Méthode “logique” multivaluée de René Thomas et logique temporelle

Gilles Bernot

University of Nice sophia antipolis, I3S laboratory, France

Université Nice SOPHIA ANTIPOLIS

i3S sophia antipolis

cnrs dépasser les frontières

Menu

1. Models and formal logic
2. Thomas'models for gene networks
3. Gene networks and temporal logic
4. Models for checking biological hypotheses
5. Extracting experiments from models
6. Model Simplifications

Mathematical models: what for ?

- Models as "Data Base" to store biological knowledge
- Models as design tools for synthetic biology
- Models as logical analysis tools of causality chains
- Models as guidelines for the choice of experiments

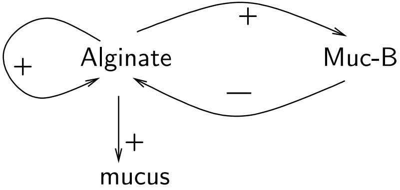
For the 2 or 3 last purposes, models can deviate far from biological descriptions while remaining very useful: "Kleenex" models...

Static Graph *v.s.* Dynamic Behaviour

Difficulty to predict the result of combined regulations

Difficulty to measure the strength of a given regulation

Example of "competitor" circuits



Multistationarity ?
Homeostasy ?

Many underlying models $\approx$ *700 qualitative behaviours*

Mathematical Models and Simulation

1. Rigorously encode sensible knowledge, into ODEs for instance
2.
    - A few parameters are approximately known
    - Some parameters are limited to some intervals
    - Many parameters are *a priori* unknown
3. Perform lot of simulations, compare results with known behaviours, and propose some credible values of the unknown parameters which produce robust acceptable behaviours
4. Perform additional simulations reflecting novel situations
5. If they predict interesting behaviours, propose new biological experiments
6. Simplify the model and try to go further

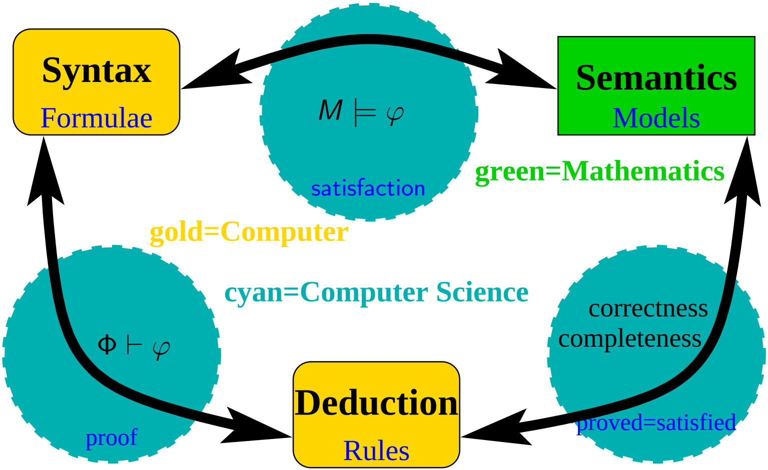
Mathematical Models and Validation

"Brute force" simulations are not the only way to use a computer. There are computer aided environments which help:

- designing simplified models that can be anatically solved
- avoiding models that can be "tuned" *ad libitum*
- validating models with a reasonable number of experiments
- defining only models that could be experimentally refuted
- proving refutability w.r.t. experimental capabilities
- establishing a *methodology*: models $\leftrightarrow$ experiments

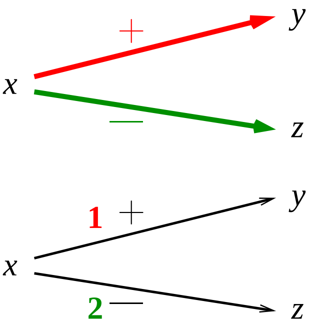
Operability and ***observability*** issues
(*Observability Group*, Epigenomics Project)

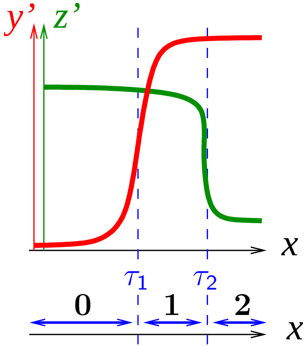
Formal Logic: syntax/semantics/deduction



