A Framework for Modelling and Simulating Networks of Cells

Sara Montagna Mirko Viroli

{sara.montagna, mirko.viroli}@unibo.it

Alma Mater Studiorum—Università di Bologna a Cesena

1st International Workshop
on Interactions between Computer Science and Biology
(CS2Bio'10)

Amsterdam, Netherlands, 10th June 2010

Outline

1. The morphogenesis of living systems
   - the crucial role of hierarchical organisation
2. Requirement
   - multilevel large-scale tool
3. A new framework
   - computational model
   - specification language
   - simulation engine
4. First results
   - the analysis of *Drosophila Melanogaster* regionalisation as a case study
5. Conclusion and future works

Outline

Biological Background

Developmental Biology researches the mechanisms of development, differentiation, and growth in animals and plants at the molecular, cellular, and genetic levels.

Animal developmental steps

1. Fertilisation of one egg
2. Mitotic division
3. Cellular differentiation
   - diverse gene expression
4. Morphogenesis
   - control of the organised spatial distribution of cell diversity

Each region of the developing organism expresses a given set of genes

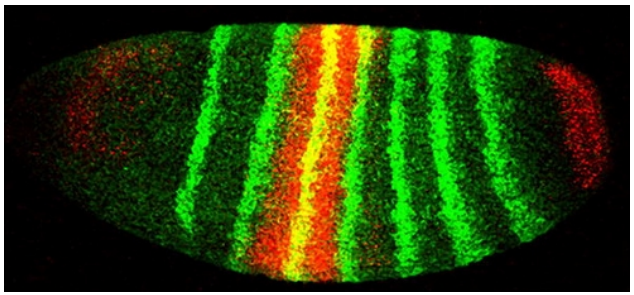


Figure: *Drosophila M.* segments

const size(20).

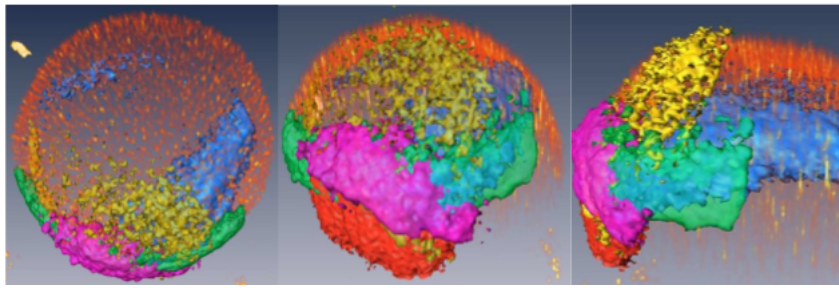


Figure: *Zebrafish* regionalisation

- Developmental Biology recognises as important actors in the emergence of embryonic patterning – self-organised structures
  - transcriptional control mechanisms
  - signalling pathways
  - cell-to-cell direct interaction
  - short and long range signals (*morphogenes*)

$\longrightarrow$ interplay between cell internal activity and cell-to-cell interactions

Figure by:

[1] taken from the web [2] An Automatic Quantification and Registration Strategy to Create a Genetic Expression Atlas of Zebrafish Embryogenesis. C. Castro et all. Accepted at IEEE Engineering in Medicine and Biology Society (EMBC'09)

On the Need of Proper Tools

Tool requirements

1. Multi-compartment / multi-scale model
   - for reproducing the interactions of system components at cellular and intracellular level
2. Diffusion / Transfer
   - for studying the effects of short and long range signals
   - for modelling the compartment membrane
3. Stochasticity
   - for capturing the aleatory behaviour peculiar of biological systems (in particular when they involve few copies of the reacting molecules)

Brief Survey on Existing Frameworks

Systems Biology Tools: See http://systems-biology.org/software/

- Based on classical mathematical models (ODE or PDE or stochastic solvers —the *Gibson and Bruck* method or the *Tau-Leap.* one)
  - CellDesigner, CellWare, COPASI, Dizzy, JDesigner, Virtual Cell, ...
  - mainly thought for intracellular aspects (biochemical pathways)
  - not practical for modelling large set of cells

Computational Systems Biology Tools

- Based on computational models (stochastic process-algebras, Petri-Nets)
  - SPIM (stochastic π-calculus), BlenX, Bio-PEPA
  - ground on Gillespie's characterisation of chemistry as CTMC
  - promote a view of "molecules as concurrent processes"
  - recent preliminary extensions towards multi-compartimentalisation

A flexible and pragmatic framework is needed to support large scale multi-compartimentalisation

Outline

An Ad-hoc Framework

Motivations

Tackle scenarios of developmental biology

- Naturally supporting systems with many compartments
- Use state-of-the-art implem. techniques for the simulation engine

