Université de Toulouse

Memorial University of Newfoundland

Introduction to Gene Regulatory Networks

Sylvain Cussat-Blanc
Wolfgang Banzhaf

GECCO'15 Companion, July 11-15, 2015, Madrid, Spain
ACM 978-1-4503-3488-4/15/07.
http://dx.doi.org/10.1145/2739482.2756586

GECCO

Outline

Gene regulation in biology
- Introduction to gene regulation
- Protein and genomic perspective
- Gene regulation in developmental biology

Artificial gene regulatory networks
- Biological models
- Computational models

Applications of gene regulatory networks
- Evo-devo
- Agent control
- Programming

Gene Regulation in Biology

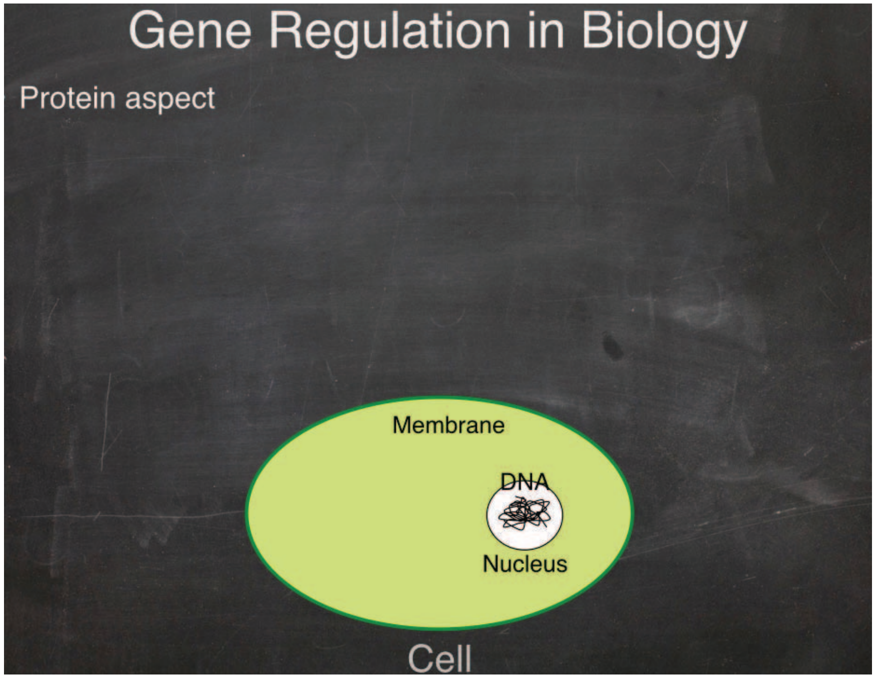
Gene Regulation in Biology

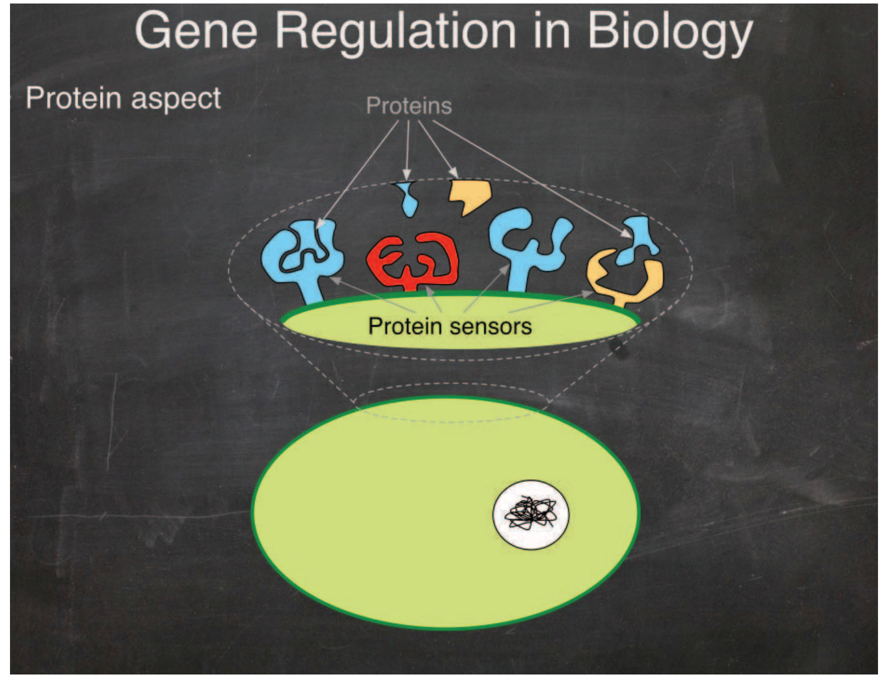
A gene regulatory network is a set of DNA segments which governs gene expression in cells

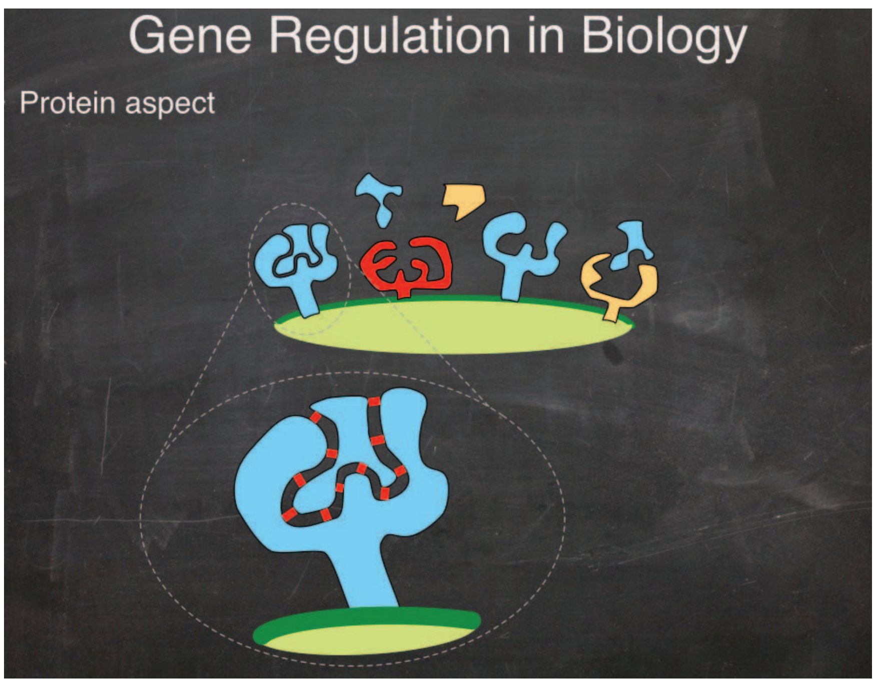
The gene expression codes for levels of mRNA of e.g., structural proteins, enzymes, or other proteins (like transcriptio factors, etc.)