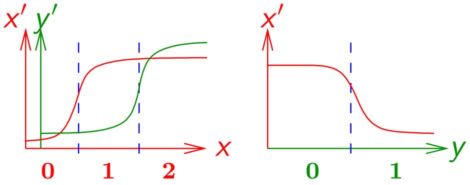
Menu

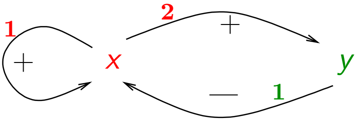
Multivalued Regulatory Graphs





Regulatory Networks (R. Thomas)





No help : K_x K_y

x helps : $K_{x,x}$ $K_{y,x}$

Absent y helps : $K_{x,\overline{y}}$

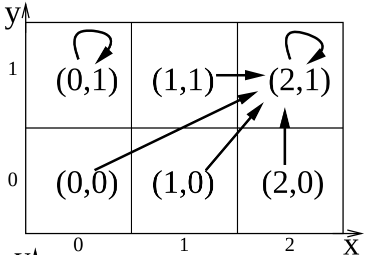
Both : $K_{x,x\overline{y}}$

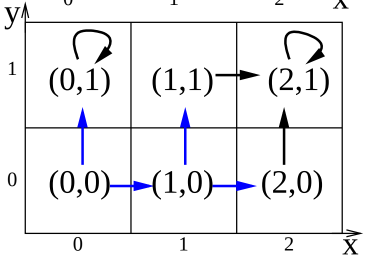
(x,y)	Focal Point
(0,0)	$(K_{x,\overline{y}}, K_y)$
(0,1)	(K_x, K_y)
(1,0)	$(K_{x,x\overline{y}}, K_y)$
(1,1)	$(K_{x,x}, K_y)$
(2,0)	$(K_{x,x\overline{y}}, K_{y,x})$
(2,1)	$(K_{x,x}, K_{y,x})$

State Graphs

(x,y)	*Focal Point*
(0,0)	$(K_{x,\overline{y}}, K_y)=(2,1)$
(0,1)	$(K_x, K_y)=(0,1)$
(1,0)	$(K_{x,x\overline{y}}, K_y)=(2,1)$
(1,1)	$(K_{x,x}, K_y)=(2,1)$
(2,0)	$(K_{x,x\overline{y}}, K_{y,x})=(2,1)$
(2,1)	$(K_{x,x}, K_{y,x})=(2,1)$

"desynchronization" $\longrightarrow$
by units of Manhattan distance





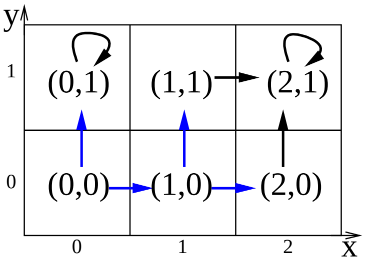
Do it by yourselves !

Example on paper sheets. . .

Menu

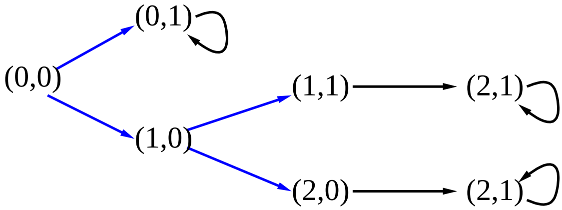
1. Models and formal logic
2. Thomas'models for gene networks
3. Gene networks and temporal logic
4. Models for checking biological hypotheses
5. Extracting experiments from models
6. Model Simplifications

Time has a tree structure. . .



As many possible state graphs as possible parameter sets. . . (huge number)

. . . from each initial state:



CTL = Computation Tree Logic

Atoms = comparaisons : (x=2) (y>0) ...