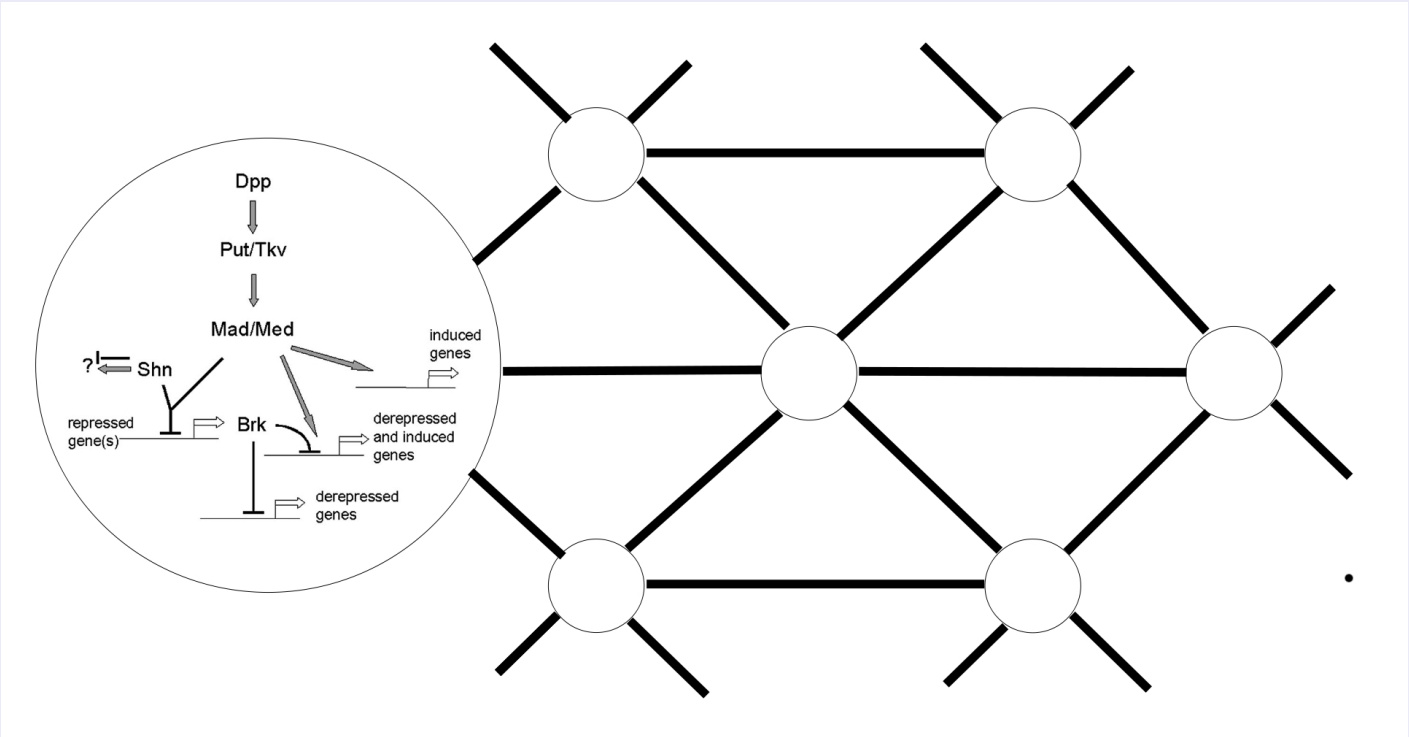
Framework's Conceptual levels

1. Computational Model: graph of compartments, with transfer reactions
2. Surface Language: systems as logic-oriented description programs
   - system structure
   - inner chemical behaviours
3. Simulation Engine: known $O(N * logN)$ version of Gillespie's SSA
   - reproducing the exact chemical evolution/diffusion of substances