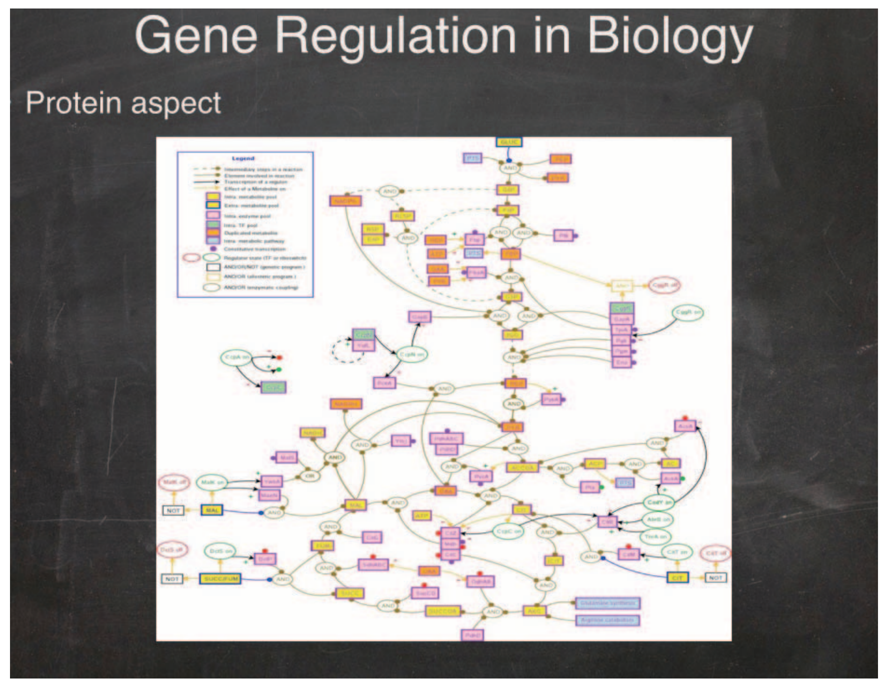
Transcription factors enhance or inhibit the production of other proteins by the cells.

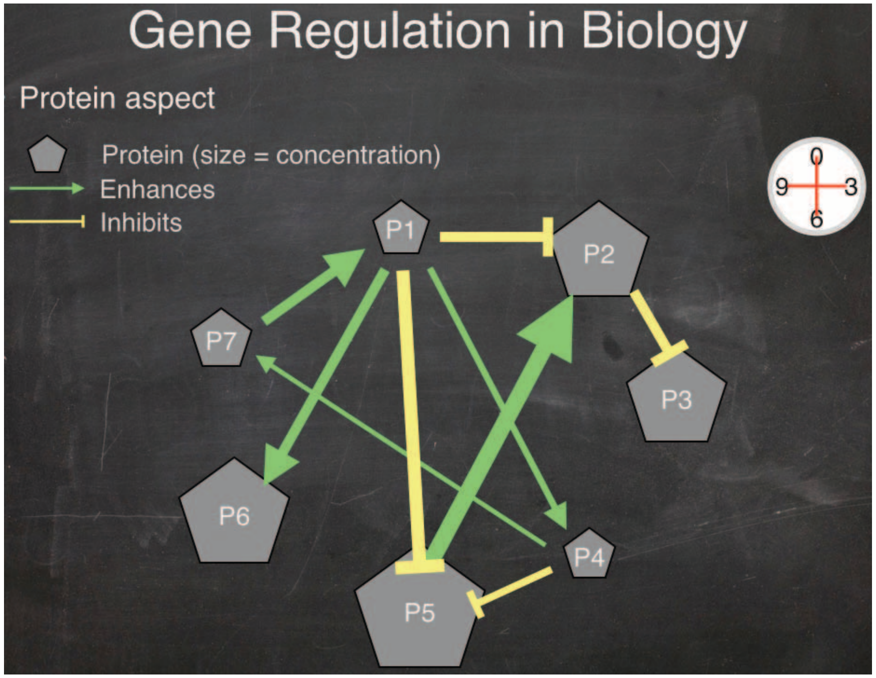
Gene regulation:
- provides the behavior of the cells (reaction to environmental conditions)
- allows for specialisation in multicellular organism by turning on an off some part of the genomes

