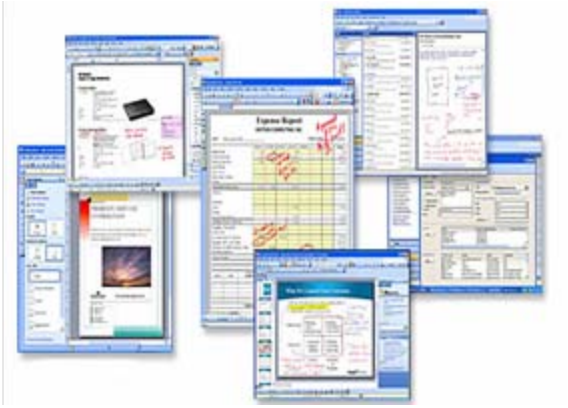
1. Self-Assembly of Pre-Patterned Components
2. Pattern Formation in Pre-Assembled Media
3. Integrating Self-Assembly and Pattern Formation Under Genetic Regulation
4. Toward Evolutionary Meta-Design

4. Evolutionary Meta-Design

- Rapid growth in size & complexity of computer systems,

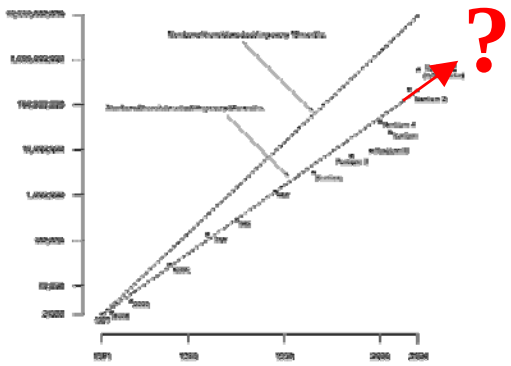


whether hardware,

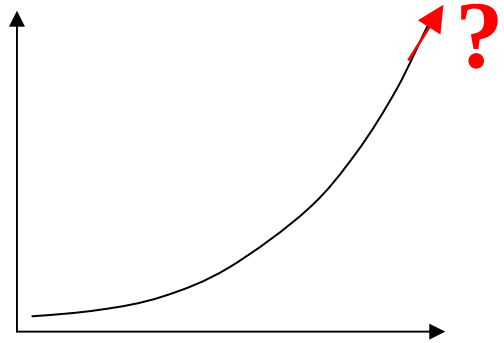


software,

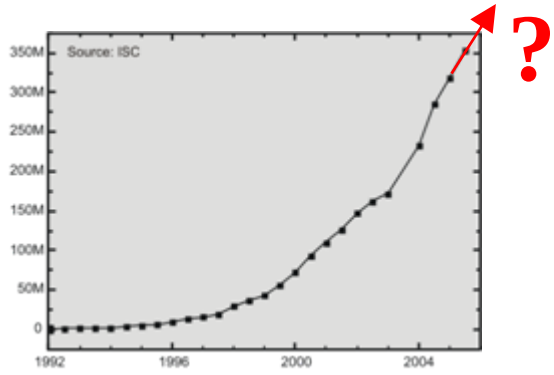
or networks, ...



number of transistors/year



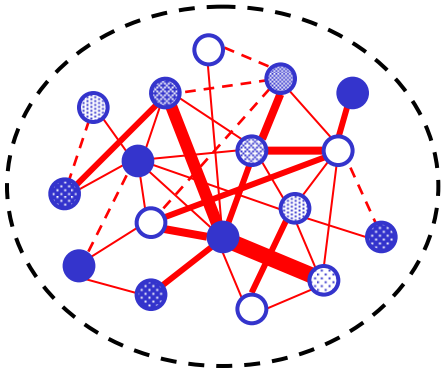
number of O/S lines of code/year



number of network hosts/year

4. Evolutionary Meta-Design

➢ ... leads us to rethink engineering as *complex systems*



- large number of elements interacting locally
- simple individual behaviors creating a complex emergent behavior
- decentralized dynamics: no master blueprint or grand architect