Logical connectives: $(\varphi_1 \wedge \varphi_2)$ $(\varphi_1 \implies \varphi_2)$ $\cdots$

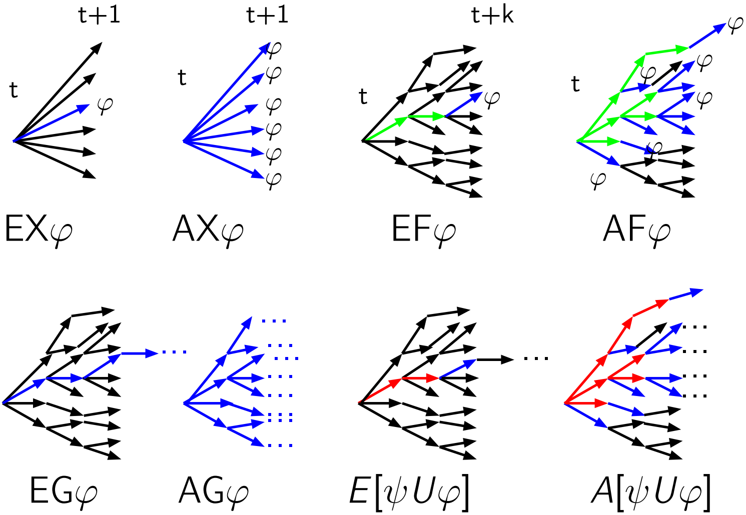
Temporal modalities: made of 2 characters

first character	second character
A = for **A**ll path choices	X = ne**X**t state
	F = for some **F**uture state
E = there **E**xist a choice	G = for all future states (**G**lobally)
	U = **U**ntil

$AX(y = 1)$: the concentration level of y belongs to the interval 1 in all states directly following the considered initial state.

$EG(x = 0)$: there exists at least one path from the considered initial state where x always belongs to its lower interval.

Semantics of Temporal Connectives



CTL to encode Biological Properties

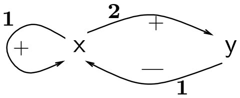
Common properties:

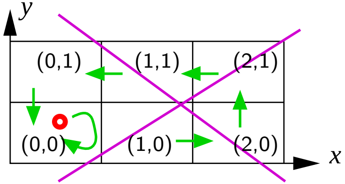
"functionality" of a sub-graph

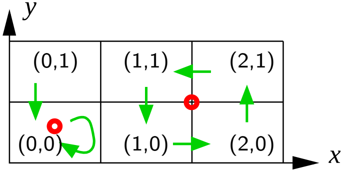
Special role of "feedback loops"

– positive: *multistationnarity* (even number of —)

– negative: *homeostasy* (odd number of —)







Characteristic properties: $\begin{cases} (x = 2) \Longrightarrow AG(\neg(x = 0)) \\ (x = 0) \Longrightarrow AG(\neg(x = 2)) \end{cases}$

They express *"the positive feedback loop is functional"*

(satisfaction of these formulas relies on the parameters $K_{...}$)

Model Checking

Efficiently computes all the states of a state graph which satisfy a given formula: $\{\ \eta \ \mid \ M \ \models_\eta \ \varphi\ \}$.

Efficiently select the models which globally satisfy a given formula.

Menu

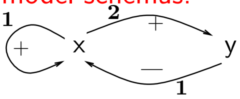
Computer Aided Elaboration of Models

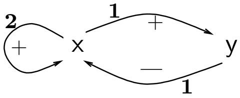
From biological knowledge and/or biological hypotheses, it comes:

- properties:

  "Without stimulus, if gene x has its basal expression level, then it remains at this level."

- model schemas:





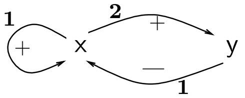
. . .

Formal logic and formal models allow us to:

- verify hypotheses and check consistency
- elaborate more precise models incrementally
- suggest new biological experiments to efficiently reduce the number of potential models

The Two Questions

$\Phi = \{\varphi_1, \varphi_2, \cdots, \varphi_n, H\}$ and $\mathcal{M} =$



. . .