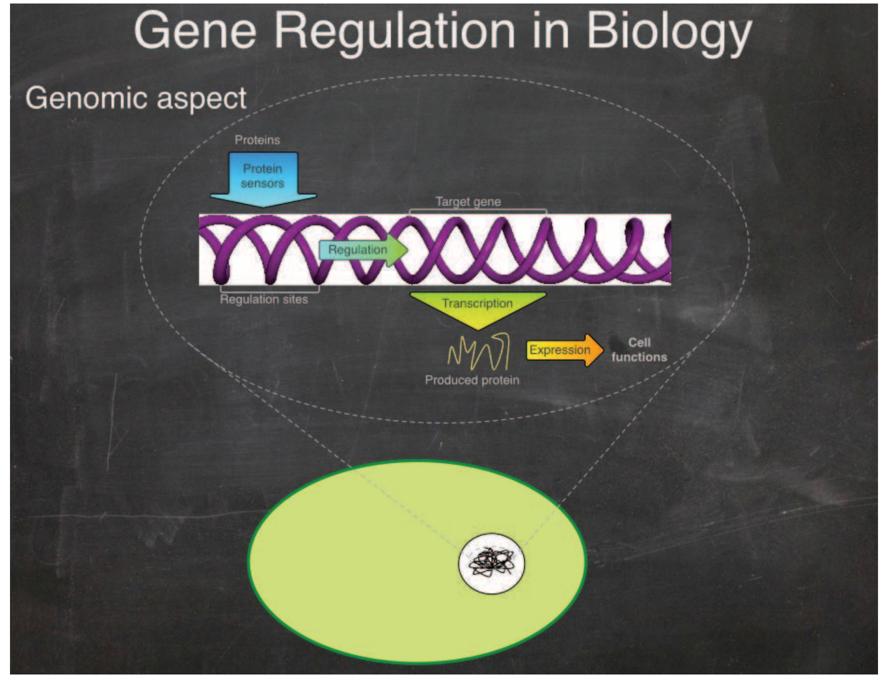

Gene Regulation in Biology
Protein aspect
Membrane
DNA
Nucleus
Cell

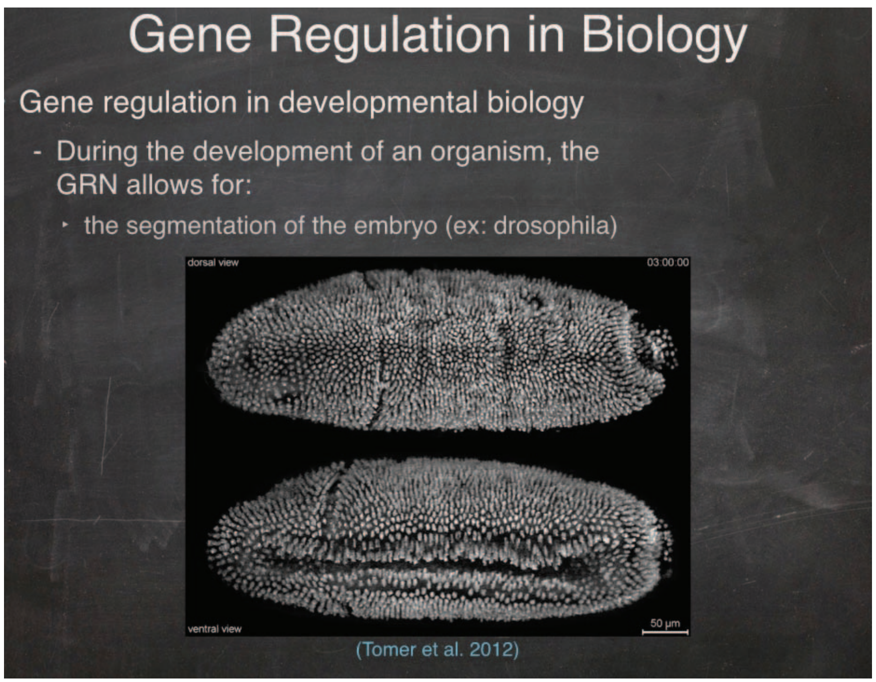

Gene Regulation in Biology
Protein aspect
Proteins
Protein sensors

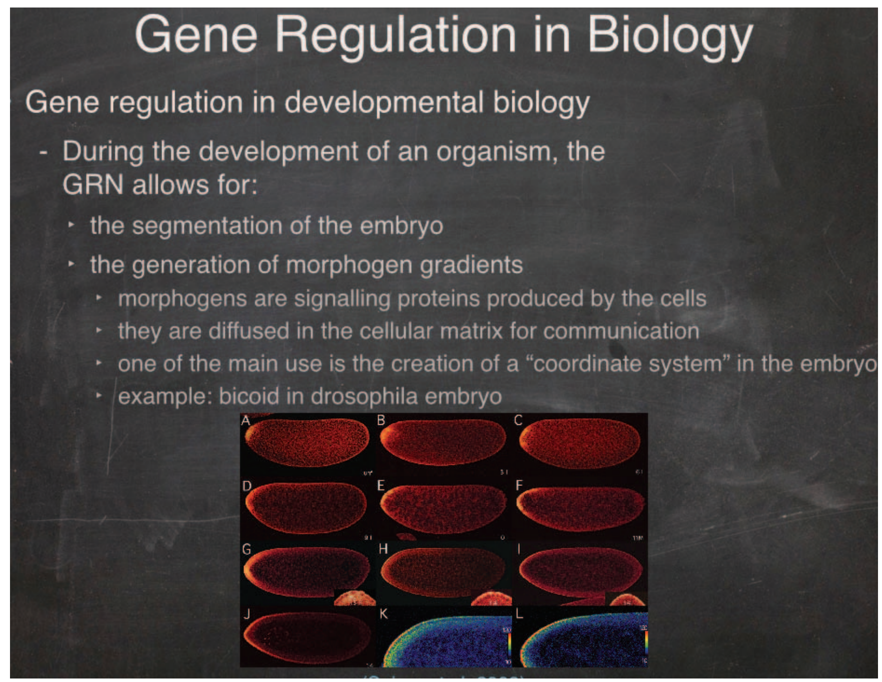

Gene Regulation in Biology
Protein aspect


Gene Regulation in Biology
Protein aspect


Gene Regulation in Biology
Protein aspect
Protein (size = concentration)
Enhances
Inhibits
P1
P2
P3
P4
P5
P6
P7
0
3
6
9


Gene Regulation in Biology
Genomic aspect
Proteins
Protein sensors
Target gene
Regulation
Regulation sites
Transcription
Expression
Cell functions
Produced protein


Gene Regulation in Biology
Gene regulation in developmental biology
- During the development of an organism, the GRN allows for:
‣ the segmentation of the embryo (ex: drosophila)
dorsal view
03:00:00
ventral view
50 µm
(Tomer et al. 2012)

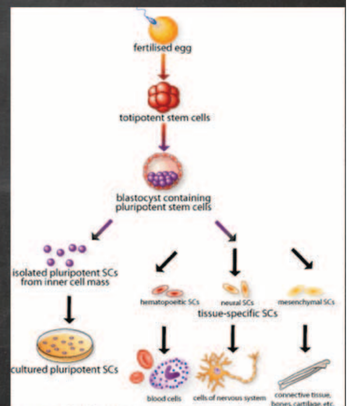

Gene Regulation in Biology
Gene regulation in developmental biology
- During the development of an organism, the GRN allows for:
‣ the segmentation of the embryo
‣ the generation of morphogen gradients
‣ morphogens are signalling proteins produced by the cells
‣ they are diffused in the cellular matrix for communication
‣ one of the main use is the creation of a “coordinate system” in the embryo
‣ example: bicoid in drosophila embryo
A
B
C
D
E
F
G
H
I
J
K
L

Gene Regulation in Biology

Gene regulation in developmental biology

- During the development of an organism, the GRN allows for:
  - the segmentation of the embryo
  - the generation of morphogen gradients
  - the differentiation of the cells into different cell types

  => 1 DNA for multiple cell functions



Source: http://www.scq.ubc.ca/stem-cell-bioengineering/

Artificial Gene Regulatory Networks

Artificial Gene Regulatory Networks

Models of gene regulation designed for:

- biological purposes
  - simulation of real GRNs
  - interactions between the protein
  - dynamics of the network
  - implication in the developmental process
  - implication in the regulation of the cell life cycle
- computational purposes
  - inspiration from the biology
  - identical structure based on interaction between proteins
  - artificial evolution of the proteins
  - generally used to control agents (cells, robots, etc.)

Artificial Gene Regulatory Networks

Biological models

- ODEs
  - Representation of gene regulation with ordinary differential equations, based on chemistry and enzymatic kinematics
- Boolean networks
  - Genes, inputs and outputs of the networks are Boolean nodes of the network
  - Edges are Boolean transition functions
- Stochastic gene networks
  - Gene expression in cells is not deterministic
    => Use of probabilistic models

Artificial Gene Regulatory Networks

Computational models

- Models used to control agents
- Agents can be:
  - Cells in evo-devo models
  - Virtual or real robots
  - etc.
- 2 groups of models
  - Biologically plausible networks
    - Bit-string representation
  - “Object-oriented” networks
    - Networks of proteins

Artificial Gene Regulatory Networks

Bit-string models

Biologically plausible model of regulation

- Encoding close to biology:
  - bit string ≈ string of nucleotides
  - Use of promoters to separate genes
- Dynamics equations close to real gene regulatory networks

But not efficient for computational purposes:

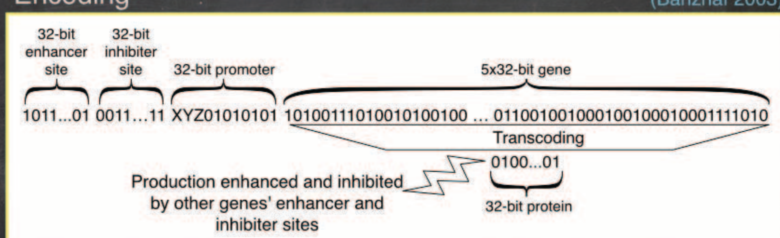
- Hard to evolve (junk DNA)

Artificial Gene Regulatory Networks

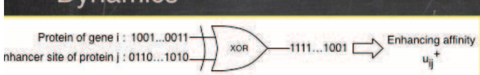
Computational models: bit-strings

- Encoding

(Banzhaf 2003)

