Property Driven Models of Gene Networks

Gilles Bernot

University Nice Sophia Antipolis, I3S laboratory, France

Menu

1. Models and Formal Logic
2. Gene Networks and Temporal Logic
3. Extracting Experiments from Models
4. Model Simplifications
5. Circadian Circle, Seasons and Jet-lag

Mathematical models: what for ?

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

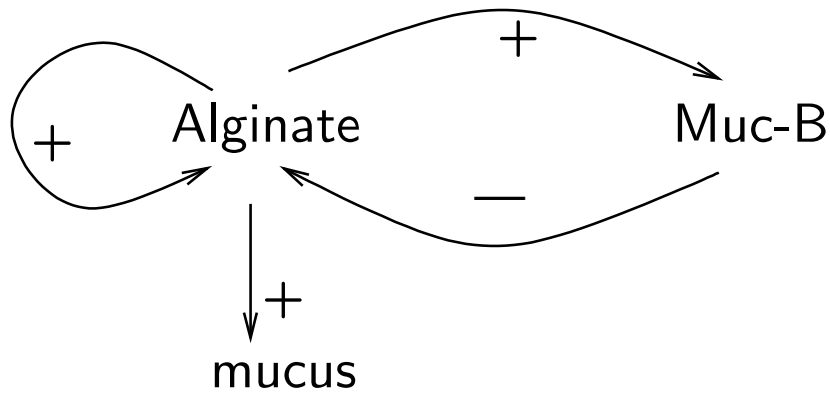
For the 2 last purposes, models can deviate far from biological descriptions but remain very useful: "Kleenex" models !

Static Graph *v.s.* Dynamic Behaviour

Difficulty to predict the result of combined regulations

Difficulty to measure the strength of a given regulation

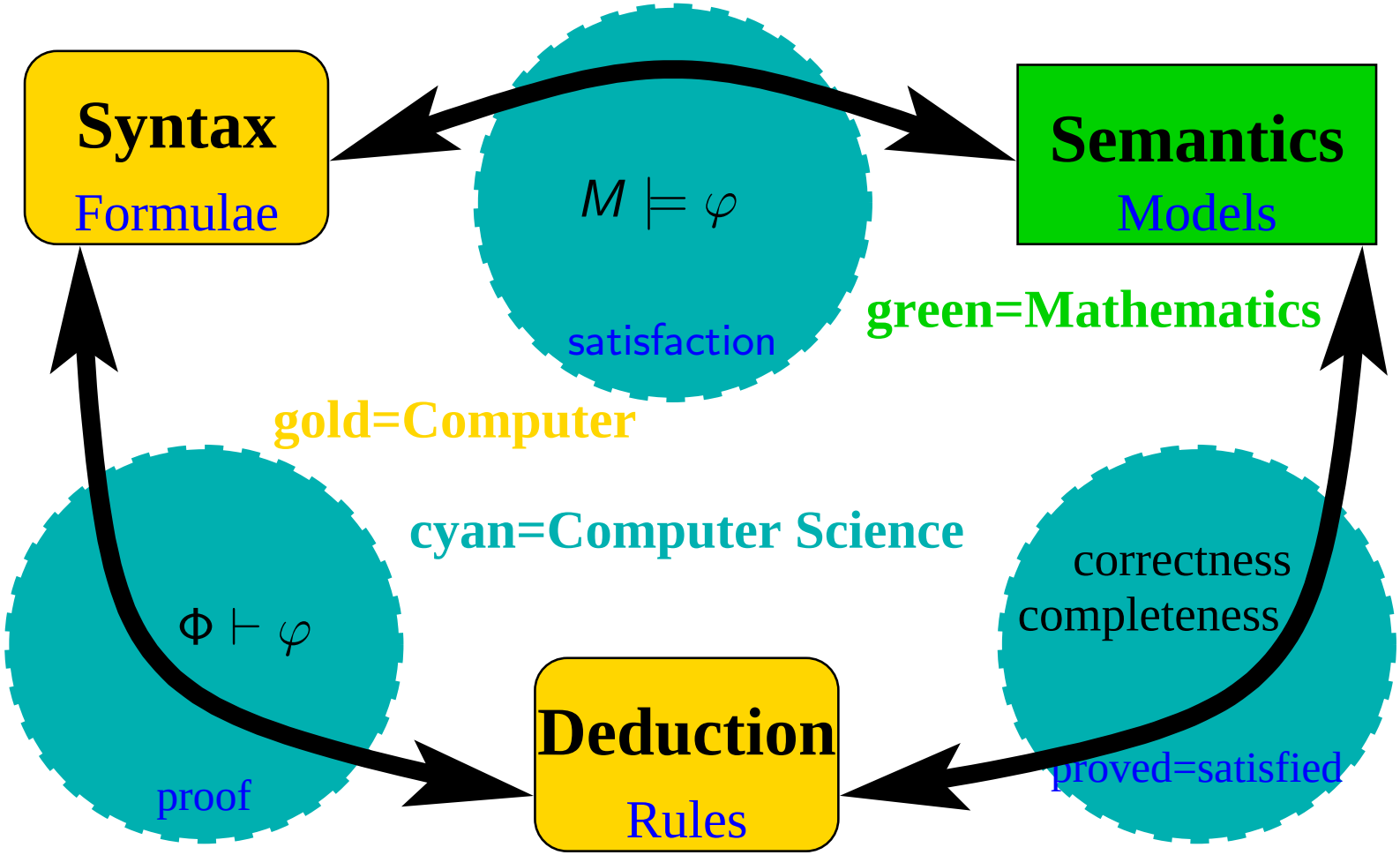
Example of "competitor" circuits



Multistationarity ?
Homeostasy ?

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

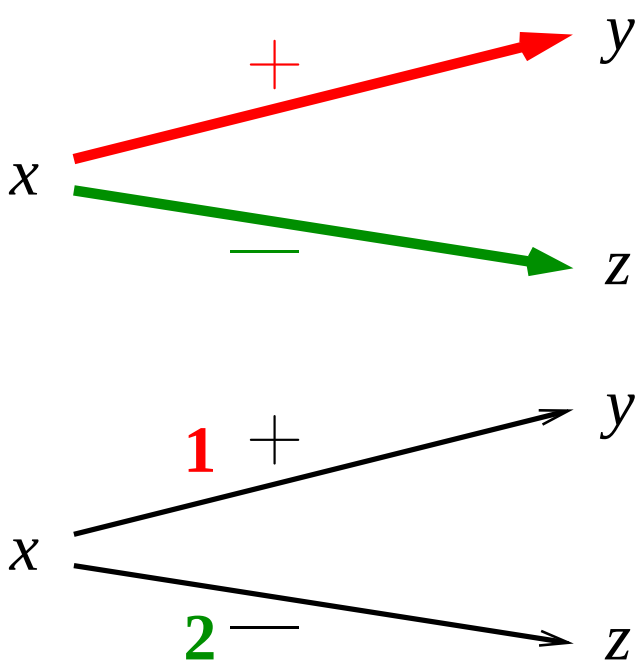
Formal Logic: syntax/semantics/deduction

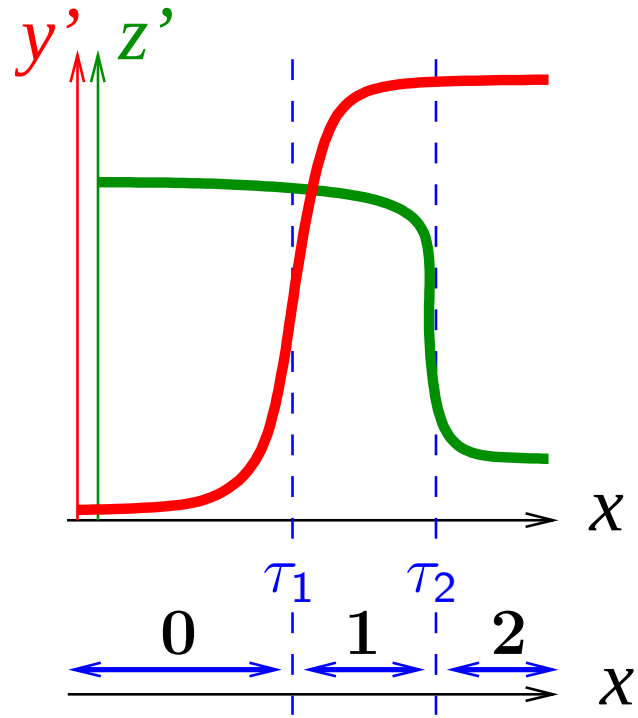


Menu

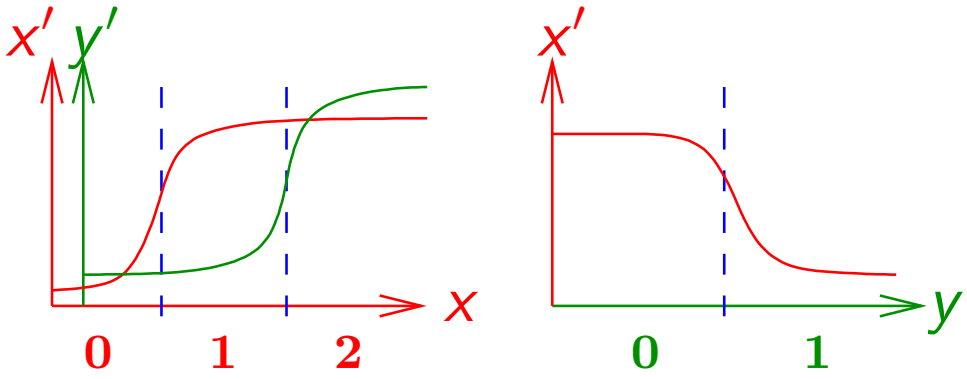
1. Models and Formal Logic
2. Gene Networks and Temporal Logic
3. Extracting Experiments from Models
4. Model Simplifications
5. Circadian Circle, Seasons and Jet-lag