$K_{\boldsymbol{x}} \cdots K_{\boldsymbol{x},\boldsymbol{x}} \cdots K_{\boldsymbol{x},\boldsymbol{xy}} \cdots$

1. Is it possible that Φ and $\mathcal{M}$?
   Consistency of knowledge and hypotheses. Means to select models belonging to the schemas that satisfy Φ.
   $(\exists?\ M \in \mathcal{M} \ \mid \ M \models \Phi)$
2. If so, is it true *in vivo* that Φ and $\mathcal{M}$?
   Compatibility of one of the selected models with the biological object. Require to propose experiments to **validate** or **refute** the selected model(s).

$\rightarrow$ Computer aided *proofs* and *validations*

Theoretical Models $\leftrightarrow$ Experiments

CTL formulas are satisfied (or refuted) w.r.t. a set of paths from a given initial state

- They can be tested against the possible paths of the theoretical models ($M \models_{Model\ Checking} \varphi$)
- They can be tested against the biological experiments ($Biological\_Object \models_{Experiment} \varphi$)

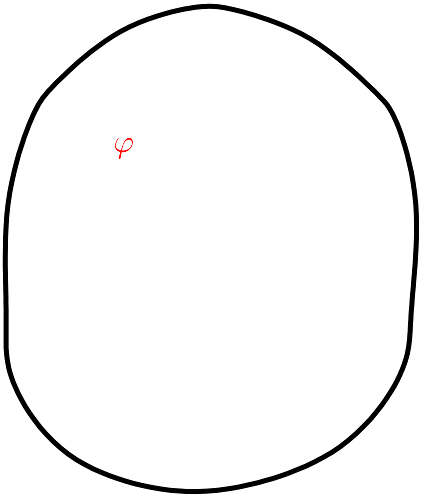
CTL is a bridge between theoretical models and biological objects

Menu

Generation of biological experiments (1)

Set of all the formulas:

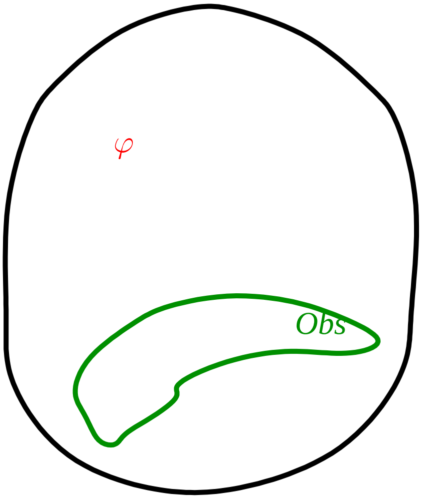
φ = hypothesis



Generation of biological experiments (2)

Set of all the formulas:

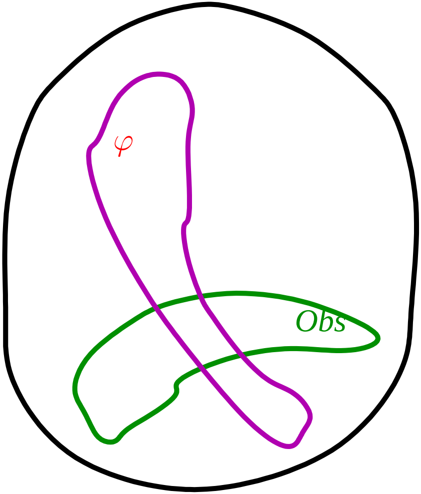
φ = hypothesis
Obs = possible experiments



Generation of biological experiments (3)

Set of all the formulas:

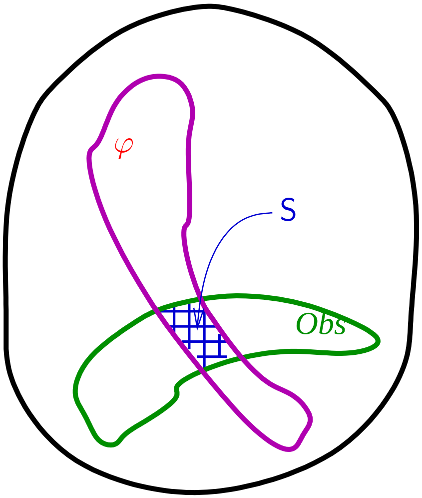
φ = hypothesis
Obs = possible experiments
$Th(\varphi) = \varphi$ inferences