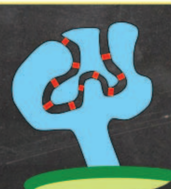


- Dynamics



$$e_i = \frac{1}{N}\sum_j^N c_j e^{\beta u^+_{ij} - u^+_{max}} \qquad h_i = \frac{1}{N}\sum_j^N c_j e^{\beta u^-_{ij} - u^-_{max}}$$

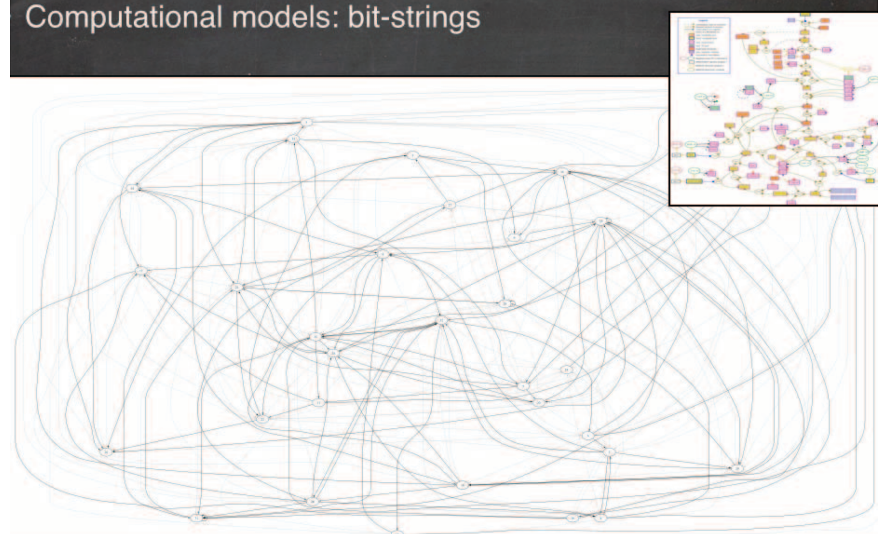
$$\frac{dc_i}{dt} = \frac{\delta(e_i - h_i)}{\Phi}$$



Matching between proteins and sensors

Artificial Gene Regulatory Networks

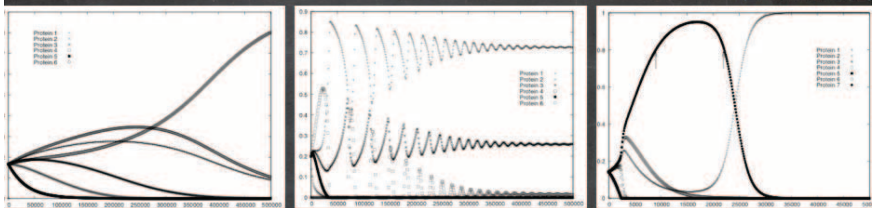
Computational models: bit-strings



Artificial Gene Regulatory Networks

Computational models: bit-strings

- Random DNA strings produce 3 types of behaviors (Banzhaf 2003)



Stable Oscillatory Transitory

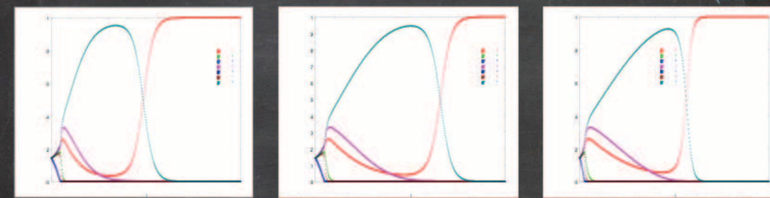
- By adding input and output proteins, possibility to control a cart pole

(Nicolau et al. 2010)

Artificial Gene Regulatory Networks

Computational models: bit-strings

- Variation of dynamical behaviors is smooth (Banzhaf 2003)



Original 1-bit mutation A further 1-bit mutation

- Both mutations were applied to the regulatory site of proteins
- Heterochrony (shift in expression timing)

Artificial Gene Regulatory Networks

Computational models: “Object-oriented” models

- Higher level of representation
- Direct encoding of proteins and the affinities between proteins
  - No promoter
  - No junk DNA
  - Easier to evolve
- Inputs and outputs can be easily represented
  - “Plug-and-play” to any agent-based problem

Artificial Gene Regulatory Networks

Computational models: “Object-oriented” models

GRN = network of proteins

- Nodes = Proteins ; characterized by:
  - identifier tag *id* ∈ [0, p]
  - Protein type (input, output ou regulatory)
  - Enhancing tag *enh* ∈ [0, p]
  - Inhibiting tag *inh* ∈ [0, p]
- Edges = matching between proteins