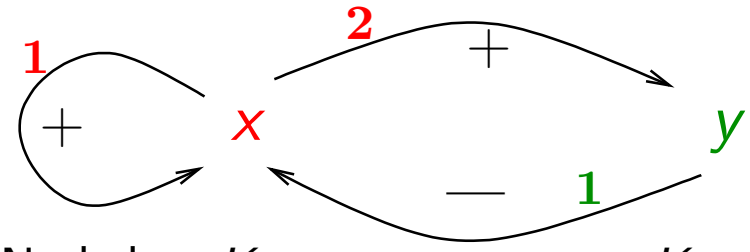
Multivalued Regulatory Graphs





Regulatory Networks (R. Thomas)





No help : K_x K_y

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

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

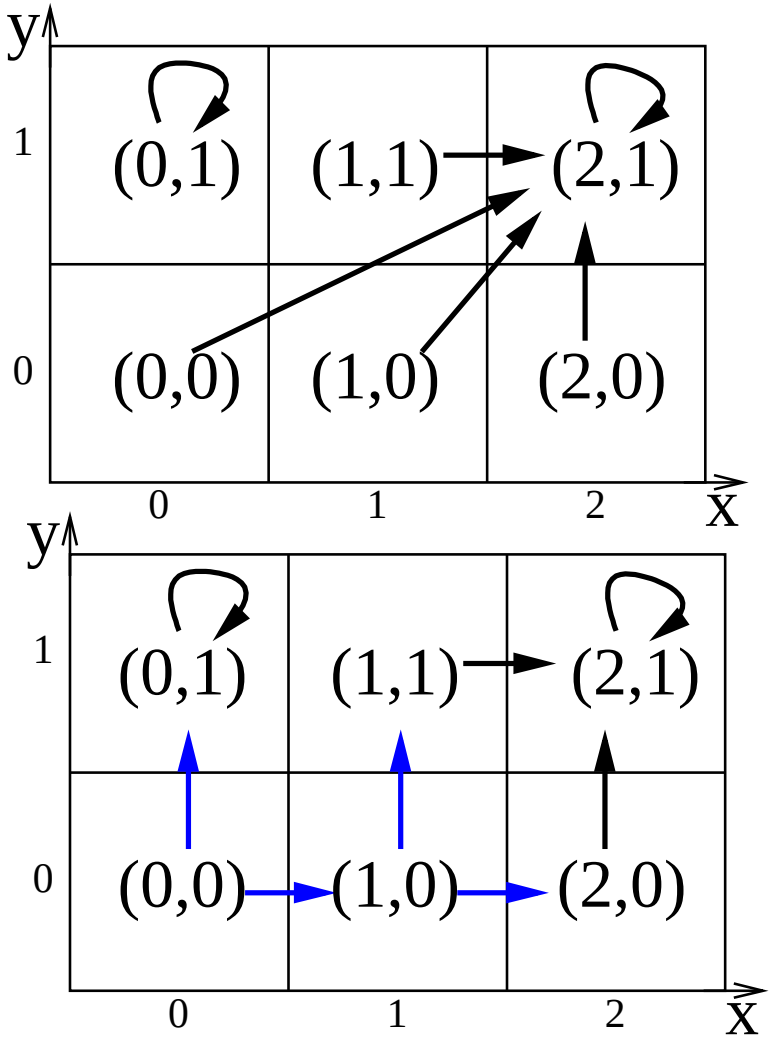
Both : $K_{x,x\overline{y}}$

(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



CTL = Computation Tree Logic

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

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

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.

CTL to encode Biological Properties

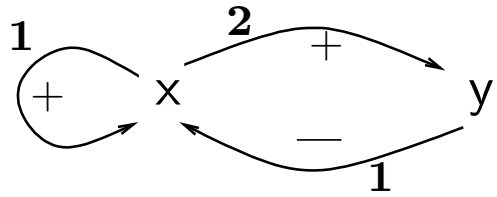
Common properties:

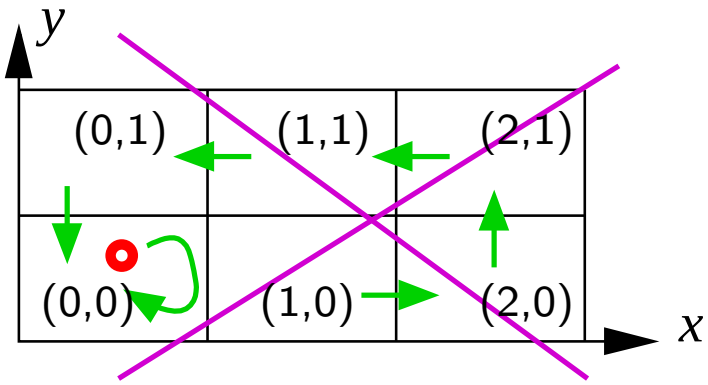
"functionality" of a sub-graph

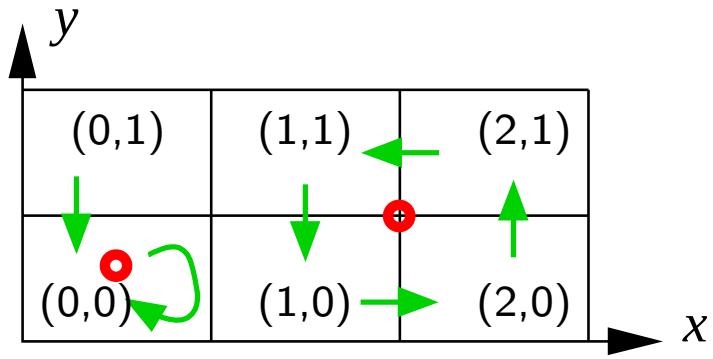
Special role of "feedback loops"

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

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







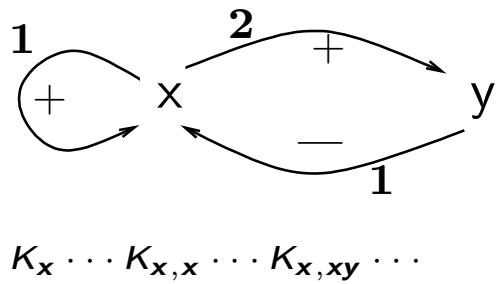
Characteristic properties: $\begin{cases} (x = 2) \Longrightarrow AG(\neg(x = 0)) \\ (x = 0) \Longrightarrow AG(\neg(x = 2)) \end{cases}$

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

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

The Two Questions

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



. . .

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 formulae 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

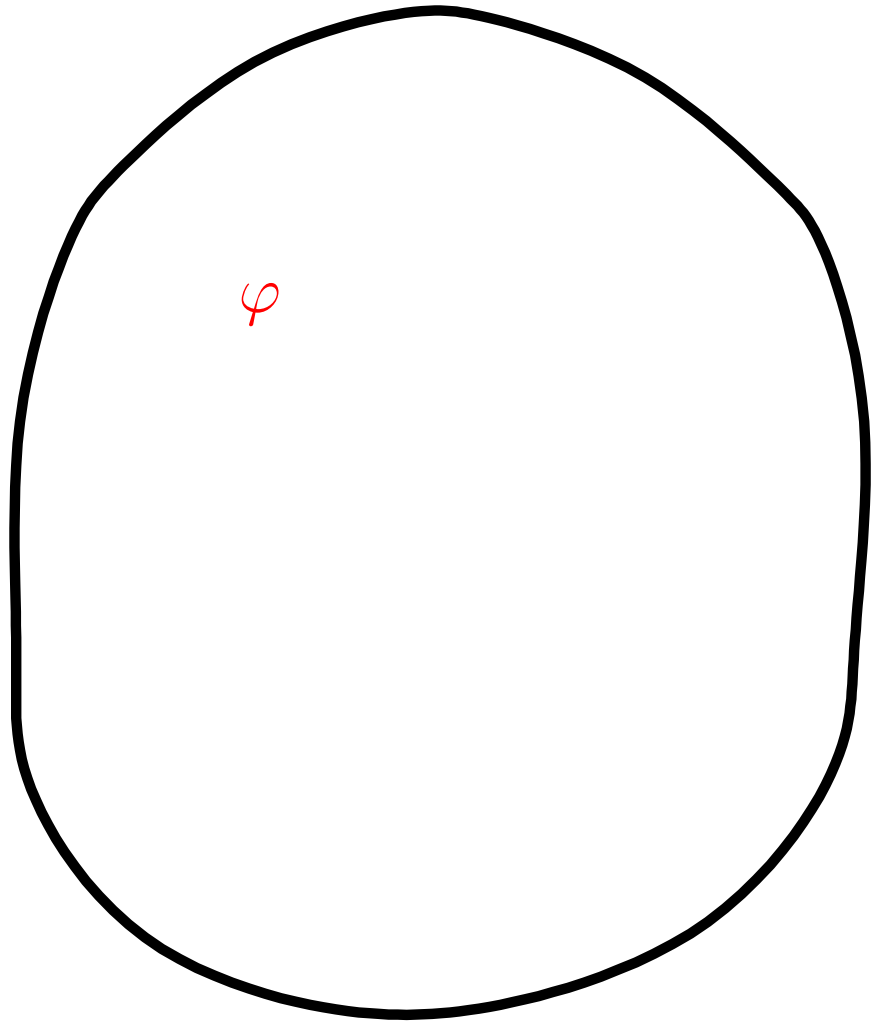
Menu

1. Models and Formal Logic
2. Gene Networks and Temporal Logic
3. Extracting Experiments from Models
4. Model Simplifications
5. Circadian Circle, Seasons and Jet-lag

Generation of biological experiments (1)

Set of all the formulae:

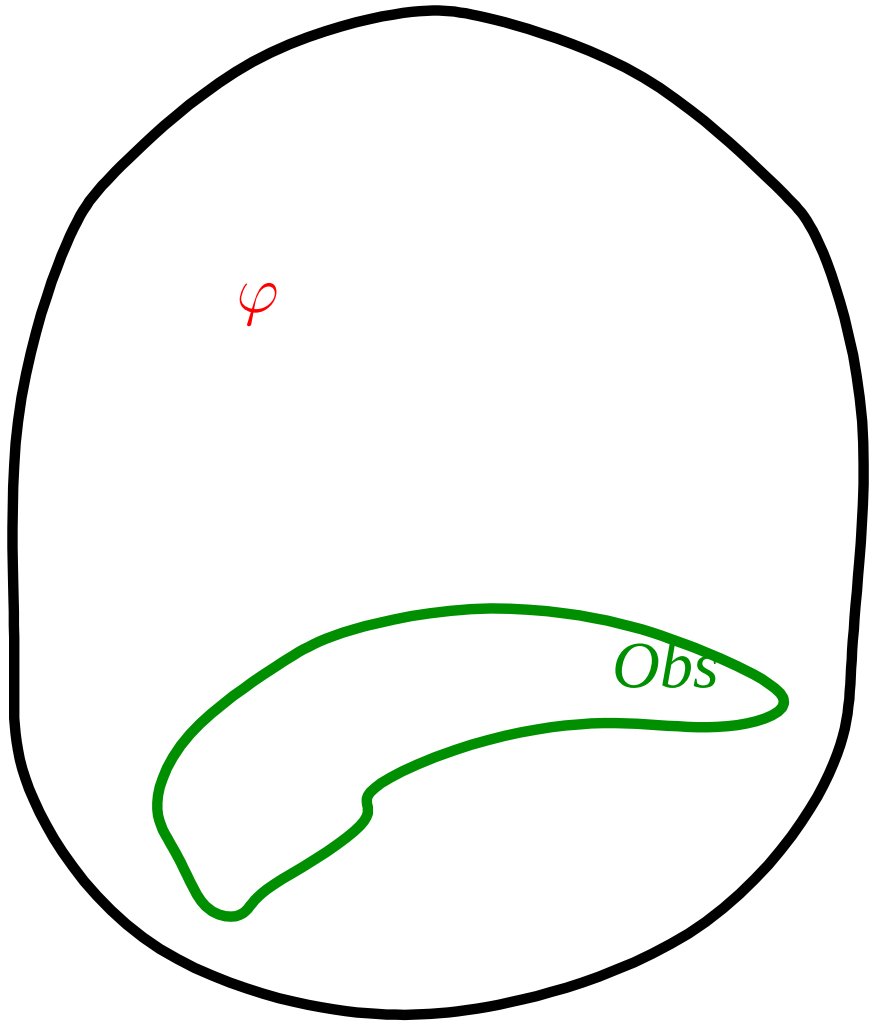
φ = hypothesis



Generation of biological experiments (2)

Set of all the formulae:

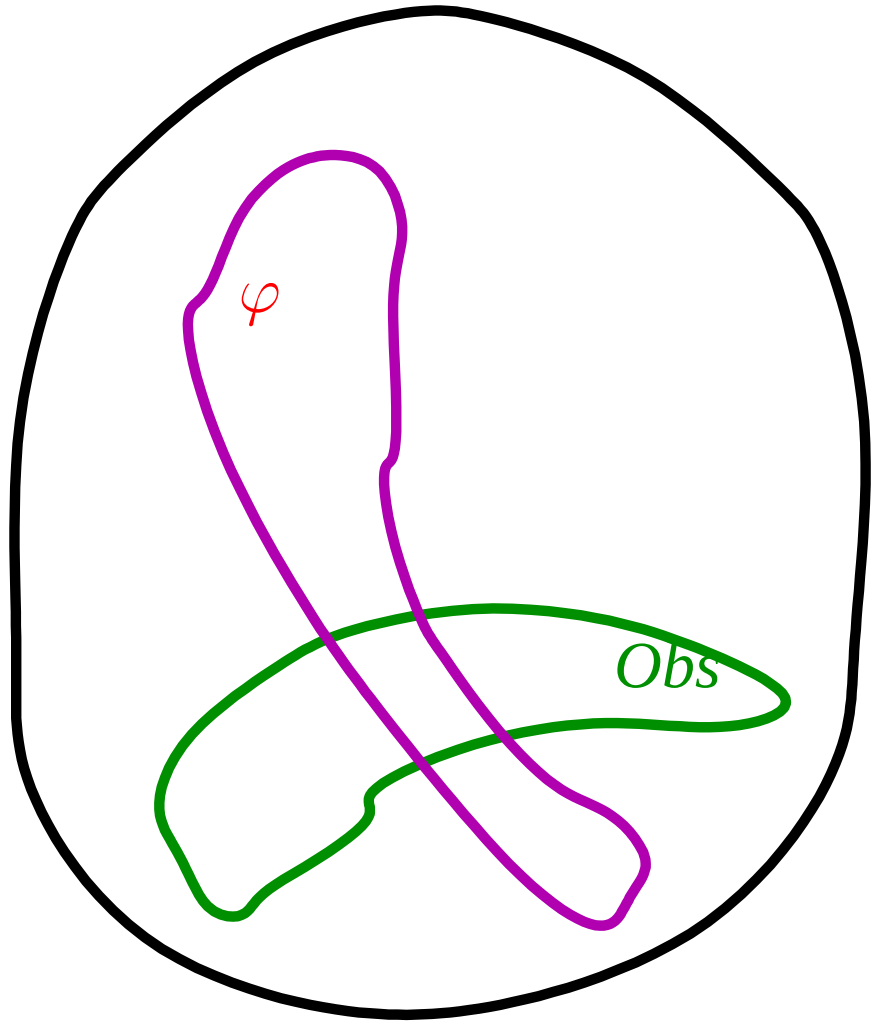
φ = hypothesis
Obs = possible experiments



Generation of biological experiments (3)

Set of all the formulae:

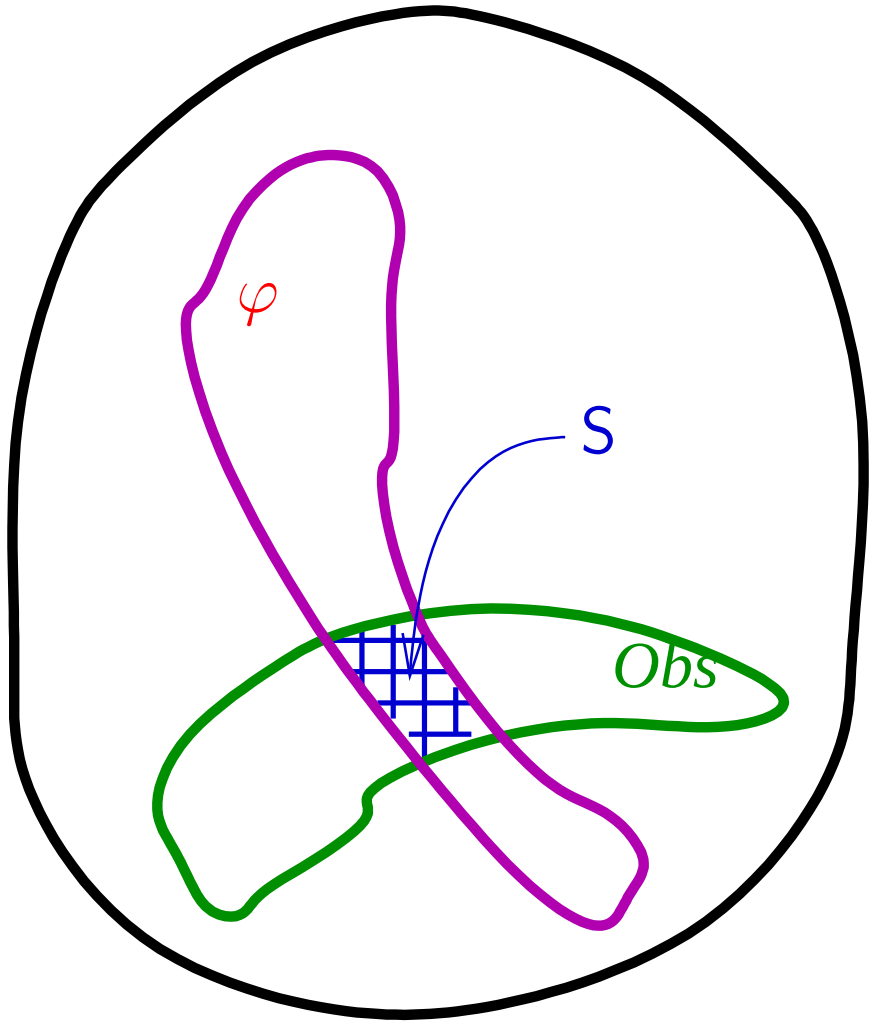
φ = hypothesis
Obs = possible experiments
$Th(\varphi)$ = φ inferences



Generation of biological experiments (4)

Set of all the formulae:

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



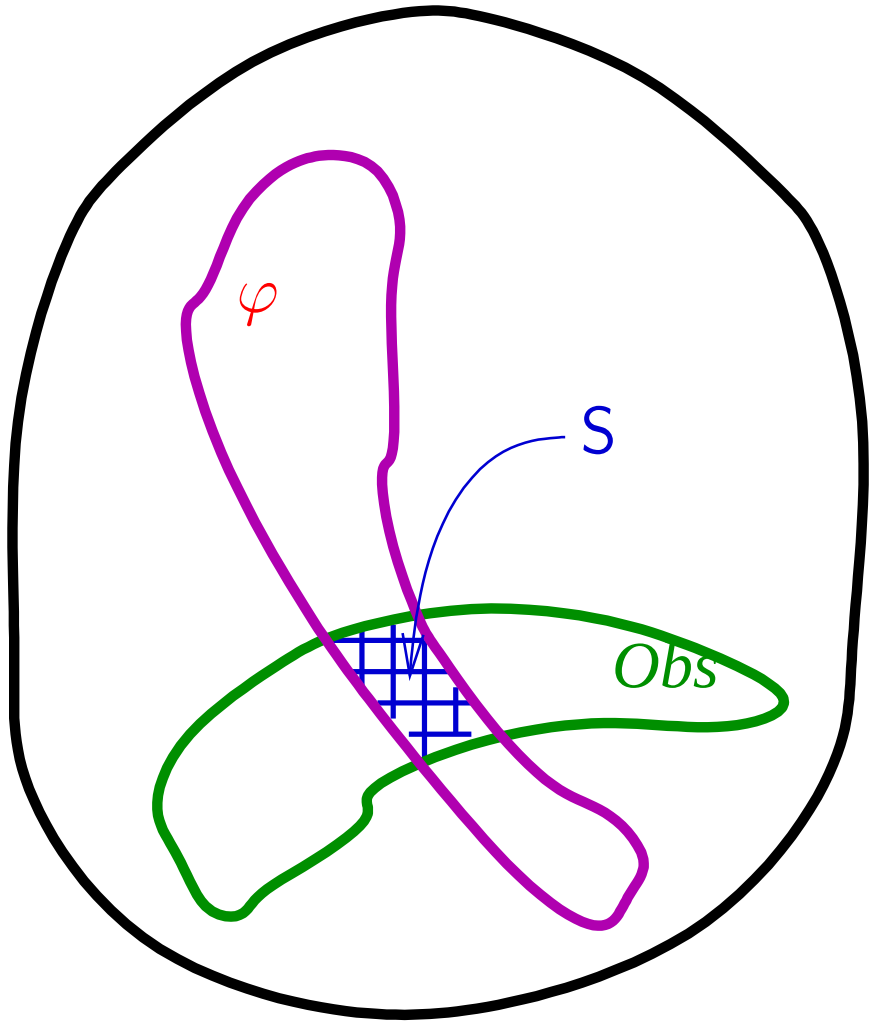
Generation of biological experiments (5)

Set of all the formulae:

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

Refutability:

$$S \Longrightarrow \varphi \ ?$$



Generation of biological experiments

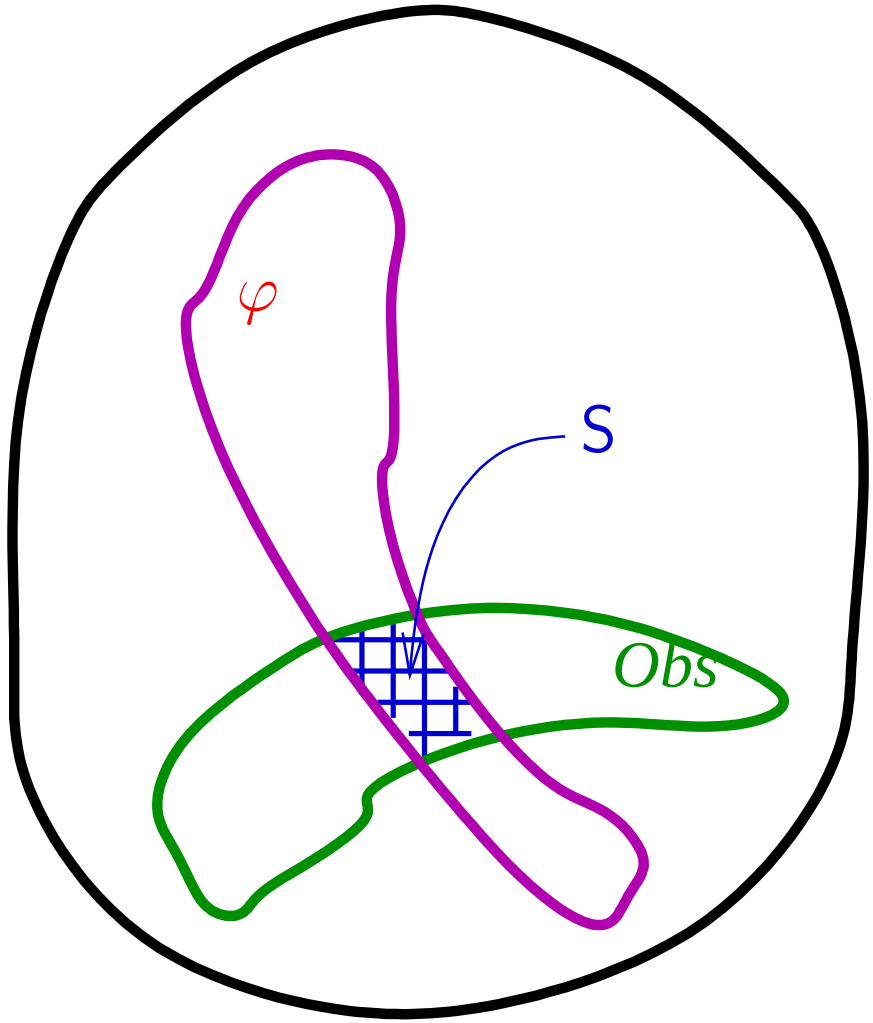
Set of all the formulae:

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

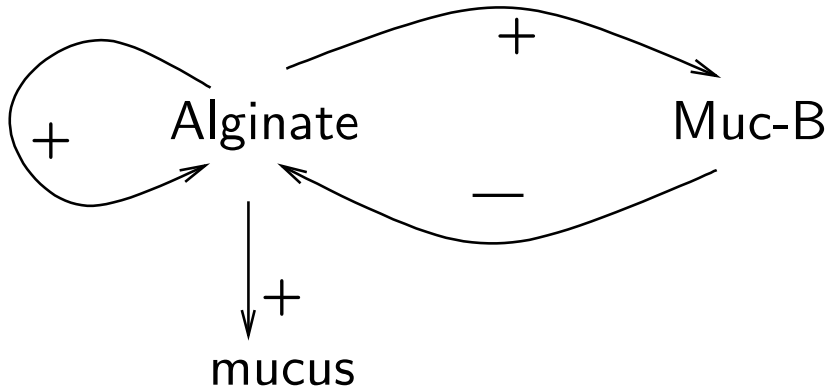
Refutability:

$$S \Longrightarrow \varphi \ ?$$

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



How to validate a multistationnarity



Hypotheses: $\begin{cases} (Alginate = 2) \Longrightarrow AG(Alginate = 2) \\ (Alginate = 0) \Longrightarrow AG(Alginate < 2) \end{cases}$

Assume that only *mucus* can be observed:

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

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

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

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

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

Karl Popper:

to validate = to try to refute

thus A=false is useless

experiments must begin with a pulse

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

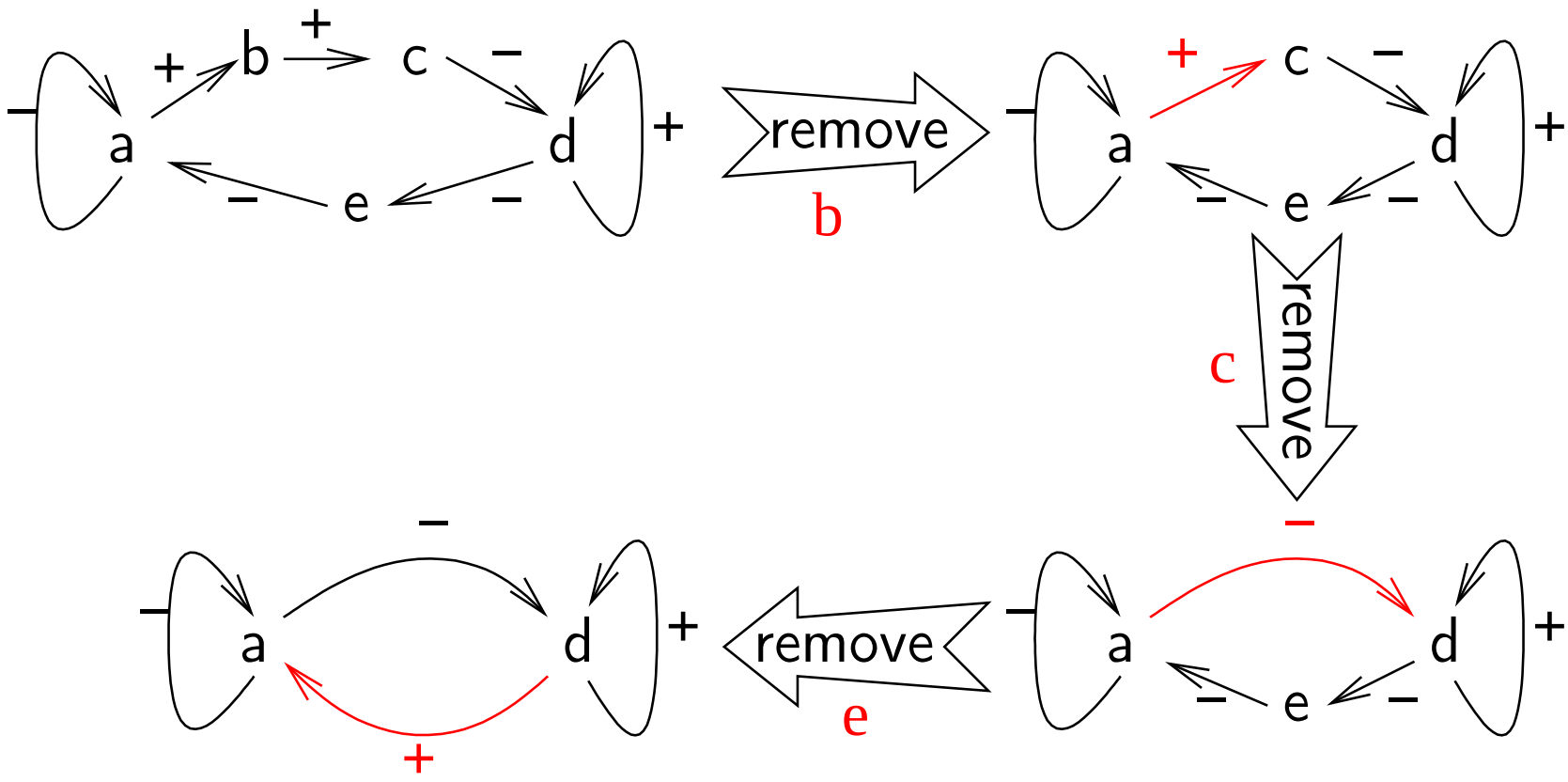
1. Models and Formal Logic
2. Gene Networks and Temporal Logic
3. Extracting Experiments from Models
4. Model Simplifications
5. Circadian Circle, Seasons and Jet-lag