Level 1: the Computational Model – Structure

A multi-compartment version of standard CMSB view

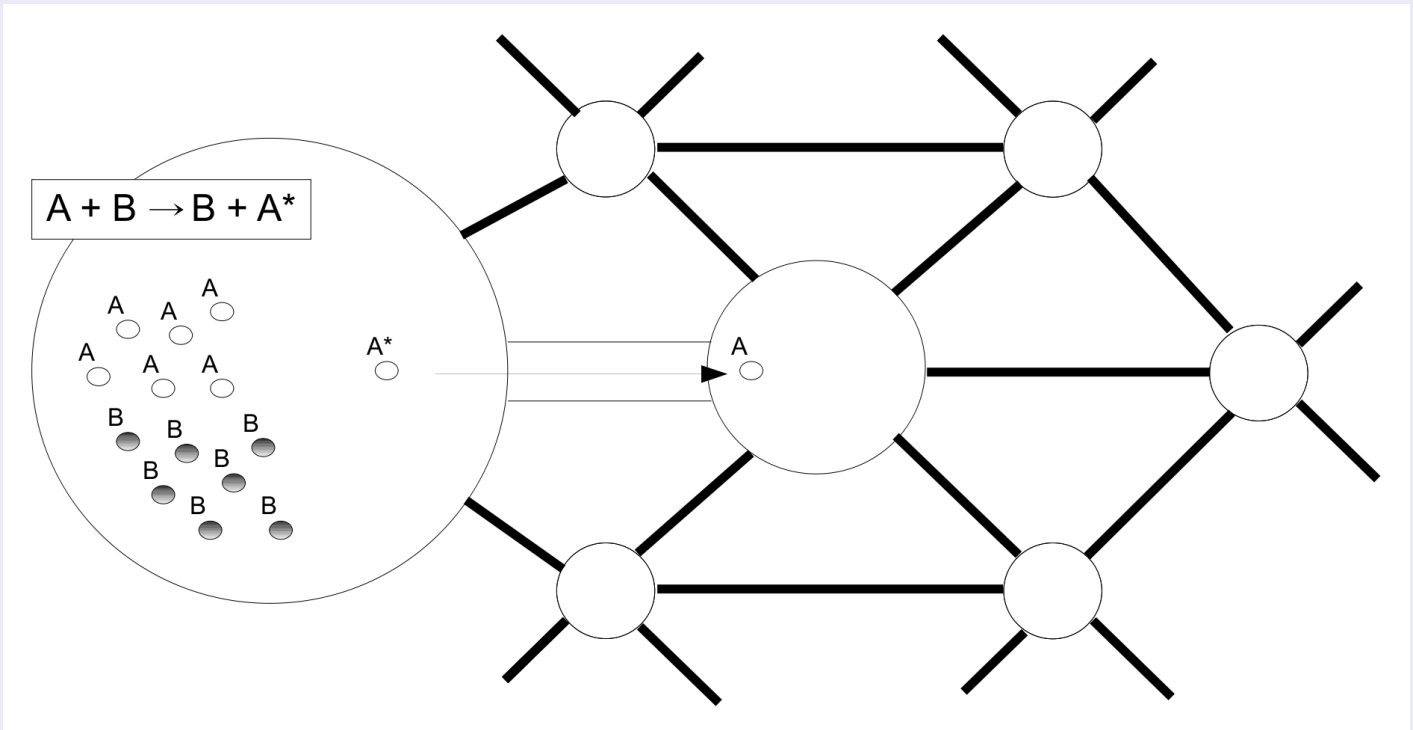
- A system is a graph-like network of compartments
- Each compartment hosts a chemical solution
- Mobility and mitotic division will be supported in future versions



Level 1: the Computational Model – Chemical Transfer

Transfer Model

- Some chemical reactions can produce so-called firing molecules
- They are sent to a neighbouring compartment picked probabilistically



Level 2: the Surface Specification Language

Structure

A program as a set of logic declarations

- Declaring molecules, reactions, compartments, links
- Configuring the simulation (initial state and parameters)
- Stating output commands

Declarations can have variables and be equipped with fully expressive preconditions, acting as constraints on declarations

- Supporting flexibility

Example Specification

Diffusion of a substance into a grid-like tissue

```
const size(20).

molecule M where (M in [pump,field]).
reaction r(pump)  : [pump] --> [pump,field] rate 10.0.
reaction r(diff)  : [field] --> [field,firing(field)] rate 0.2.
reaction r(decay) : [field] --> [] rate 0.1.

compartment c(X,Y) where (const size(N), X in {1..N}, Y in {1..N}).
link c(X,Y) >>> c(X,Y1) molecule field rate 10000.0 where ( Y1 in [Y-1,Y+1] ).
link c(X,Y) >>> c(X1,Y) molecule field rate 10000.0 where ( X1 in [X-1,X+1] ).

concentration 1 of pump into c(M,M) where (const size(N), M is N//2).
place AnyReaction into AnyCompartment.

final_steps 100000.
sample_steps 100.

out [molecule(c(X,Y),field),Delimiter] where (
    const size(N),
    inspect(compartment c(X,Y)),
    (Y=N -> Delimiter=end_of_line ; Delimiter=space)
).
```

Level 3: the Simulation Engine – Under the hood

Main Elements

- An available chemical reaction in a compartment is picked up probabilistically based on its rate
  - selected in $O(logN)$ time, via a special binary search tree
- The transition modifies a small portion of data structures
- The transition duration is drawn with exponential distribution
  - exactly modelling chemical dynamics according to [1]

Based on:

[1] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem., 81(25), 1977.
[2] M. A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. J. Phys. Chem. A, 104(9), 2000.

Level 3: the Simulation Engine – Output (1)

Produces a textual result from out commands

```
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 4 3 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 1 7 54 7 3 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 1 12 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Level 3: the Simulation Engine – Output (2)

Produces a textual result from `out` commands

```
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 1 1 3 1 2 0 0 0 0 0 0 0 0
0 0 0 0 0 0 2 4 7 9 11 6 0 0 0 0 0 0 0 0
0 0 0 0 0 0 5 6 24 29 30 9 1 0 0 0 0 0 0 0
0 0 0 0 0 1 3 8 31 85 56 19 4 0 0 0 0 0 0 0
0 0 0 0 0 1 3 20 83 166 73 21 9 1 1 0 0 0 0 0
0 0 0 0 0 1 2 9 58 90 56 14 6 1 0 0 0 0 0 0
0 0 0 0 0 0 0 2 19 27 26 10 3 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 5 7 5 1 2 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