$$u^+_{ab} = p - |id_a - enh_b| \qquad u^-_{ab} = p - |inh_a - id_b|$$

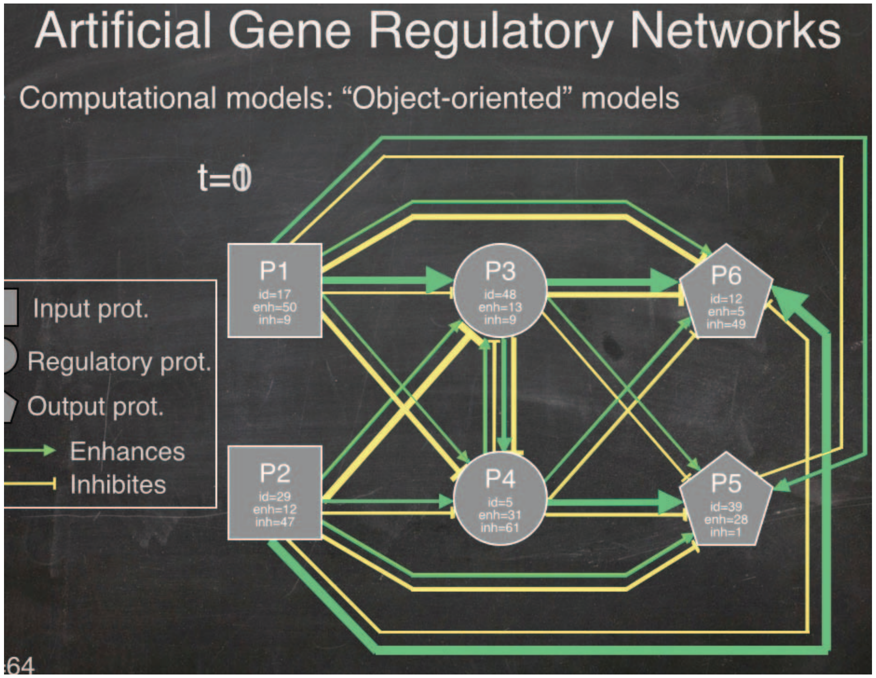
Dynamics of the network

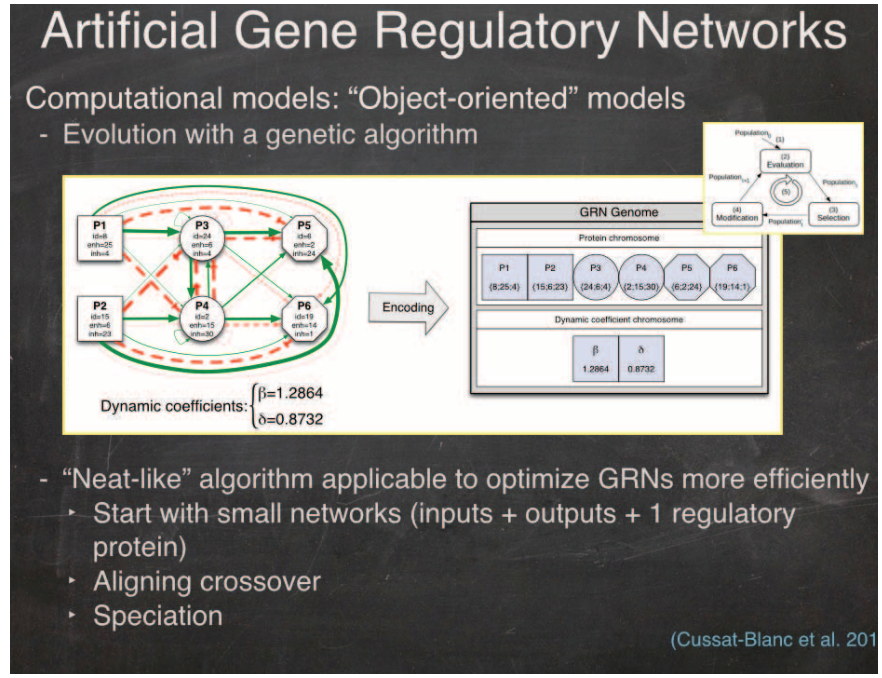
- Enhancing and inhibiting coefficients of protein i:

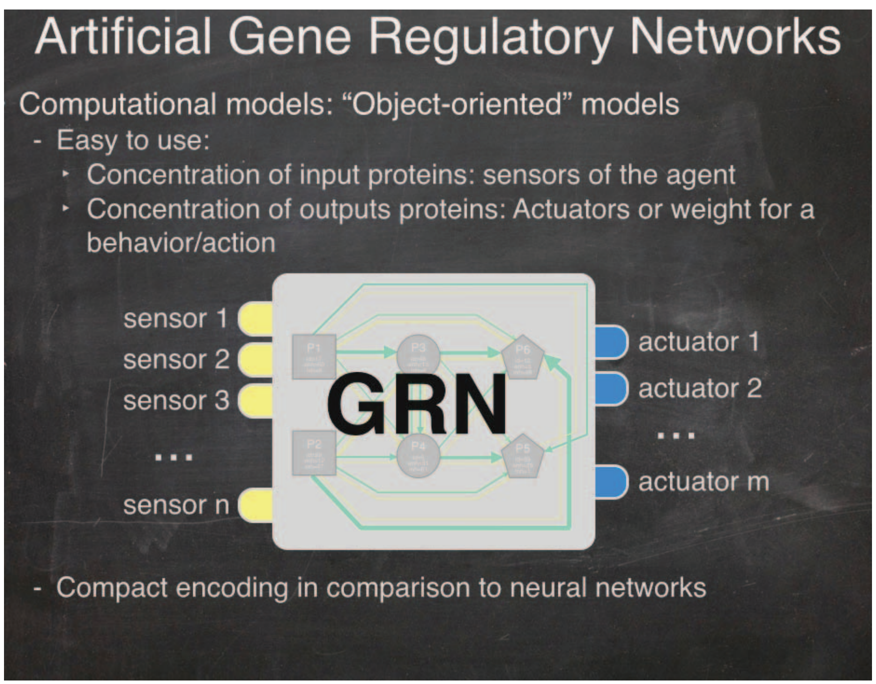
$$e_i = \frac{1}{N}\sum_j^N c_j e^{\beta u^+_{ij} - u^+_{max}} \qquad h_i = \frac{1}{N}\sum_j^N c_j e^{\beta u^-_{ij} - u^-_{max}}$$

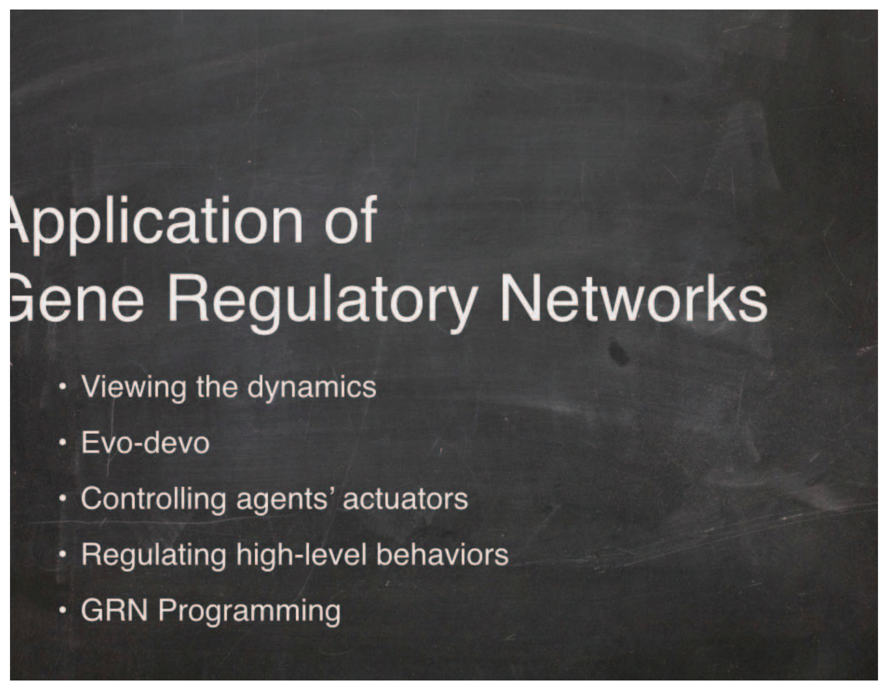
- Differential evolution of the concentration of protein i:

$$\frac{dc_i}{dt} = \frac{\delta(e_i - h_i)}{\Phi}$$


Artificial Gene Regulatory Networks
Computational models: "Object-oriented" models
t=0
Input prot.
Regulatory prot.
Output prot.
Enhances
Inhibites
P1
id=17
enh=50
inh=9
P2
id=29
enh=12
inh=47
P3
id=48
enh=13
inh=9
P4
id=5
enh=31
inh=61
P5
id=39
enh=28
inh=1
P6
id=12
enh=5
inh=49
64