Hypothesis driven model simplifications

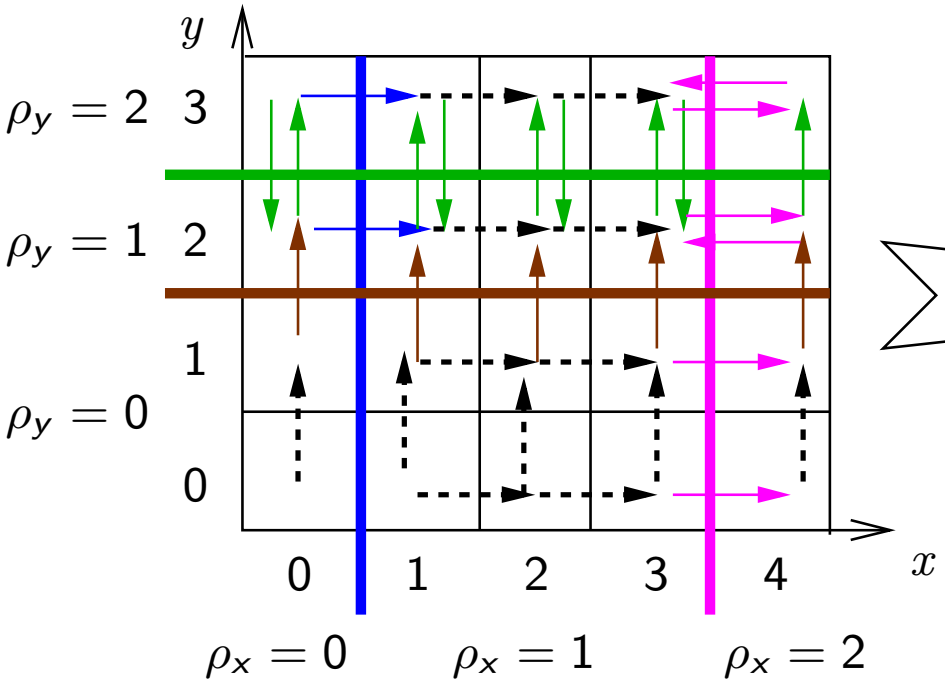
Successive simplified views of the studied biological object:

$$\begin{array}{ccccccc} \text{Model} & & \text{Model} & & \text{Model} & & \\ M_1 & \Longleftrightarrow & M_2 & \Longleftrightarrow & M_3 & \Longleftrightarrow & \ldots \\ \text{satisfies} & & \text{satisfies} & & \text{satisfies} & & \\ \varphi_1 & & \varphi_2 & & \varphi_3 & & \end{array}$$

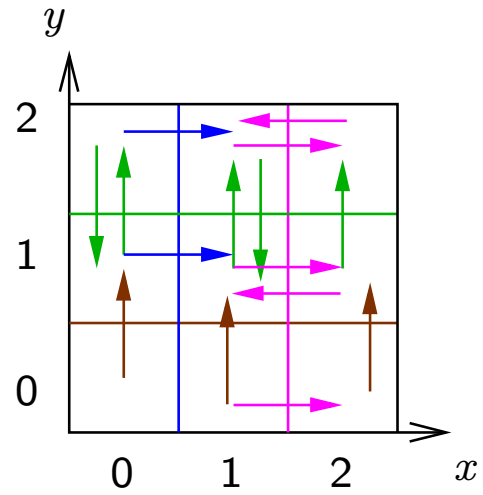
Simplifications *via* gene removing (Naldi)

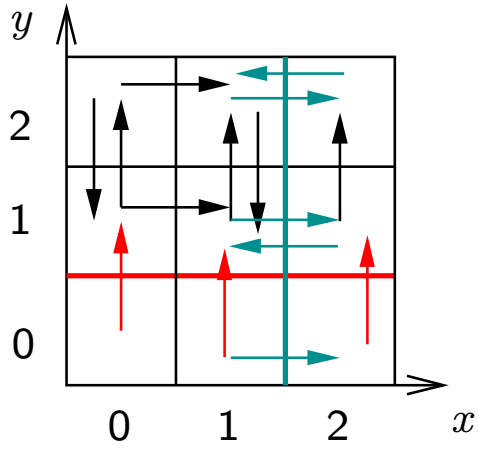


Simplifications *via* level folding

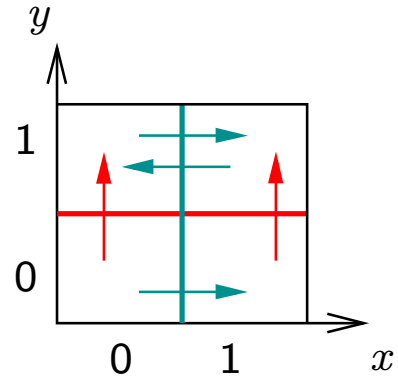


folding



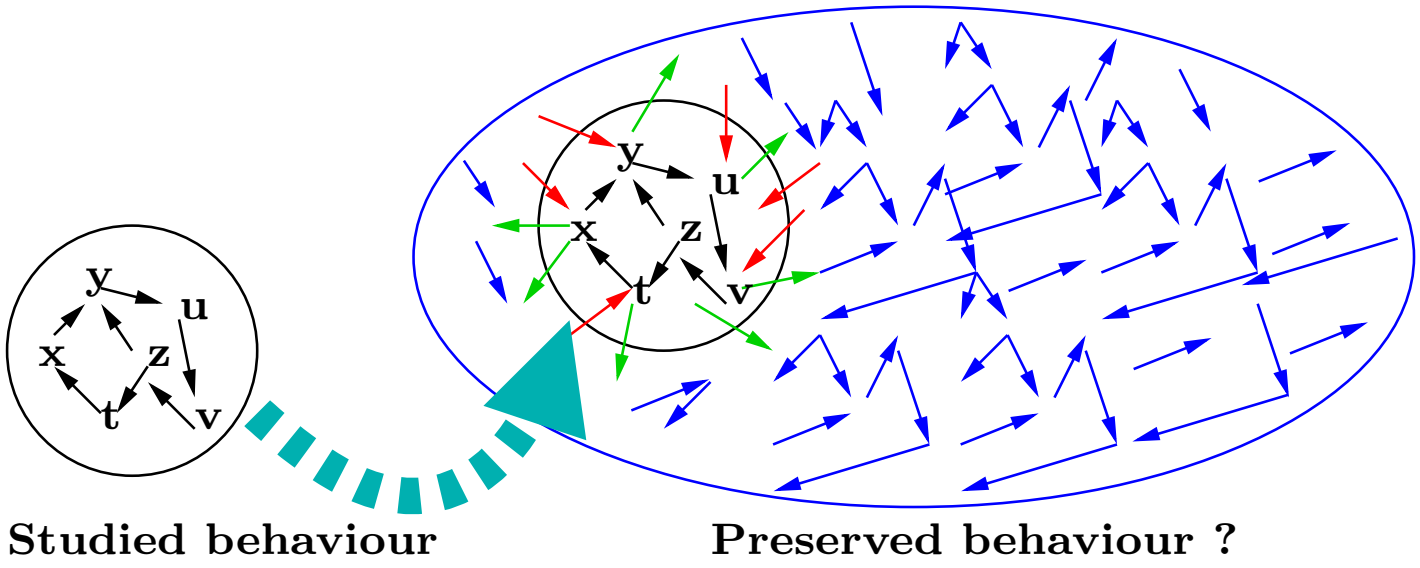


folding



Simplifications *via* subgraphs

Embeddings of Regulatory Networks:



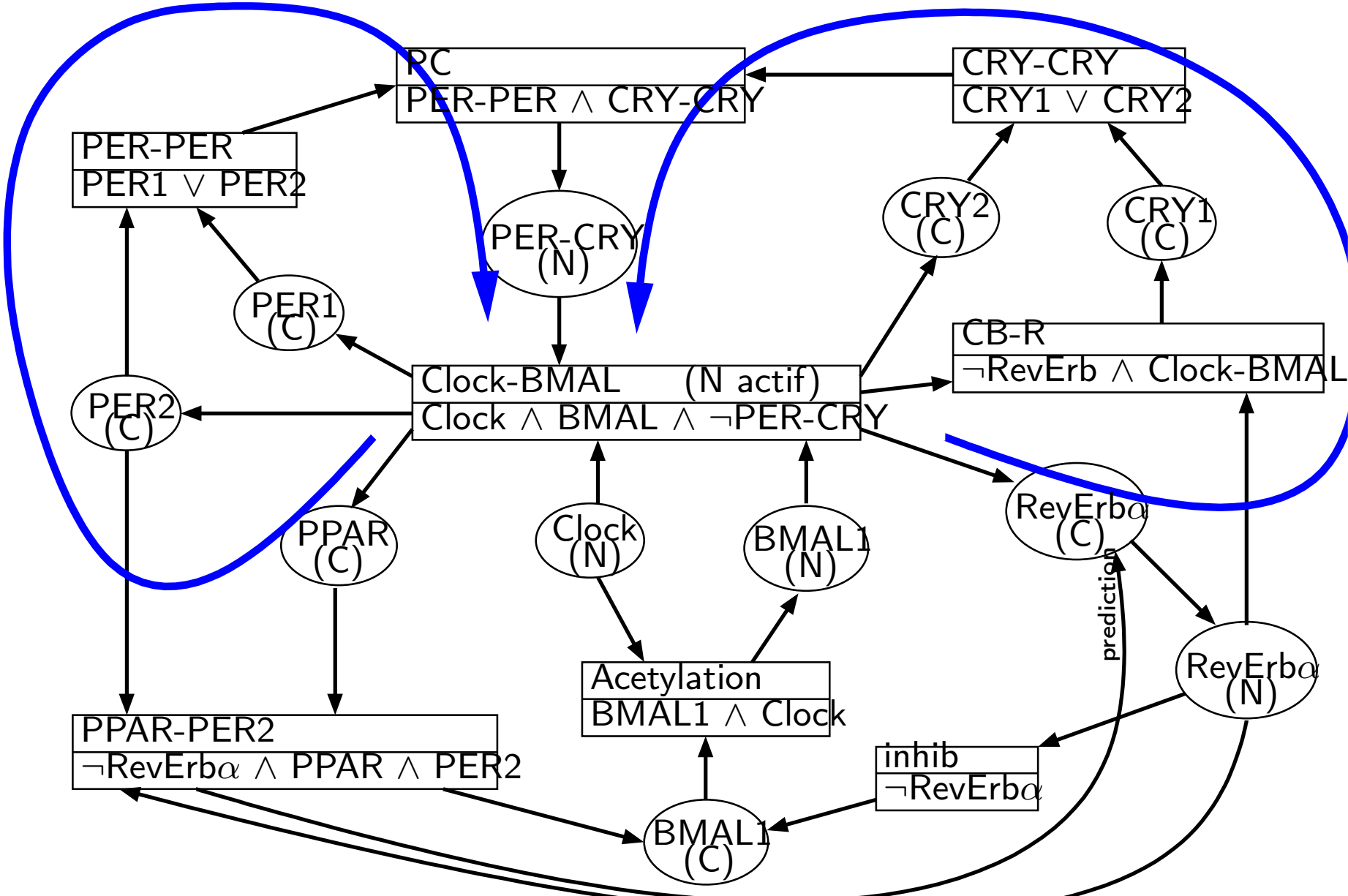
Necessary and sufficient condition on the *local* dynamics of the "input frontier"