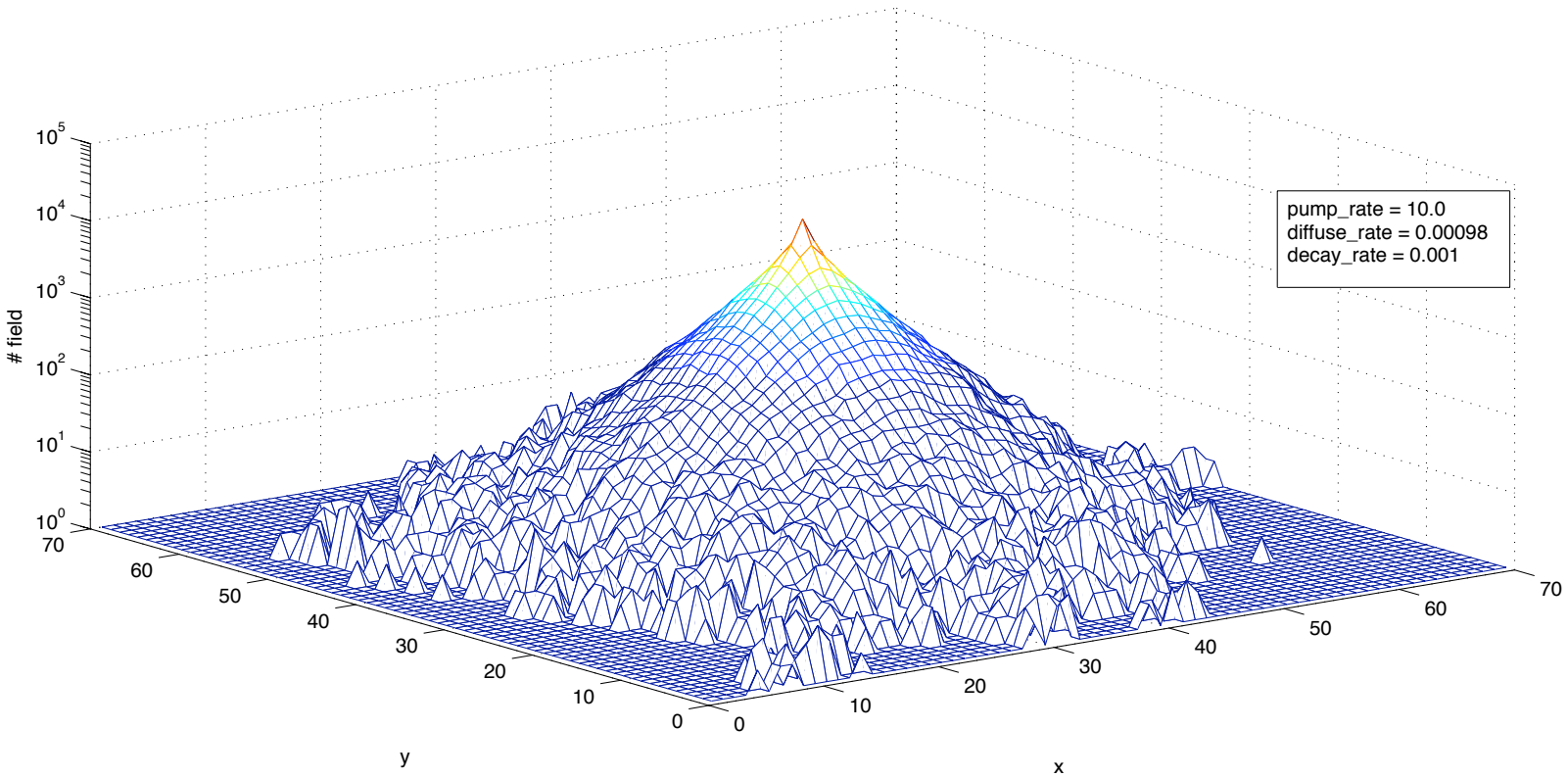
Level 3: the Simulation Engine – Output (3)