Generation of biological experiments (4)

Set of all the formulas:

φ = hypothesis
Obs = possible experiments
$Th(\varphi)$ = φ inferences
S = sensible experiments



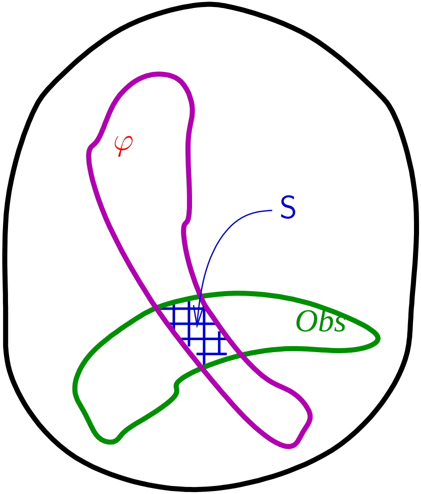
Generation of biological experiments (5)

Set of all the formulas:

φ = hypothesis
Obs = possible experiments
$Th(\varphi)$ = φ inferences
S = sensible experiments

Refutability:

$$S \Longrightarrow \varphi \ ?$$



Generation of biological experiments

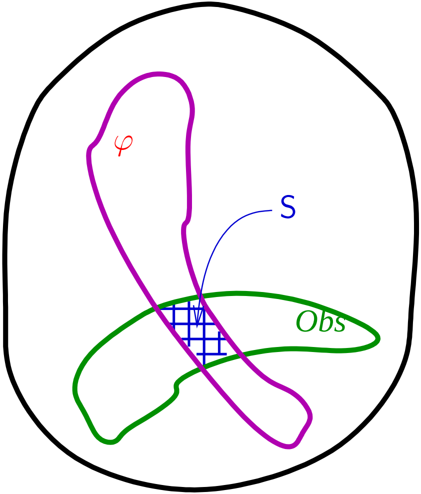
Set of all the formulas:

φ = hypothesis
Obs = possible experiments
$Th(\varphi)$ = φ inferences
S = sensible experiments

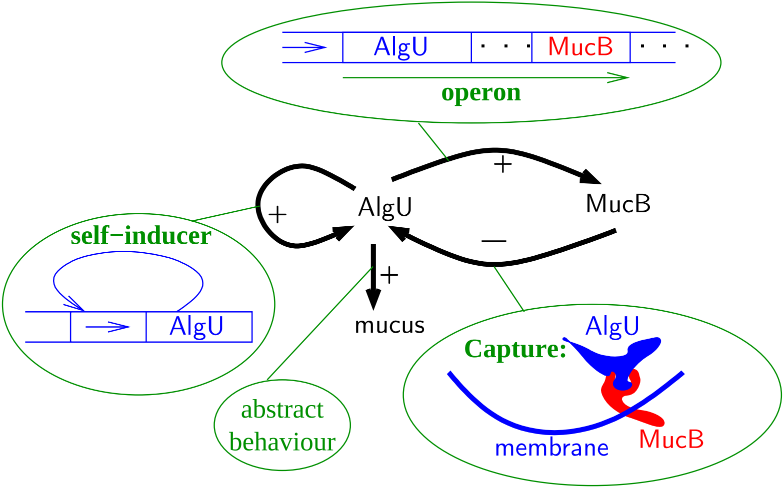
Refutability:

S $\Longrightarrow$ φ ?

Best refutations:
Choice of experiments in S ?
... optimisations

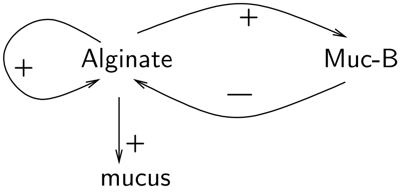


Example: Mucus Production in *P. aeruginosa*



How to validate a multistationnarity

$\mathcal{M}$: *(unknown thresholds)*



$$\Phi: \begin{cases} (Alginate = 2) \Longrightarrow AG(Alginate = 2) & (hypothesis) \\ (Alginate = 0) \Longrightarrow AG(Alginate < 2) & (knowledge) \end{cases}$$

Assume that only *mucus* can be observed:

Lemma: $AG(Alginate = 2) \Longleftrightarrow AFAG(mucus = 1)$

(... formal proof by computer ...)