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

Menu

1. Models and Formal Logic
2. Gene Networks and Temporal Logic
3. Extracting Experiments from Models
4. Model Simplifications
5. Circadian Circle, Seasons and Jet-lag

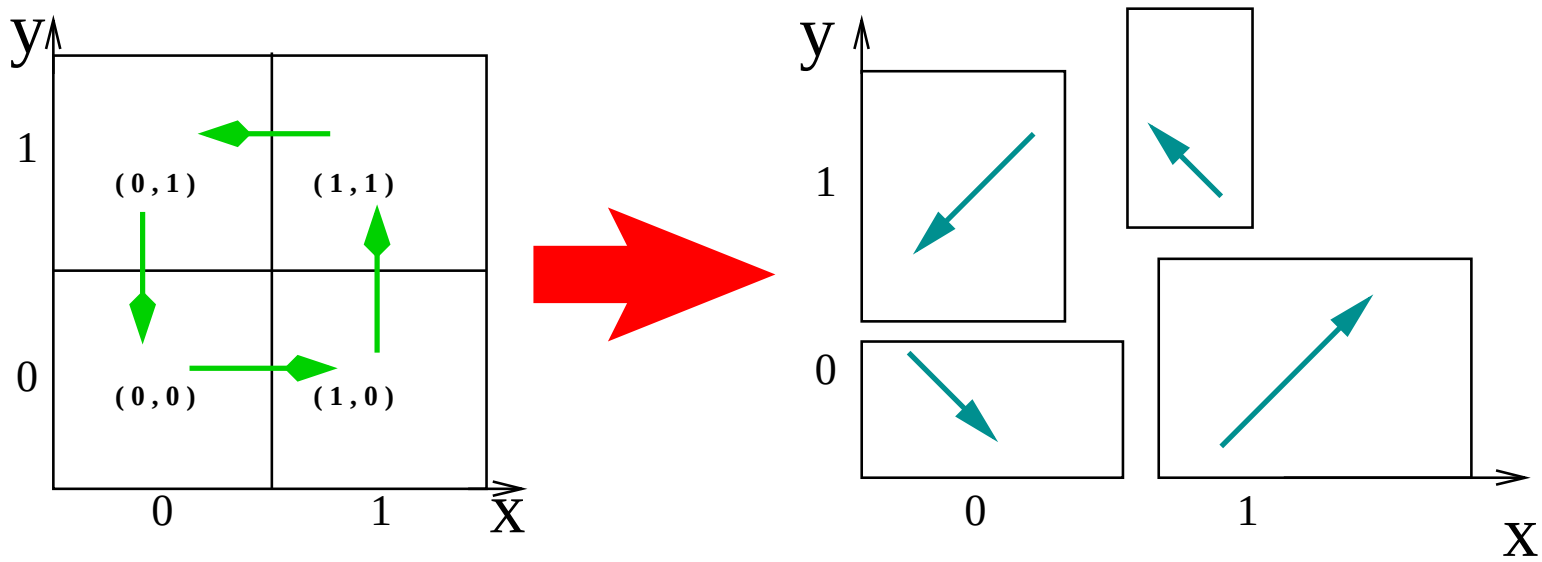
"Data Base" interaction graph



The target question

Impact of the day length on the persistence of the circadian circle ?

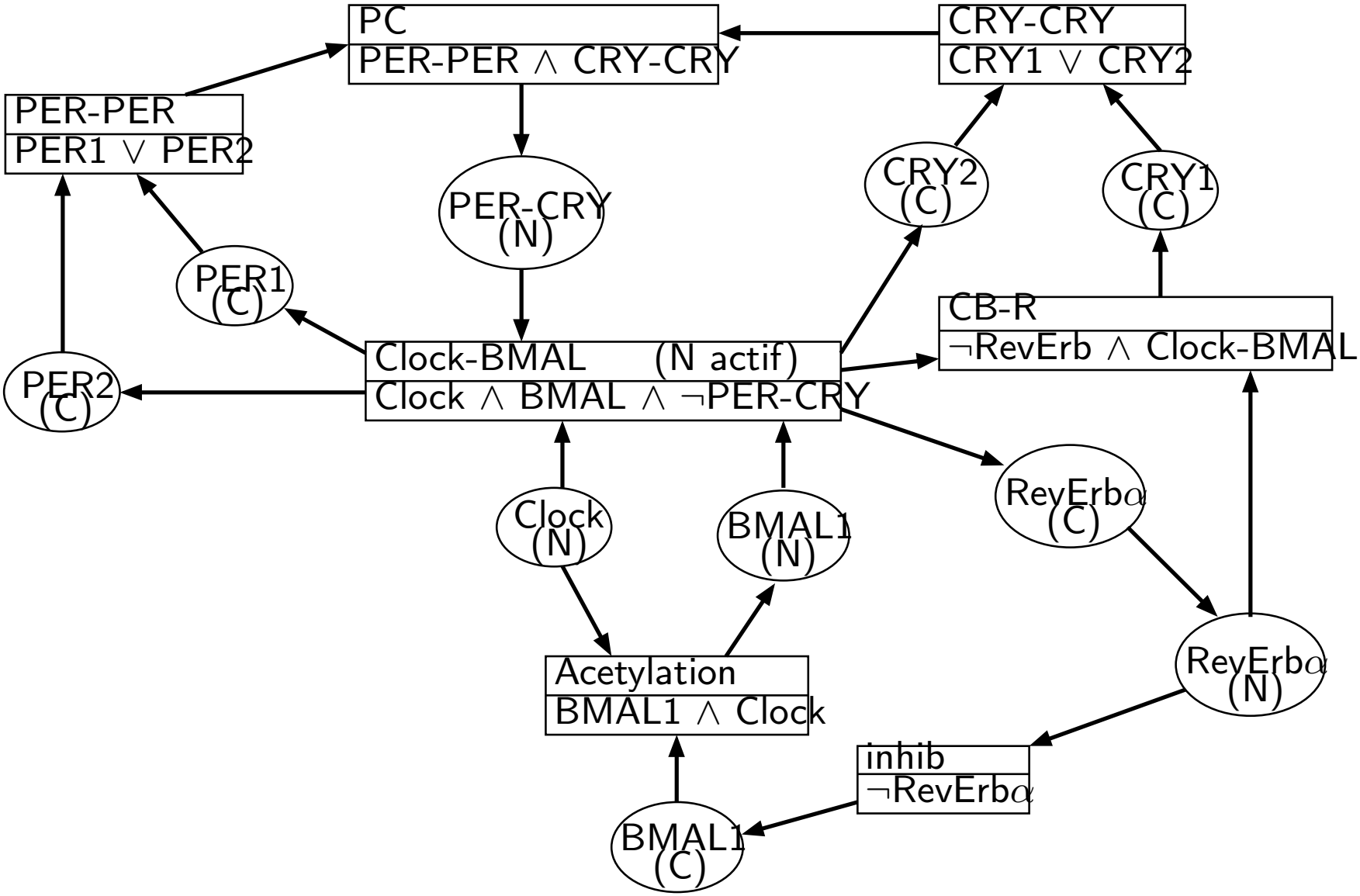
$\Longrightarrow$ framework with time delays:



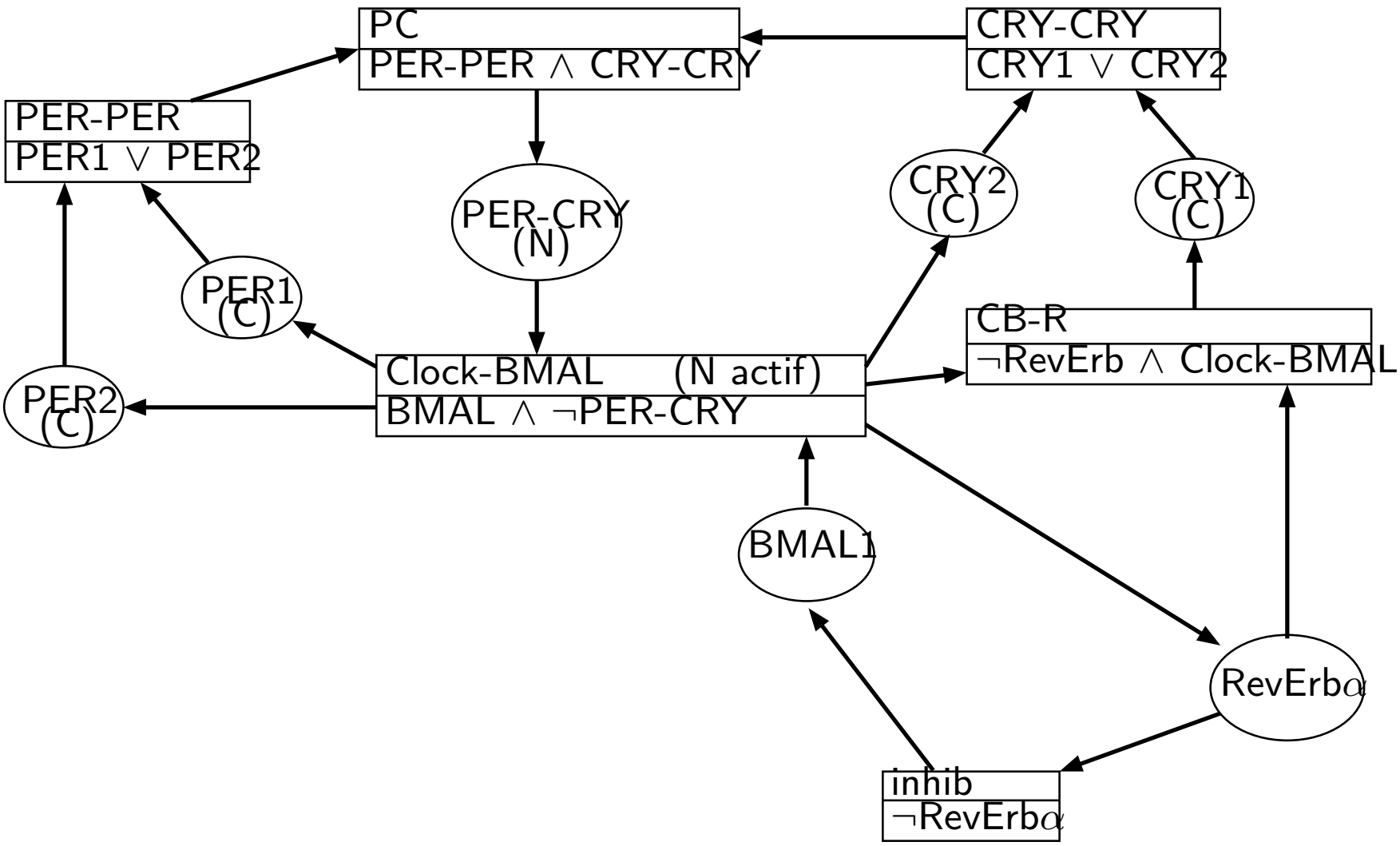
(size of rectangular areas = delays)

+ extension of temporal logic with delays. . .

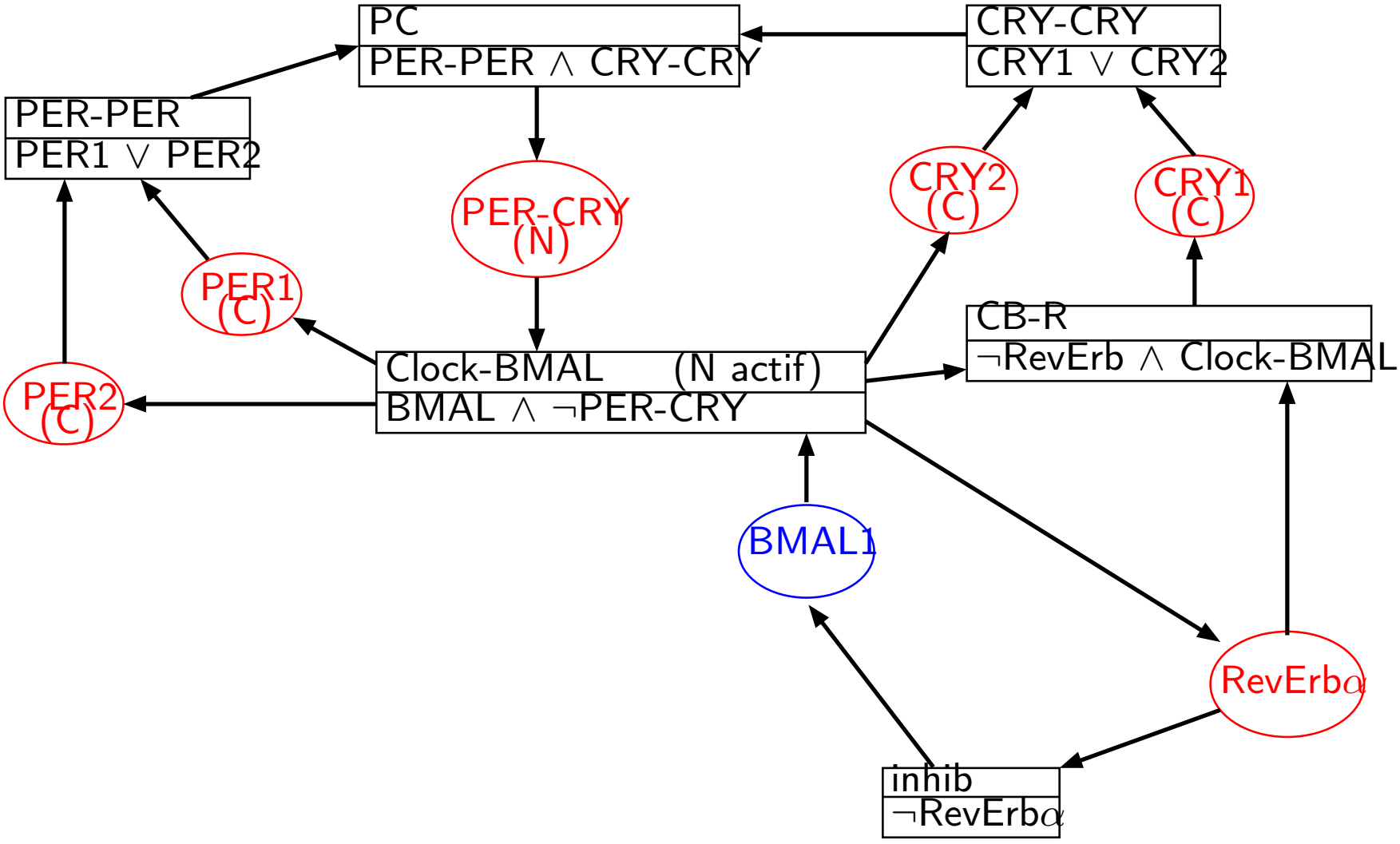
Fold levels and remove PPAR



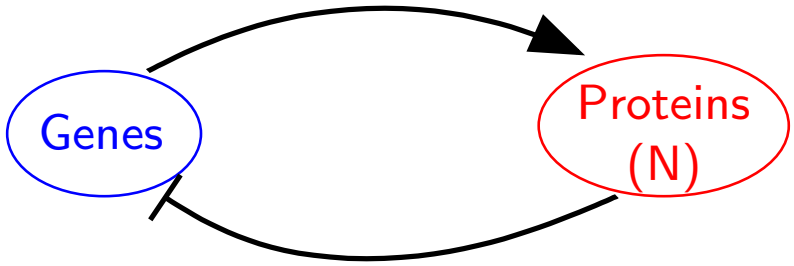
Remove Clock and "tunnel" pathways



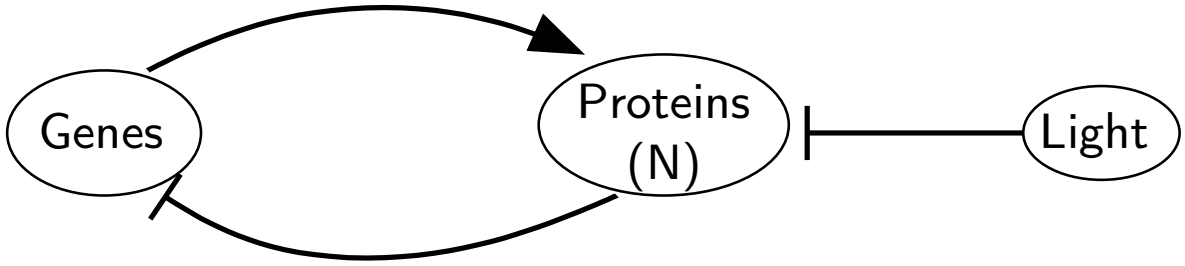
Separate inhibitors/activators of Clock-BMAL