Produces a textual result from out commands

```
0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 1 0 0 1 1 1 0 0 0 0 0 0 0 0
0 0 0 0 0 0 2 1 3 6 5 2 0 1 0 0 0 0 0 0
0 0 0 0 0 2 6 11 11 14 12 2 0 1 3 2 0 0 0 0
0 0 0 1 4 8 17 36 64 57 46 20 7 5 2 1 0 0 0 0
0 0 0 3 9 15 31 77 120 142 98 47 36 16 5 1 0 0 0 0
1 1 2 6 20 28 78 162 217 238 221 145 78 43 14 6 1 0 0 0
0 1 1 12 33 83 171 261 401 497 419 265 130 52 16 3 1 0 0 0
1 5 9 20 52 93 227 399 658 779 587 381 228 84 29 5 2 0 0 0
0 0 6 16 44 100 243 460 768 966 737 495 269 98 38 10 4 4 0 0
0 0 3 11 44 106 191 385 638 803 677 431 234 94 36 11 0 0 0 0
0 0 3 2 32 51 120 251 412 466 445 338 160 61 23 16 1 0 0 0
0 0 2 1 8 27 48 117 205 289 246 171 75 39 15 11 5 1 0 0
0 0 1 2 2 11 25 46 86 134 117 64 36 14 5 5 7 1 0 0
0 0 0 0 0 2 6 24 31 37 26 23 9 4 0 0 0 0 0 0
0 0 0 0 0 0 1 4 6 10 4 7 4 2 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 3 2 1 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Level 3: the Simulation Engine – Drawing Charts

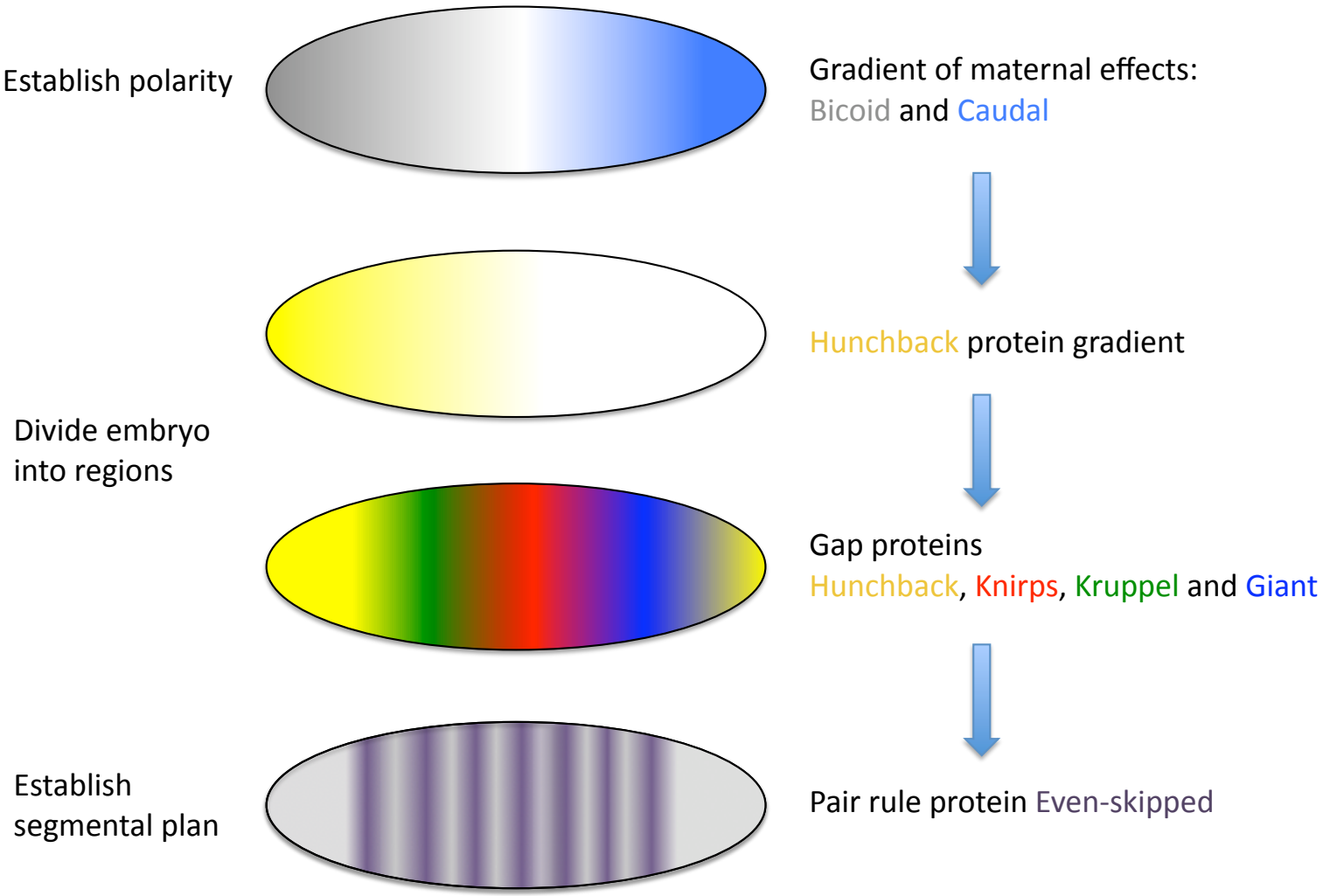
Charting using any existing tool (Matlab, gnuplot,..)



Outline

Biological Background - Gene Expression Pattern

- Egg of Drosophila already polarised by maternal effects



Goal of the Model

- Reproducing the gene expression pattern of the gap genes at Cleavage Cycle 14A - temporal class 8...
