Towards an automatic conversion from SBML core to SBML qual

Athénaïs Vaginay, Malika Smail-Tabbone, Taha Boukhobza

Towards an automatic conversion from SBML core to SBML qual

Athénaïs Vaginay[1, 2], Malika Smail-Tabbone[1], Taha Boukhobza[2]
[1]LORIA (Université de Lorraine, CNRS, Inria); [2]CRAN (Université de Lorraine, CNRS)
Contact: <athenais.vaginay@inria.fr>

Abstract: The SBML format is the *de facto* standard to encode biological model in different formalisms. It was first developed to encode quantitative models like Differentials Equations (DEs), but the current release allows the definition of packages extending the core format. One of these packages, named qual, encodes qualitative models like Boolean Networks (BNs). To the best of our knowledge, there is no automatic pipeline to convert a quantitative model encoded in SBML core into a qualitative model encoded with the qual package. Here, we explore such a pipeline on a relatively simple system: the cell division of fission yeast, which has been studied both with a set DEs [1] and with a BN [2]. Our approach consists in extracting the model topology from the set of DEs and in solving them numerically in order to retrieve the time course data of species' concentrations on which we apply a discretization. Then we extract from these data a scarce state transition table. We are currently investigating ways to synthesize a BN fitting both topology knowledge and state transitions.

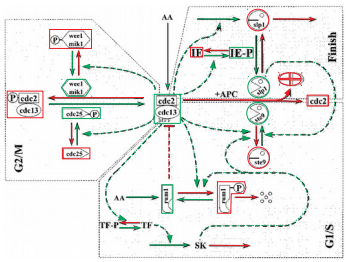
Quantitative model (e.g. DE)

$$\begin{cases} da/dt = \alpha(t,a,b,c,d) \\ db/dt = \beta(t,a,b,c,d) \\ dc/dt = \gamma(t,a,b,c,d) \\ dd/dt = \delta(t,a,b,c,d) \end{cases}$$

From [1] :
9 differential equations
+ 6 algebraic equations
+ parameters
+ 2 events

Biological system

E.g. the cell division cycle of wild type fission yeast



From [1]:

wiring diagram of the agents involved in the system.

Qualitative models (e.g. BN)

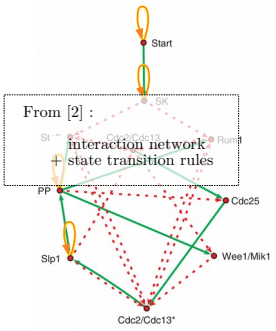
$$G(X,F); F = \begin{cases} x_a(t+1) = f_a(X(t)) \\ x_b(t+1) = f_b(X(t)) \\ x_c(t+1) = f_c(X(t)) \\ x_d(t+1) = f_d(X(t)) \end{cases}$$

$$x_i(t+1) \begin{cases} 0, & \sum_j a_{ij}x_j(t) > \theta_i \\ 1, & \sum_j a_{ij}x_j(t) < \theta_i \\ x_i(t), & \sum_j a_{ij}x_j(t) = \theta_i \end{cases}$$

θ_i=1 for Cdc2/Cdc13* and Cdc2/Cdc13
otherwise 0

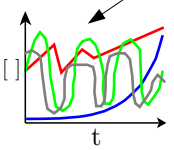
model agents
$j \to i$ positive interaction ($a_{ij}=1$)
$j \dashrightarrow i$ negative interaction ($a_{ij}=-1$)
$i \to i$ self-degradation ($a_{ii}=1$)

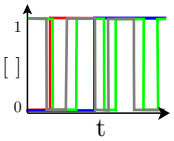
From [2] :
interaction network
+ state transition rules



SBML core → SBML qual (Encoding (LibSBML))

Simulation (COPASI) → Discretisation (BoolNet)

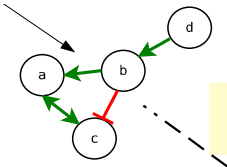




	X(t)	X(t+1)
	abcd	abcd
1	1000	0101
2	1010	1001
	...	
m	1000	0011

State Transition Table (STT)

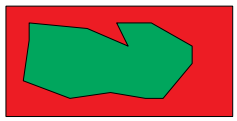
Extraction → use knowledge about topology



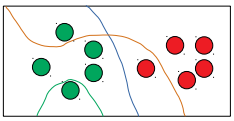
G → Model selection → $G_1, G_2, ..., G_x$
Possibility of multiple models

boolean function synthesis

Case 1: Complete consistent STT
→ Convertor to canonical boolean formula (BioLQM)

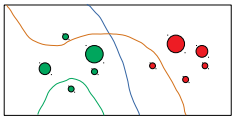


Case 2: Incomplete consistent STT
→ REVEAL (BoolNet)

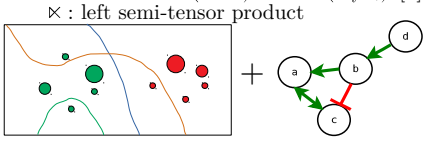


(weighted) False examples
(weighted) True examples
boolean classifier

Case 3: Incomplete inconsistent STT
→ Lähdesmäki's best-fit extension (BoolNet)



Case 4: Incomplete inconsitent STT + topology
→ Find M such that $X(t+1) = M \ltimes (\ltimes_i^n x_i)$ [3]
$\ltimes$: left semi-tensor product



Next objectives:
- Use the knowledge about topology derived from the kinetic model to synthetize boolean functions
- Define pipeline validation criteria such as attractor identification and matching
- Assess the pipeline on other models
- Fully automatize the pipeline → to be submitted to SBFC convertor

References:
[1] Novak *et al.* 'Mathematical model of the cell division cycle of fission yeast', Chaos, Mar. 2001.
[2] Davidich and Bornholdt, 'Boolean network model predicts cell cycle sequence of fission yeast', PLoS ONE, Feb. 2008.
[3] Leifeld *et al.* 'Identification of Boolean Network Models From Time Series Data Incorporating Prior Knowledge', Front. Physiol., Jun. 2018.