Property Driven Models: Experimental Validations and Simplifications

Gilles Bernot
University of Nice sophia antipolis, I3S laboratory, France
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

Syntax

- Formulae

Deduction

- Rules

Semantics

- Models

\[ M \models \varphi \]

satisfaction

green=Mathematics

gold=Computer

cyan=Computer Science

\[ \phi \vdash \varphi \]

proof

correctness

completeness

proved=satisfied
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)

\[ x' = x_0 + x_1 + x_2 \]

- No help: \( K_x \)
- \( x \) helps: \( K_{x,x} \)
- Absent \( y \) helps: \( K_{x,y} \)
- Both: \( K_{x,x,y} \)

<table>
<thead>
<tr>
<th>( (x,y) )</th>
<th>Focal Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0)</td>
<td>( (K_x, y, K_y) )</td>
</tr>
<tr>
<td>(0,1)</td>
<td>( (K_x, K_y) )</td>
</tr>
<tr>
<td>(1,0)</td>
<td>( (K_x, x, K_y) )</td>
</tr>
<tr>
<td>(1,1)</td>
<td>( (K_x, x, K_y) )</td>
</tr>
<tr>
<td>(2,0)</td>
<td>( (K_x, x, y, K_y) )</td>
</tr>
<tr>
<td>(2,1)</td>
<td>( (K_x, x, K_y, x) )</td>
</tr>
<tr>
<td>$(x,y)$</td>
<td>Focal Point</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>(0,0)</td>
<td>$(K_x, y, K_y) = (0,1)$</td>
</tr>
<tr>
<td>(0,1)</td>
<td>$(K_x, x, K_y) = (2,1)$</td>
</tr>
<tr>
<td>(1,0)</td>
<td>$(K_{x,x}, K_y) = (2,1)$</td>
</tr>
<tr>
<td>(1,1)</td>
<td>$(K_{x,x}, K_{y,x}) = (2,1)$</td>
</tr>
<tr>
<td>(2,0)</td>
<td>$(K_{x,x}, K_{y,x}) = (2,1)$</td>
</tr>
<tr>
<td>(2,1)</td>
<td>$(K_{x,x}, K_{y,x}) = (2,1)$</td>
</tr>
</tbody>
</table>

“desynchronization” → by units of Manhattan distance
CTL = Computation Tree Logic

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

Logical connectives: (\varphi_1 \land \varphi_2) (\varphi_1 \implies \varphi_2) ...

Temporal modalities: made of 2 characters

<table>
<thead>
<tr>
<th>first character</th>
<th>second character</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = for All path choices</td>
<td>X = neXt state</td>
</tr>
<tr>
<td>E = there Exist a choice</td>
<td>F = for some Future state</td>
</tr>
<tr>
<td></td>
<td>G = for all future states (Globally)</td>
</tr>
<tr>
<td></td>
<td>U = Until</td>
</tr>
</tbody>
</table>

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.
Common properties:
“functionality” of a sub-graph
Special role of “feedback loops”
– positive: multistationnarity (even number of — )
– negative: homeostasy (odd number of — )

Characteristic properties: \[
\begin{align*}
(x = 2) & \implies AG(\neg(x = 0)) \\
(x = 0) & \implies AG(\neg(x = 2))
\end{align*}
\]
They express “the positive feedback loop is functional”
(satisfaction of these formulae relies on the parameters $K$...
The Two Questions

\[ \Phi = \{ \varphi_1, \varphi_2, \cdots, \varphi_n, H \} \text{ and } M = \cdots \]

1. Is it possible that \( \Phi \) and \( M \)?

**Consistency** of knowledge and hypotheses. Means to select models belonging to the schemas that satisfy \( \Phi \).

\( (\exists \ M \in M \mid M \models \Phi) \)

2. If so, is it true *in vivo* that \( \Phi \) and \( M \)?

Compatibility of one of the selected models with the biological object. Require to propose experiments to **validate** or **refute** the selected model(s).

\[ \rightarrow \text{ Computer aided *proofs and validations*} \]
Theoretical Models ↔ 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 ($M \models_{\text{Model Checking}} \varphi$)
- They can be tested against the biological experiments ($\text{Biological Object} \models_{\text{Experiment}} \varphi$)

CTL is a bridge between theoretical models and biological objects
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
Set of all the formulae:

\[ \varphi = \text{hypothesis} \]
Set of all the formulae:

\[ \varphi = \text{hypothesis} \]
\[ Obs = \text{possible experiments} \]
Set of all the formulae:

$\varphi = \text{hypothesis}$  
$Obs = \text{possible experiments}$  
$Th(\varphi) = \varphi \text{ inferences}$
Set of all the formulae:

\( \varphi = \text{hypothesis} \)
\( \text{Obs} = \text{possible experiments} \)
\( Th(\varphi) = \varphi \text{ inferences} \)
\( S = \text{sensible experiments} \)
Set of all the formulae:

\( \varphi = \text{hypothesis} \)

\( Obs = \text{possible experiments} \)

\( Th(\varphi) = \varphi \text{ inferences} \)

\( S = \text{sensible experiments} \)

Refutability:

\( S \implies \varphi ? \)
Generation of biological experiments

Set of all the formulae:

\[ \varphi = \text{hypothesis} \]
\[ Obs = \text{possible experiments} \]
\[ Th(\varphi) = \varphi \text{ inferences} \]
\[ S = \text{sensible experiments} \]

Refutability:
\[ S \implies \varphi ? \]

Best refutations:
Choice of experiments in \( S \) ?
\[ \ldots \text{optimisations} \]
How to validate a multistationnarity

Hypotheses:

\[ \begin{align*}
(\text{Alginate} = 2) & \implies AG(\text{Alginate} = 2) \\
(\text{Alginate} = 0) & \implies AG(\text{Alginate} < 2)
\end{align*} \]

Assume that only \textit{mucus} can be observed:

\textbf{Lemma: } \text{AG}(\text{Alginate} = 2) \iff \text{AFAG}(\text{mucus} = 1)

(\ldots formal proof by computer \ldots)

\rightarrow \text{To validate: } (\text{Alginate} = 2) \implies AXAG(\text{mucus} = 1)
(Alginate = 2) \iff AXAG(mucus = 1)

<table>
<thead>
<tr>
<th>A \iff B</th>
<th>true</th>
<th>false</th>
</tr>
</thead>
<tbody>
<tr>
<td>true</td>
<td>true</td>
<td>false</td>
</tr>
<tr>
<td>false</td>
<td>true</td>
<td>true</td>
</tr>
</tbody>
</table>

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:

\[(\text{something reachable}) \iff (\text{Alginate} = 2)\]

General form of a test:

\[(\text{something reachable}) \iff (\text{something observable})\]
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:

Model $M_1$ satisfies $\varphi_1 \iff$ Model $M_2$ satisfies $\varphi_2 \iff$ Model $M_3$ satisfies $\varphi_3 \iff \ldots$
Simplifications via level folding

\[ \rho_y = 0 \quad \rho_y = 1 \quad \rho_y = 2 \]

\[ \rho_x = 0 \quad \rho_x = 1 \quad \rho_x = 2 \]
Embeddings of Regulatory Networks:

Simplifications via subgraphs

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

... *Also fusion of genes, etc.*
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
The target question

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

⇒ framework with time delays:

(size of rectangular areas = delays)

+ extension of temporal logic with delays...
Fold levels and remove PPAR

PER-PER
PER1 \lor PER2

PER-CRY
PER1 \land CRY

PER-CRY (N)

Clock-BMAL (N actif)

Clock \land BMAL \land ¬PER-CRY

BMAL1 \land Clock

Acetylation

¬RevErb(\alpha)(C)

¬RevErb(\alpha)(N)

RevErb(\alpha)(N)

RevErb(\alpha)(C)

CRY1 (C)
CRY2 (C)

CRY (C)

CB-R

¬RevErb(\alpha) \land Clock-BMAL

Clock (N)
BMAL (N)
Remove Clock and “tunnel” pathways

\[ \text{PER-CRY} \land \neg \text{RevErb} \land \text{Clock-BMAL} \]

\[ \text{PC} \quad \text{PER-PER} \land \text{CRY-CRY} \]

\[ \text{CRY-CRY} \quad \text{CRY1} \lor \text{CRY2} \]

\[ \text{PER1} \lor \text{PER2} \]

\[ \text{PER} \quad \text{PER-PER} \land \text{CRY-CRY} \]

\[ \text{ Cry2} \quad \text{ Cry1} \]

\[ \text{CB-R} \quad \neg \text{RevErb} \land \text{Clock-BMAL} \]

\[ \text{BMAL1} \]

\[ \text{RevErb} \]

\[ \text{inhib} \quad \neg \text{RevErb} \]
Separate inhibitors/activators of Clock-BMAL

PC

PER-PER \land CRY-CRY

CRY-CRY

CRY1 \lor CRY2

CRY1 (C)  CRY2 (C)

CB-R

\neg RevErb \land Clock-BMAL

BMAL \land \neg PER-CRY

Clock-BMAL (N actif)

PER1 (N)

PER1 (C)

PER2 (C)

PER-PER

PER1 \lor PER2

PER-PER

\neg RevErb

RevErbα

BMAL1

inhib

\neg RevErbα
Fusion of all inhibitors

and Light prevents PER-CRY to enter the nucleus:
12 hours model

L=1 (day)

L=0 (night)
Winter model

L = 1 (day)

L = 0 (night)

G

P
Summer model

L=1 (day)

L=0 (night)
Jet lag + training

L = 1 (day)

L = 0 (night)
Yet far from automatic simplifications but...

Abstract interpretation at INRIA

Model reduction at sobios
Acknowledgements

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

Option BIMB en Génie Biologique
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