  - *hunchback* (hb), *Krüppel* (Kr), *knirps* (kni) and *giant* (gt)

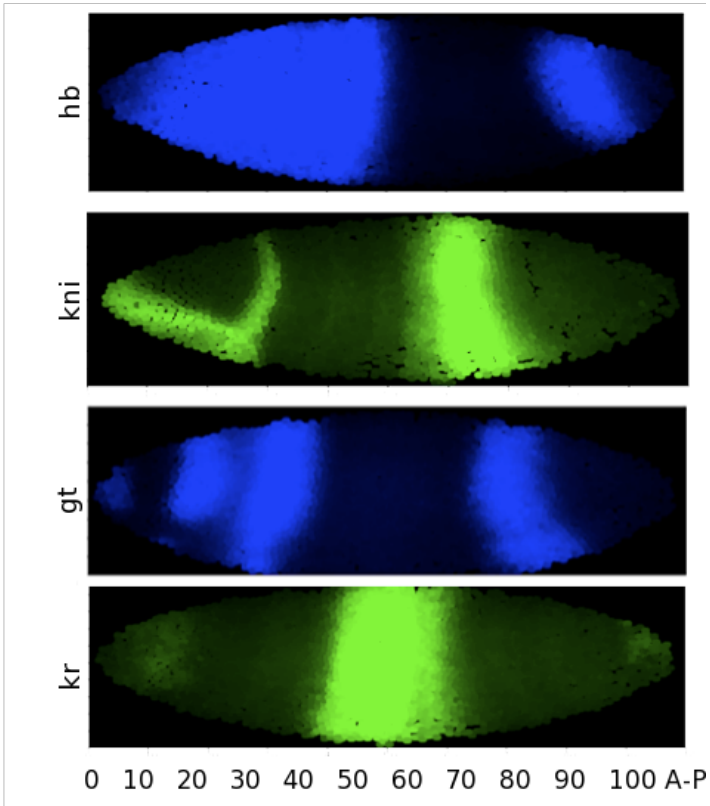


Figure: 2D data from the FlyEx database[1]

[1] http://flyex.ams.sunysb.edu/flyex/index.jsp

Initial Condition

- . . . Beginning with expression data at Cleavage Cycle 11

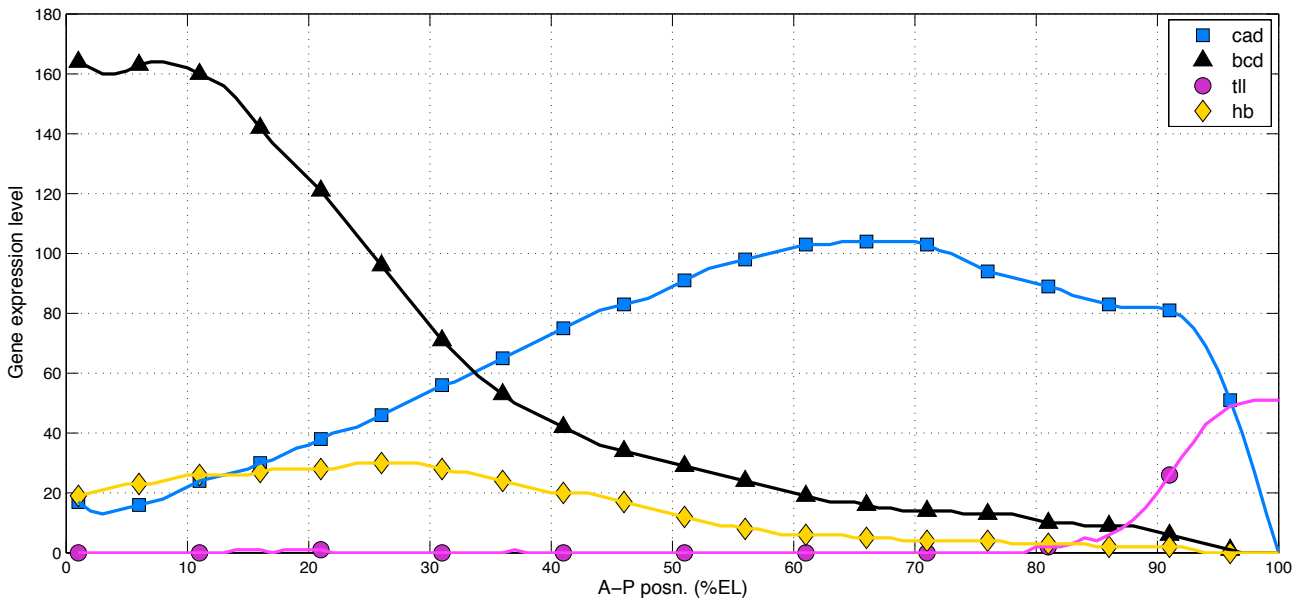


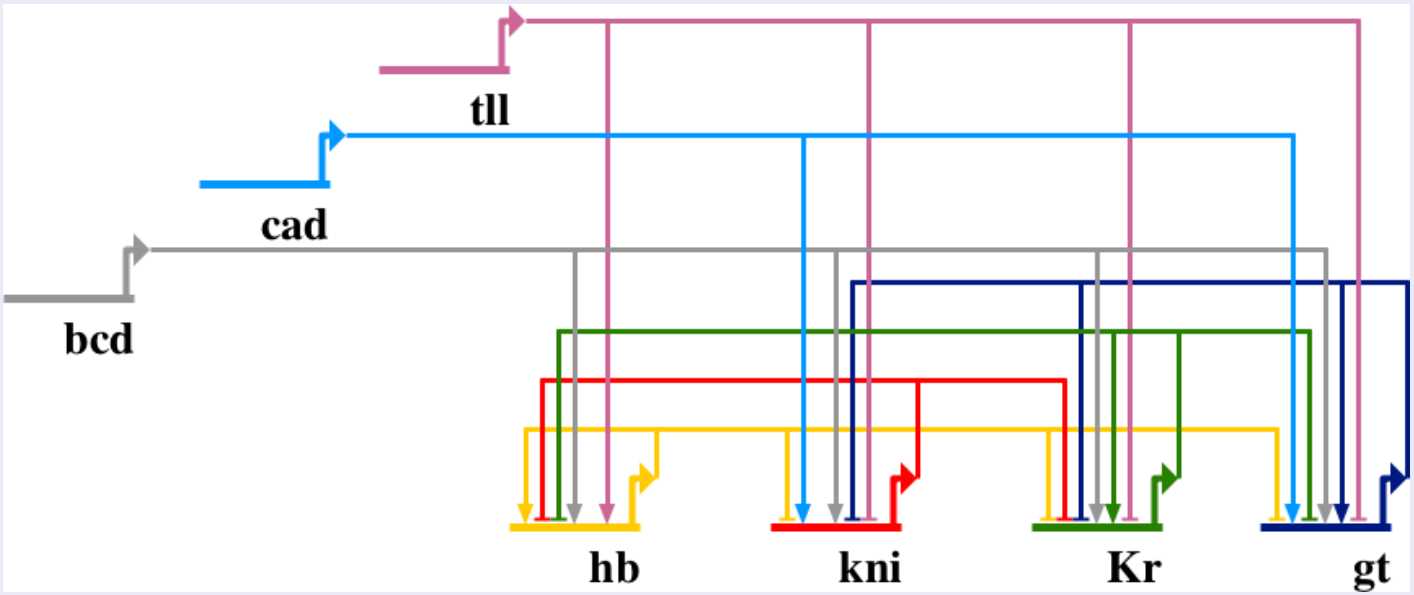
Figure: Experimental data from the FlyEx database of genes with non-zero concentration. The concentration of proteins are unitless, ranging from 0 to 255, at space point x, ranging from 0 to 100 % of embryo length.

Working Hypotheses of the Intracellular Model

- *caudal* and *bicoid* are maternal effectors
- They drive the expression of the gap genes *hunchback* (hb), *Krüppel* (Kr), *knirps* (kni) and *giant* (gt)
- *tailess* (tll) is a gap gene that we model as an input of the network

Intracellular Network from literature [a]

[a] T. J. Perkins, J. Jaeger, J. Reinitz, and L. Glass. 2006. Reverse engineering the gap gene network of *Drosophila Melanogaster*. PLoS Comput Biol, 2(5)



Working Hypotheses of the Cellular-System

- The system size is 10x100
  - y corresponds to the central portion of D-V axis 45%-55%
  - x corresponds to the 0%-100% of the A-P axis
- Hb, Kr, Kni and Gt are able to diffuse
- Maternal factors do not diffuse

The model specification (1/2)

Hunchback (hb) model