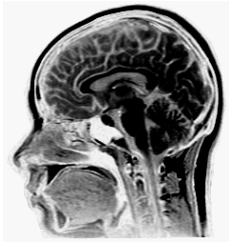
✓ in particular, seek inspiration from biological and social systems

physical pattern formation

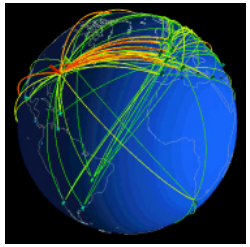
organism development

insect colonies

the brain



World Wide Web



social networks

4. Evolutionary Meta-Design

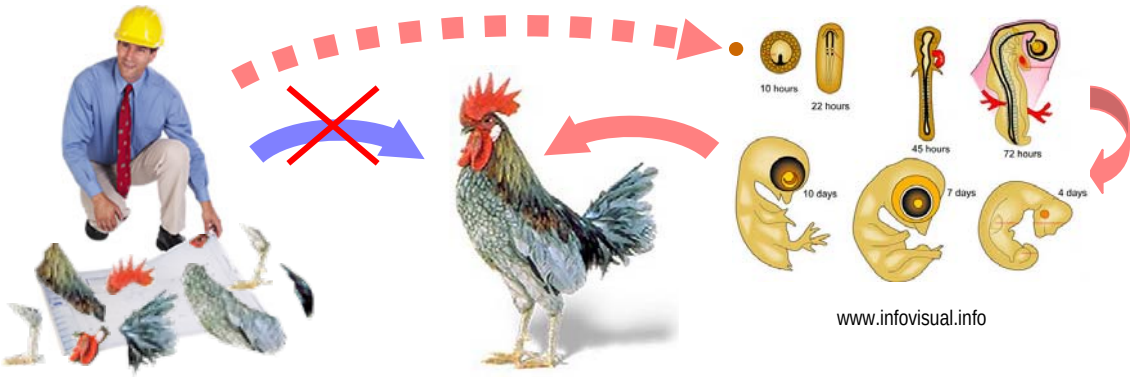
- Natural adaptive systems as a new paradigm for ICT
  - ✓ natural complex adaptive systems, biological or social, can become a new and powerful source of inspiration for future IT in its transition toward autonomy
  - ✓ "emergent engineering" will be less about direct design and more about developmental and evolutionary meta-design
  - ✓ it will also stress the importance of constituting fundamental laws of *development* and developmental *variations* before these variations can even be selected upon in the evolutionary stage
  - ✓ it is conjectured that fine-grain, hyperdistributed systems will be uniquely able to provide the required "solution-rich" space for successful evolution by selection

4. Evolutionary Meta-Design

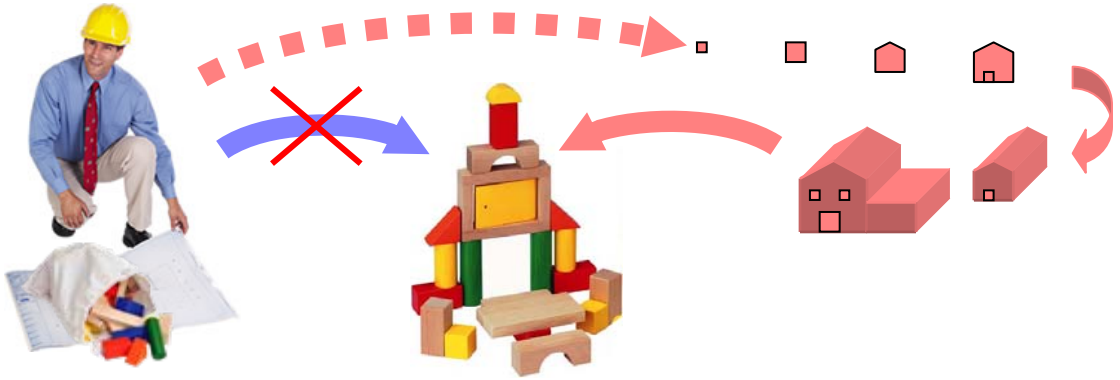
- From centralized heteromy to decentralized autonomy
  - ✓ artificial systems *are built* exogenously, organisms endogenously *grow*

systems design

systems "meta-design"