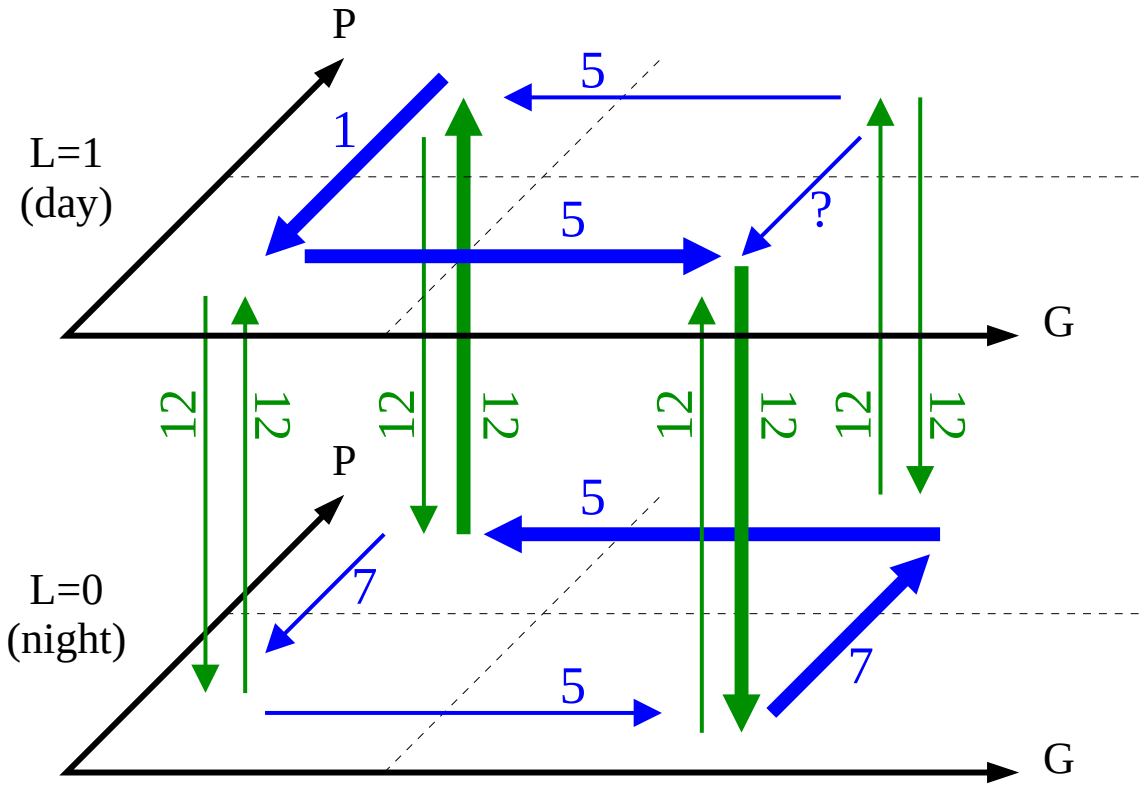
Fusion of all inhibitors



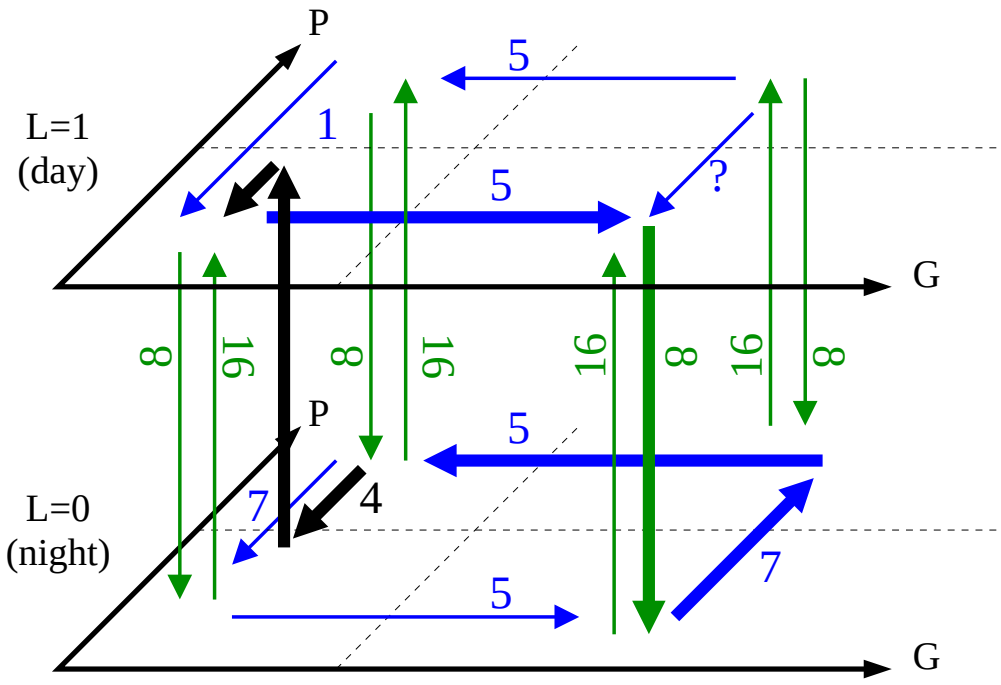
and Light prevents PER-CRY to enter the nucleus:



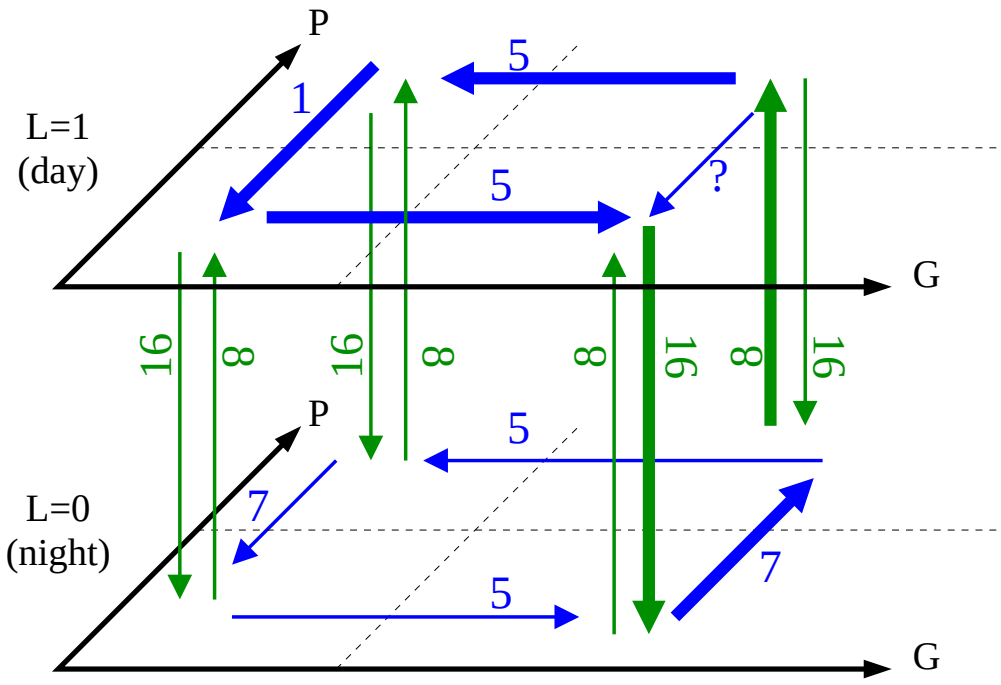
12 hours model



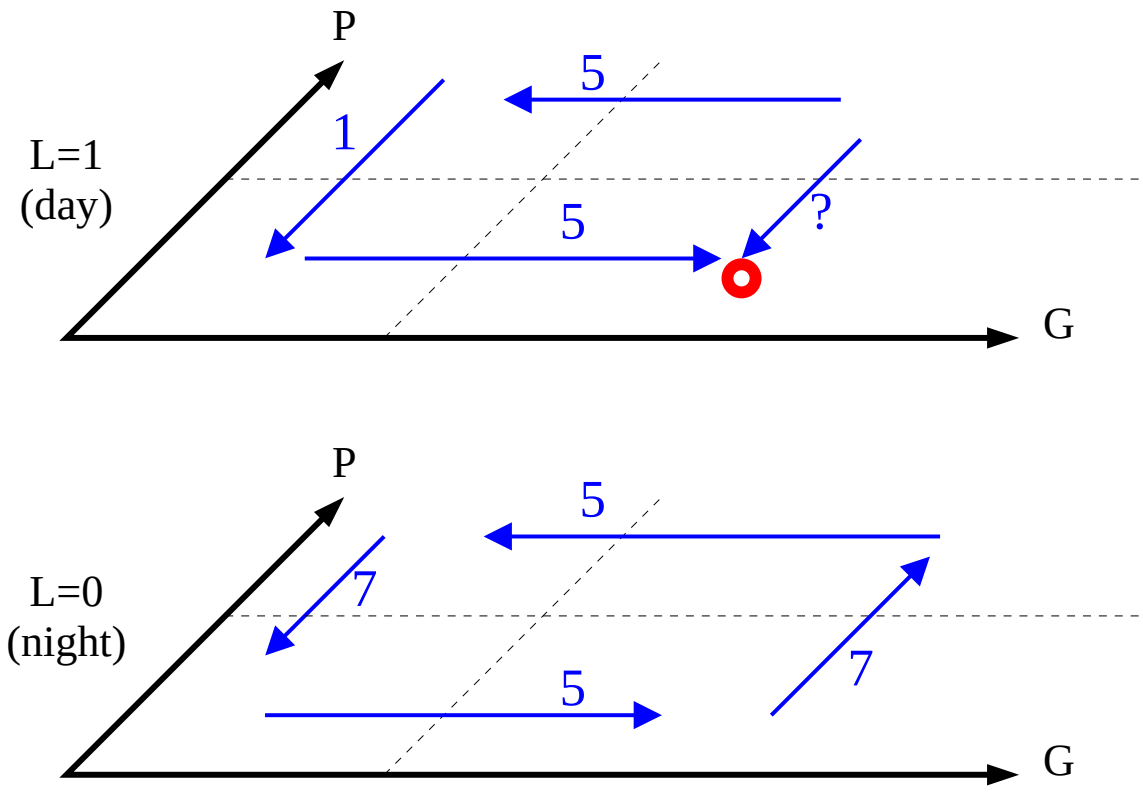
Winter model



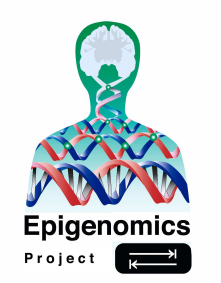
Summer model



Jet lag + training



Acknowledgements

- Epigenomics Project 

(Genopole®)
- Janine Guespin (Rouen)
- Franck Delaunay (Nice)
- Jean-Paul Comet (Nice)
- Camille Massot (Polytech, BIMB)
- Amélie Cessieux (Polytech, BIMB)

POLYTECH

Take Home Messages

Make explicit the hypotheses that motivate your research

A far as possible formalize them to get a computer aided approach

Behavioural *properties* are as much important as *models*

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

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

Formal proofs can suggest wet experiments

"Kleenex" models help understanding main behaviours