```
% Hb can be found in three different "forms"
% Protein --> pHb
% Gene active --> gHb1, gene inactive --> gHb0
molecule M where (M in [pHb, gHb0, gHb1]).

% Hb reactions

reaction r(gHbAct00)  : [pBcd, gHb0] --> [pBcd, gHb1] rate 0.1114.
reaction r(gHbAct01)  : [pTll, gHb0] --> [pTll, gHb1] rate 0.1.
reaction r(gHbAct02)  : [pHb, gHb0] --> [pHb, gHb1] rate 0.0293.

reaction r(gHbDeAct00)  : [pKni, gHb1] --> [pKni, gHb0] rate 0.3903.
reaction r(gHbDeAct01)  : [pKr, gHb1] --> [pKr, gHb0] rate 0.0124.

reaction r(pHbSynth) : [gHb1] --> [gHb1, pHb] rate 32.03.

reaction r(pHbDegr) : [pHb] --> [] rate 0.136.

reaction r(pHbMove)  : [pHb] --> [firing(pHb)] rate 2.25.
```

The model specification (2/2)

Setting the initial conditions

```
elem(1,[H|T],H).
elem(I,[H|T],X) :- I1 is I-1, elem(I1,T,X).

const size(10,100).

const dataHb([19,20,21,22,23,23,23,24,25,26,26,26,26,26,26,27,28,28,28,28,28,28,
29,30,30,30,30,30,30,29,28,27,27,26,25,24,23,22,21,20,20,20,20,19,18,17,16,15,
14,13,12,11,10,9,9,8,8,7,6,6,6,6,6,6,5,5,5,5,4,4,4,4,4,4,4,4,4,3,3,3,3,3,3,3,2,
2,2,2,2,2,2,2,1,0,0,0,0,0,0,0]).