$\rightarrow$ To validate: $(Alginate = 2) \Longrightarrow AFAG(mucus = 1)$

$(Alginate = 2) \Longrightarrow AFAG(mucus = 1)$

Karl Popper:
to validate = to try to refute
thus A=false is useless
experiments must begin with a pulse

$A \Longrightarrow B$	*true*	*false*
true	true	false
false	true	true

The pulse forces the bacteria to reach the initial state $Alginate = 2$.
If the state is not directly controlable we need to prove lemmas:

$$(something\ reachable) \Longrightarrow (Alginate = 2)$$

General form of a test:

$$(something\ \underline{reachable}) \Longrightarrow (something\ \underline{observable})$$

Menu

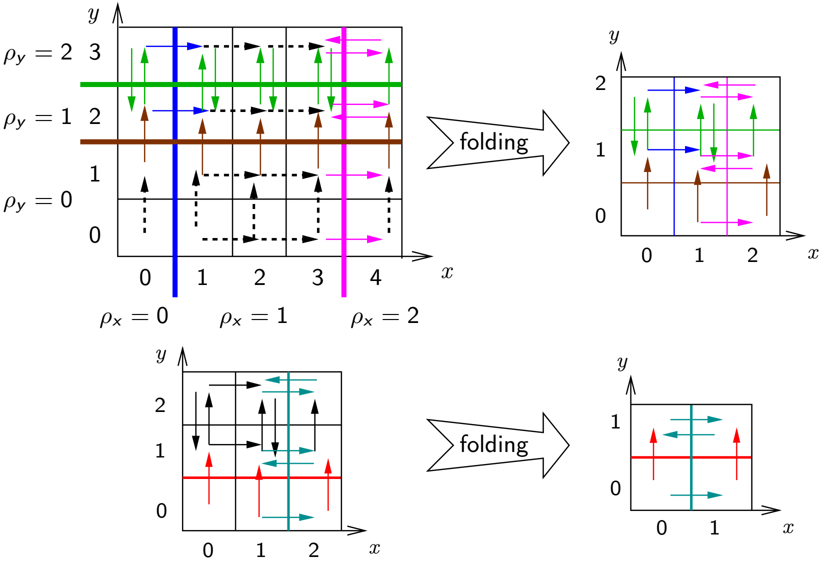
Hypothesis driven model simplifications

Successive simplified views of the studied biological object:

Model M_1 satisfies φ_1	$\Longleftrightarrow$	Model M_2 satisfies φ_2	$\Longleftrightarrow$	Model M_3 satisfies φ_3	$\Longleftrightarrow$	$\ldots$

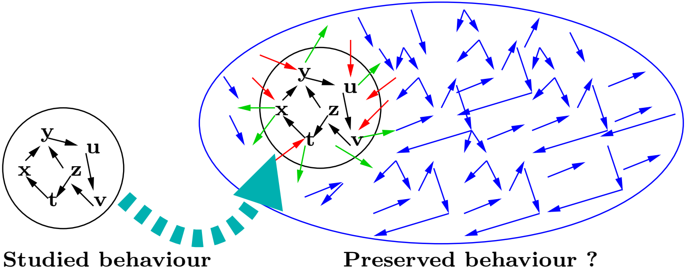
Example: gene removal often preserves the number of attraction basins [Naldi&al.2011]

Simplifications *via* level folding



Simplifications *via* subgraphs

Embeddings of Regulatory Networks:



Necessary and sufficient condition on the *local* dynamics of the "input frontier"

. . . *Also* fusion of genes, *etc.*

Take Home Messages

Formalize the hypotheses that motivate the biological research

Behavioural *properties* (Φ) are as much important as *models* ($\mathcal{M}$)

Symbolic parameter identification is essential

Modelling is significant only with respect to the considered experimental *reachability* and *observability* (for refutability)

Formal proofs can suggest wet experiments

Mathematical models are not reality: let's use this freedom ! (simplified views of a biological object)

"Kleenex" models help understanding main behaviours