  - ✓ future engineers should "step back" from their creation and only set *generic* conditions for systems to self-assemble and evolve



4. Evolutionary Meta-Design

➢ Growth, function, selection

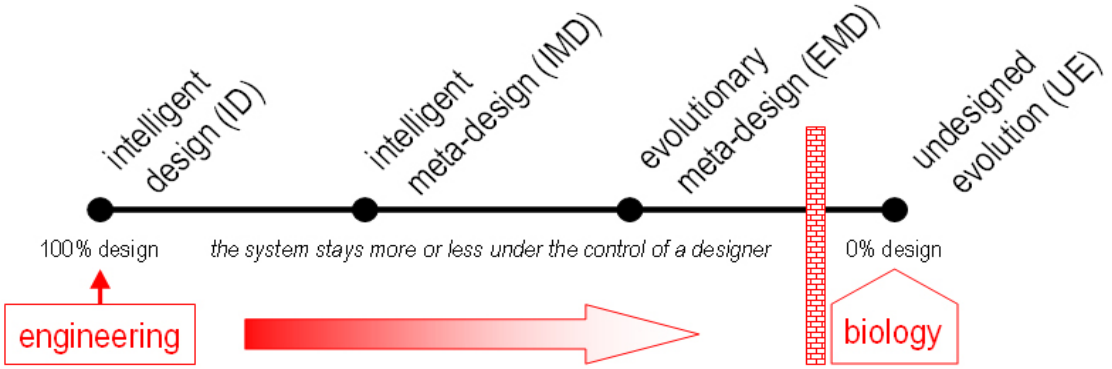
✓ the three challenges of complex systems engineering:

1. how does the system **grow**?
   - development results from a combination of elementary mechanisms: elements change internal state, communicate, travel, divide, die, etc.
   - starting from a single element, a complex and organized architecture develops by repeatedly applying these rules inside each element
   - task (1) consists of combining these principles and designing their dynamics and parameters
2. how does the system **function**?
   - task (2) is about defining the nature of the elements their functionality: hardware components? software modules? robot parts? are they computing? or physically moving? etc.
3. how does the system **evolve** and how is it **selected**?

parameters = "genetic code"

4. Evolutionary Meta-Design

- Pushing engineering toward evolutionary biology



intelligent design		intelligent & evolutionary "meta-design"
heteronomous order	▪	autonomous order
centralized control	▪	decentralized control
manual, extensional design	▪	automated, intentional design
engineer as a micromanager	▪	engineer as a lawmaker
rigidly placing components	▪	allowing fuzzy self-placement
tightly optimized systems	▪	hyperdistributed & redundant systems
sensitive to part failures	▪	insensitive to part failures
need to control	▪	prepare to adapt & self-regulate
need to redesign	▪	prepare to learn & evolve
complicated systems: planes, ~~computers~~	▪	*complex* systems: Web, market ... computers?

4. Evolutionary Meta-Design

- The paradox of complex systems engineering

How can we control complexity?

How can we both "let go" and still have requirements at the same time?

How can we "optimize" the parameters (genetic code) of a self-organized process?

4. Evolutionary Meta-Design

➢ Selecting without expectations

✓ different degrees of fitness constraints

a) selecting for a specific **organism** (shape, pattern)

- reverse problem: given the phenotype, what should be the genotype?
- **direct** recipe; ex: Nagpal's macro-to-microprogram Origami compilation
- otherwise: **learn** or **evolve** under strict fitness → difficult to achieve!

b) selecting for a specific **function**, leaving freedom of architecture

- given a task, optimize performance (computing, locomotion, etc.)
- be surprised by pattern creativity; ex: Avida, GOLEM, Framsticks

c) selecting the **unexpected**

- create a "solution-rich" space by diversifying the requirements
- "harvest" interesting organisms from a free-range menagerie

The Self-Made Puzzle

1. Self-Assembly of Pre-Patterned Components
2. Pattern Formation in Pre-Assembled Media
3. Integrating Self-Assembly and Pattern Formation Under Genetic Regulation
4. Toward Evolutionary Meta-Design