concentration C of pHb into c(X,Y) where
( const dataHb(List), const size(K,L), X in {1..K}, Y in {1..L}, elem(Y,List,C) ).
```

Results

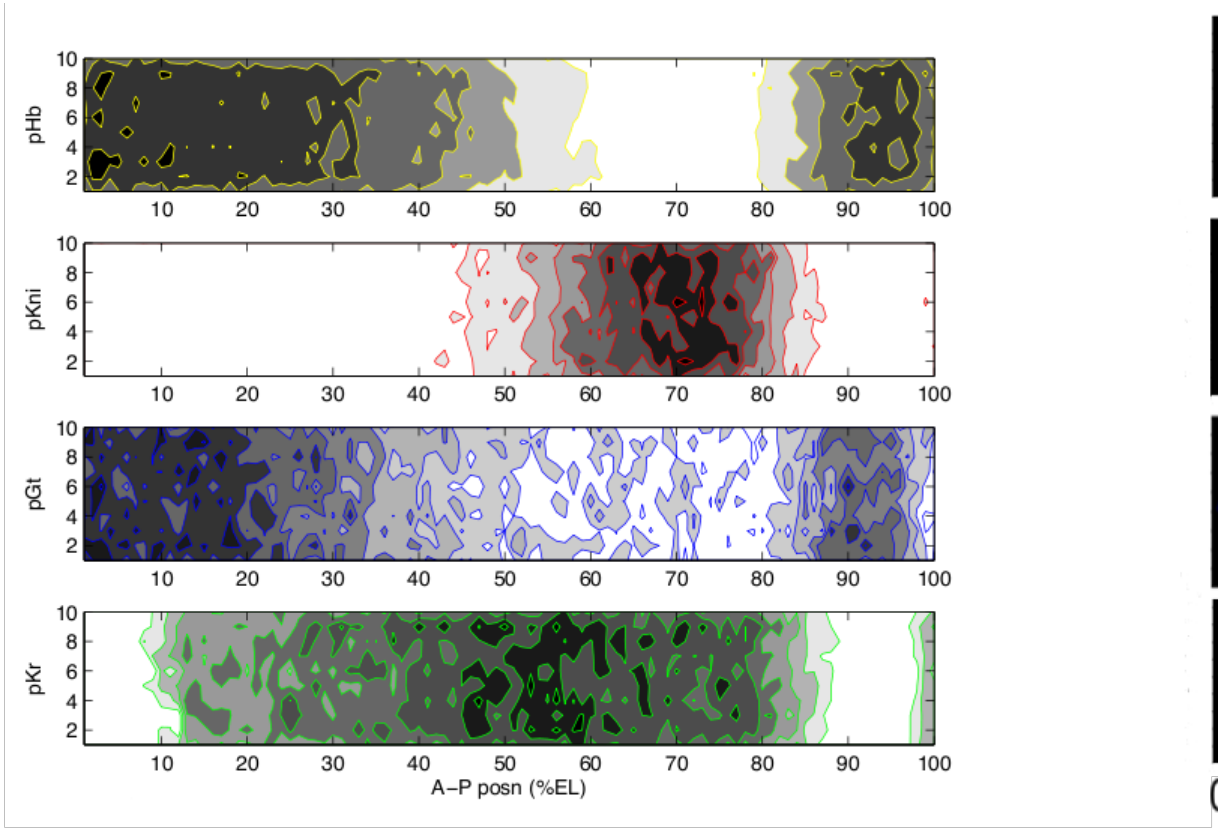


Figure: Simulation results for the four gap genes *hb*, *kni*, *gt*, *Kr* at a simulation time equivalent to the eighth time step of Cleavage Cycle 14A (left) and the corresponding experimental data (right)—% A-P length on the x and % D-V width on the y

Outline

Discussion

- There are tools that can support similar scenarios
  - SPiM, Bio-PEPA, COPASI ...
- Large-scale networks $\longrightarrow$ large-scale specification
  - need to be automatically generated
  - efficiency is not guaranteed
    $\longrightarrow$ our approach translates a high-level specification into a simulation engine tackling large sets of compartments
- Different engines can be considered in the future, e.g., including optimisations approaches like *Tau-Leaping*

Future Works

- Tool side
  - re-engineer the tool towards a community release
  - support dynamic networks and mitotic division
  - improve chemical transfer model
- Biological systems model side
  - studying the *even-skipped* stripes formation
  - introducing cellular phenomena driving the cell sorting – chemotaxis, cell adhesion, mitosis...

A Framework for Modelling and Simulating Networks of Cells

Sara Montagna Mirko Viroli

{sara.montagna, mirko.viroli}@unibo.it

Alma Mater Studiorum—Università di Bologna a Cesena

1st International Workshop
on Interactions between Computer Science and Biology
(CS2Bio'10)

Amsterdam, Netherlands, 10th June 2010