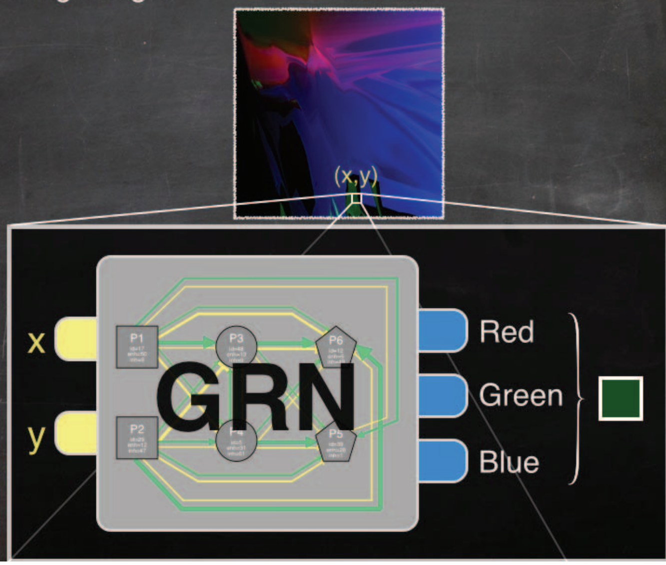

Artificial Gene Regulatory Networks
Computational models: "Object-oriented" models
- Evolution with a genetic algorithm
GRN Genome
Protein chromosome
Dynamic coefficient chromosome
Encoding
Dynamic coefficients: β=1.2864 δ=0.8732
- "Neat-like" algorithm applicable to optimize GRNs more efficiently
Start with small networks (inputs + outputs + 1 regulatory protein)
Aligning crossover
Speciation
(Cussat-Blanc et al. 201


Artificial Gene Regulatory Networks
Computational models: "Object-oriented" models
- Easy to use:
Concentration of input proteins: sensors of the agent
Concentration of outputs proteins: Actuators or weight for a behavior/action
sensor 1
sensor 2
sensor 3
...
sensor n
GRN
actuator 1
actuator 2
...
actuator m
- Compact encoding in comparison to neural networks


Application of
Gene Regulatory Networks
Viewing the dynamics
Evo-devo
Controlling agents' actuators
Regulating high-level behaviors
GRN Programming

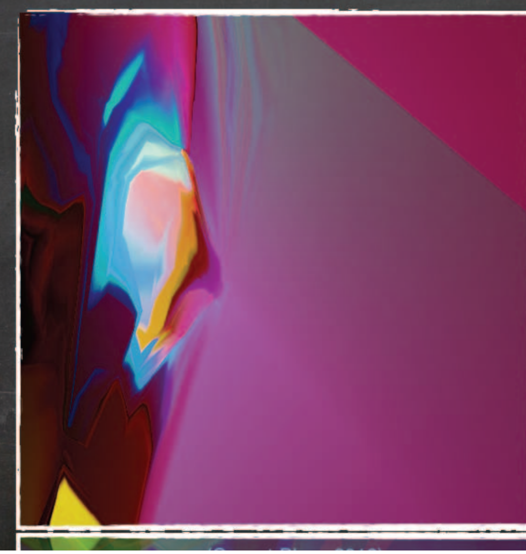
Application of GRNs

Generating images



Application of GRNs

Generating images



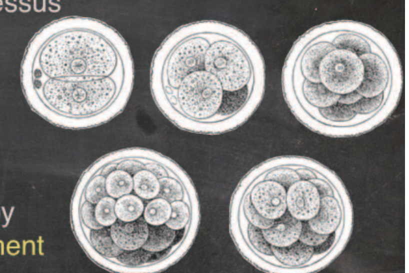
Application of GRNs

Generating videos

Application of GRNs

Evo-Devo: Evolution and Development

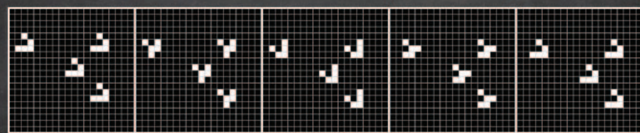
- Definition: grow virtual organism with evolved GRNs
- Embryogenesis: Formation Processus of multicellular organisms from the egg cell to the autonomous living being
- Artificial embryogenesis: Generation of virtual organisms by taking inspiration of the development process of living beings
  => based on biological concepts



Application of GRNs

Evo-Devo: cellular automata

- Evolution of a matrix with simple rules based on the neighborhood
- The state of the cells at time *t* determines their state at time *t+1*
- Example: Conway's game of life



(Conway 1970)

- Shapes can be generated by evolving the rules

$$fitness = \frac{Nb_{in} - \frac{1}{2}Nb_{out}}{des}$$

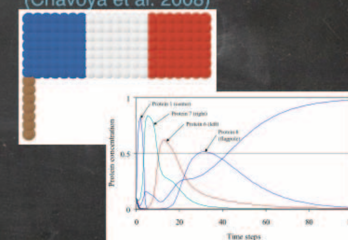
(De Garis 1999)

Application of GRNs

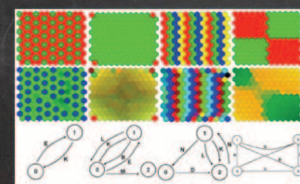
Evo-Devo: use of GRNs

- Development of specialization patterns with a cellular automata
- French flag problem
  - Different cell specialization, depending on the position of the cells
  - Morphogen gradient pre-positioned or produced by the organisms

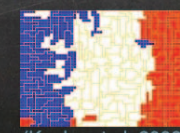
(Chavoya et al. 2008)



(Flann et al. 2005)



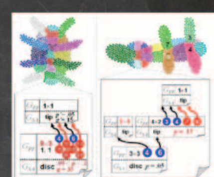
(Joachimczak et al. 2008)



Application of GRNs

Evo-Devo: use of GRNs

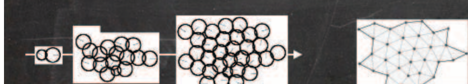
- Complexification of the shape by introducing a developmental process
  - Probability of division for each cell
  - Intercellular adhesion (mass/spring system) to aggregate the cells



(Doursat 2009)

- Development and control of the morphology
  - Morphology produce by an Evo-Devo process
  - Once stable, the morphology is moved by a GRN in a fluidic environment

(Joachimczak et al. 2012)



Growth

Final shape

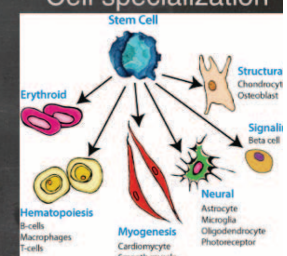
Motion

Application of GRNs

Evo-Devo: Generating artificial creatures

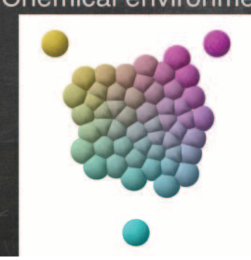
- GRN used to regulate the cell's life cycle
- Optimization fo the GRN to produce creatures adapted to their environment
- Functional differentiation of the cells
- Simulated physics and chemistry in the environment

Cell specialization



Physical environment

Chemical environment



Application of GRNs

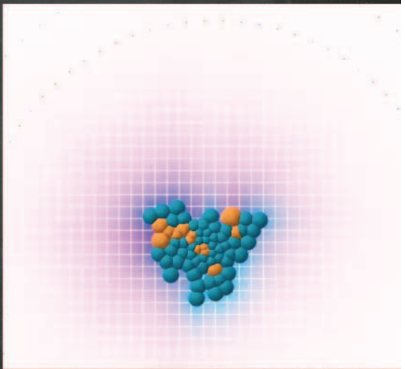
Evo-Devo: Generating artificial creatures

- In each cell, the GRN manages:
  - High-level actions (division, apoptosis, quiescence, differentiation)
  - The chemical production of the cells (morphogens, nutrients, energy)
- For regulation, GRN uses:
  - information on the cell's internal state (energy, stock of nutrient, etc.)
  - information on its local environment local (morphogens)
- Optimization of GRN to generate complex multicellular organisms
  - Possible optimization objectives (fitness):
    - Generate colored shapes (color = differentiation state)
    - Generate user-defined functions (harvest a protein, move to a point, etc.)
    - SURVIVE
- What does it mean to survive?
  - Have at least 1 living cell in the organism
    => being able to adapt to environmental conditions
    => the complexity is not in the fitness function anymore but move to the definition of the environment

Application of GRNs

Evo-Devo: Generating artificial creatures

- Example:
  - 2 types of cells:
    - Nutritive => Extract energy from the environment
    - Defensive => Resist to external agressions

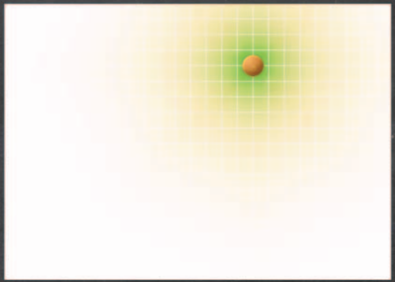


- Environment :
  - Contains nutrients
  - Contains nocive particules that kills nutritive cells
- Fitness:
  - Survive == simulation duration

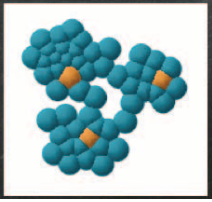
Application of GRNs

Evo-Devo: Generating artificial creatures

- At the beginning of the evolution, very simple strategy:



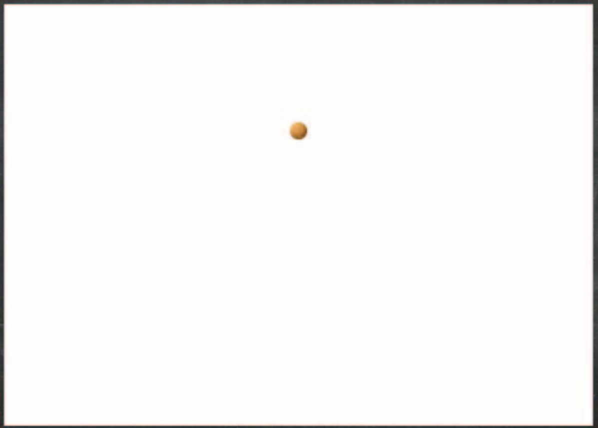
- Then, progressive complexification



Application of GRNs

Evo-Devo: Generating artificial creatures

- Finally, motion emerges (with no cell migration!)



(Disset et al. 2014)

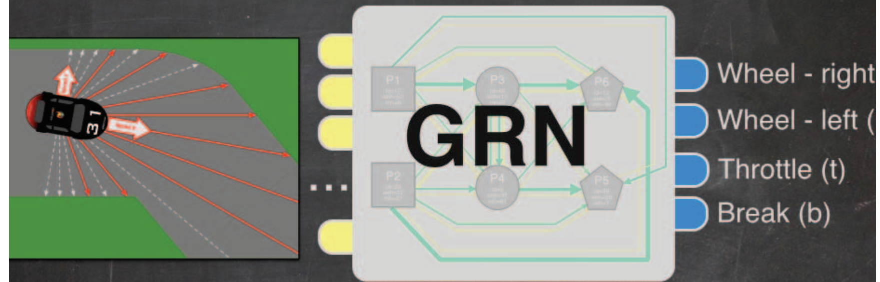
Application of GRNs

Simulated car racing

- Direct connection of the GRN to a virtual car
- Use of TORCS for the car physics
- Available sensors:
  - 18 track sensors
  - 3 speed sensors
  - 8 state sensors (rpm, position, time, etc.)
- Available actuators:
  - Wheel angle ∊ [-1, 1]
  - Throttle ∊ [0, 1]
  - Break ∊ [0, 1]
- Only keep the necessary sensors

Application of GRNs

Simulated car racing



Inputs must be normalized

Final actuator values given by:

- Wheel angle = (r-l)/(r+l)
- Throttle = max(0, (t-b)/(t+b))
- Break = max(0, (b-t)/(b+t))

Application of GRNs

Simulated car racing

- How to teach the GRN to drive? What is a good fitness?
- Solution: incremental evolution
  1. Teach to drive on 1 simple track
  2. Generalize the behavior on other tracks
  3. Polish the behavior of the network

(Sanchez et al. 2014)

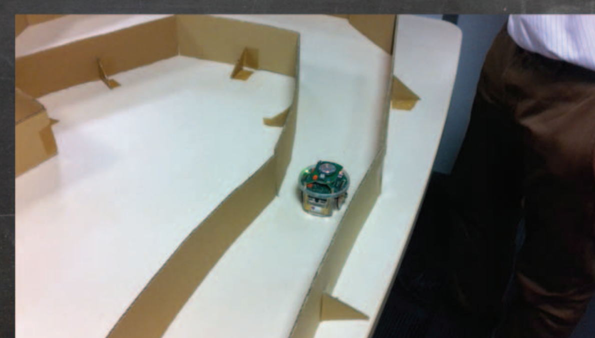
Application of GRNs

Features of system

- Easy to use
- Adaptative to the track and its surface
- Robust resistant to the noise

Example of robustness

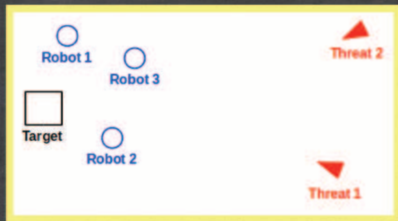
- Use of the controller trained in the game to drive a robot



Application of GRNs

Regulating high-level behaviors

- Example: protect a target against enemies
  - Decentralized control of a group of defenders
  - Multiple threats incoming from the environment



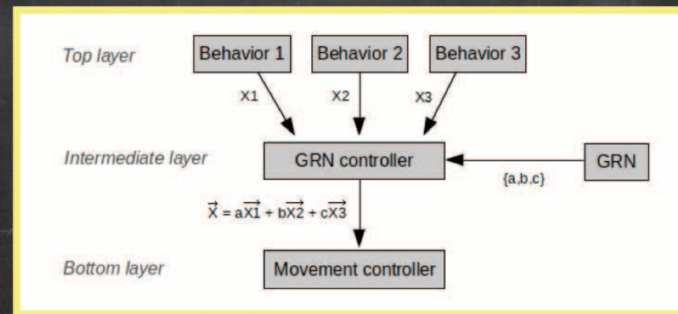
Use of a GRN to regulate high level behavior such as:

- Defence
- Attack
- Scatter
- Regroup

Application of GRNs

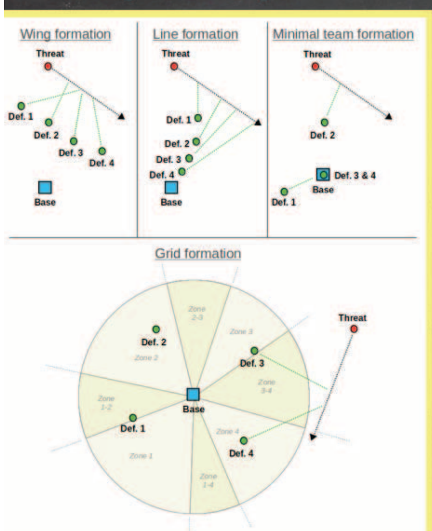
Regulating high-level behaviors

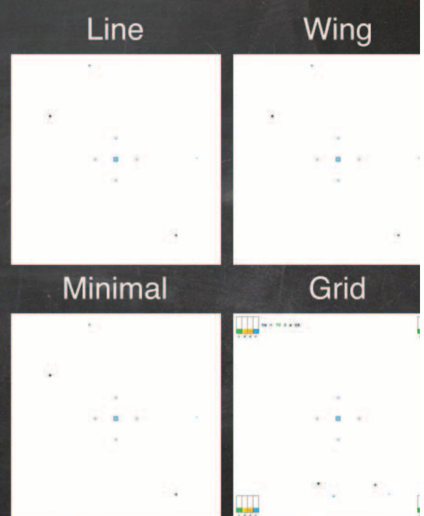
- Hybrid and modular architecture
- Behaviors generate direction vector
- The GRN aggregate the vectors to produce the final move



Application of GRNs

4 strategies emerge:



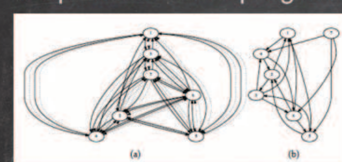


Application of GRNs

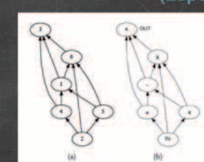
GRN programming

- Map network into program

(Lopes and Costa 201



Original Simplification



Mapping into executable graph

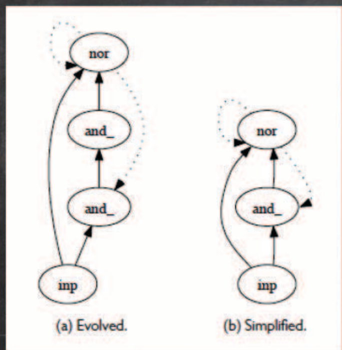
- Here, the static network connectivity is used to determine program structure, including feedback connections
- Input and output nodes are fixed
- Input data stream is applied and output data stream read off.

Application of GRNs

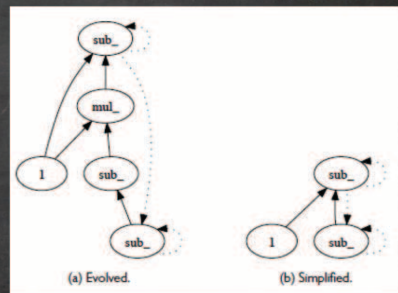
GRN programming

- Output at top node

(Lopes 2015)



Even n-bit Parity



Fibonacci Series

Gene Regulation - Pros and Cons

Pros:

- Plug-and-play
- Temporal aspect: Many tasks are dynamic
- Natively continuous
- Many different behaviors possible
- Close to natural systems

Cons:

- Can be difficult to evolve
- Black-box system

References

omer et al. 2012) Tomer, Raju, et al. "Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet icroscopy." *Nature methods* 9.7 (2012): 755-763.

pirov et al. 2009) Spirov, Alexander, et al. "Formation of the bicoid morphogen gradient: an mRNA gradient dictates the protein gradient." evelopment 136.4 (2009): 605-614.

anzhaf 2003) Banzhaf, Wolfgang. "Artificial regulatory networks and genetic programming." *Genetic programming theory and practice.* pringer US, 2003. 43-61.

licolau et al. 2010) Nicolau, Miguel, Marc Schoenauer, and Wolfgang Banzhaf. "Evolving genes to balance a pole." *Genetic Programming.* pringer Berlin Heidelberg, 2010. 196-207.

ussat-Blanc et al. 2015) Cussat-Blanc, Sylvain, Kyle Harrington, and Jordan Pollack. "Gene Regulatory Network Evolution Through ugmenting Topologies." *IEEE Transaction on Evolutionary Computation.* IEEE, 2015 (to be published)

ussat-Blanc 2012) Cussat-Blanc, Sylvain, and Jordan Pollack. "Using pictures to visualize the complexity of gene regulatory networks." *rtificial Life.* Vol. 13. 2012.

onway 1970) Conway, John. "The game of life." *Scientific American* 223.4 (1970): 4.

e Garis 1999) De Garis, Hugo. "Artificial embryology and cellular differentiation." *Evolutionary design by computers* (1999): 281-295.

havoya et al. 2008) Chavoya, Arturo, and Yves Duthen. "A cell pattern generation model based on an extended artificial regulatory network." *iosystems* 94.1 (2008): 95-101.

lann et al. 2005) Flann, Nicholas, et al. "Biological development of cell patterns: Characterizing the space of cell chemistry genetic regulator etworks." *Advances in Artificial Life.* Springer Berlin Heidelberg, 2005. 57-66.

oachimczak et al. 2008) Joachimczak, Michal, and Borys Wróbel. "Evo-devo in silico-a Model of a Gene Network Regulating Multicellular evelopment in 3D Space with Artificial Physics." *ALIFE.* 2008.

nabe et al. 2008) Knabe, Johannes F., Chrystopher L. Nehaniv, and Maria J. Schilstra. "Do motifs reflect evolved function?—No convergent volution of genetic regulatory network subgraph topologies." *Biosystems* 94.1 (2008): 68-74.

isset et al. 2014) Disset, Jean, Sylvain Cussat-Blanc, and Yves Duthen. "Toward organogenesis of artificial creatures." *Proceedings of the 014 conference companion on Genetic and evolutionary computation companion.* ACM, 2014.

oursat 2009) Doursat, René. "Facilitating evolutionary innovation by developmental modularity and variability." *Proceedings of the 11th nnual conference on Genetic and evolutionary computation.* ACM, 2009.

oachimczak et al. 2012) Joachimczak, Michał, and Borys Wróbel. "Co-evolution of morphology and control of soft-bodied multicellular nimats." *Proceedings of the 14th annual conference on Genetic and evolutionary computation.* ACM, 2012.

anchez et al. 2014) Sanchez, Stéphane, and Sylvain Cussat-Blanc. "Gene regulated car driving: using a gene regulatory network to drive a irtual car." *Genetic Programming and Evolvable Machines* 15.4 (2014): 477-511.

elecluse 2013) Martin Delecluse. "Autonomous and Robust Control Architecture for Multi-Robot System." Master thesis (2013).

opes and Costa 2012) Lopes, Rui L., and Ernesto Costa. "The regulatory network computational device." *Genetic Programming and volvable Machines* 13.3 (2012): 339-375.

opes 2015) Lopes, Rui L.. "A computational model inspired by gene regulatory networks." PhD Thesis, University of Coimbra, Portugal, 2015