Model Revision of Boolean Regulatory Networks at Stable State

Filipe Gouveia, Inês Lynce, and Pedro T. Monteiro

{filipe.gouveia,ines.lynce,pedro.tiago.monteiro}@tecnico.ulisboa.pt

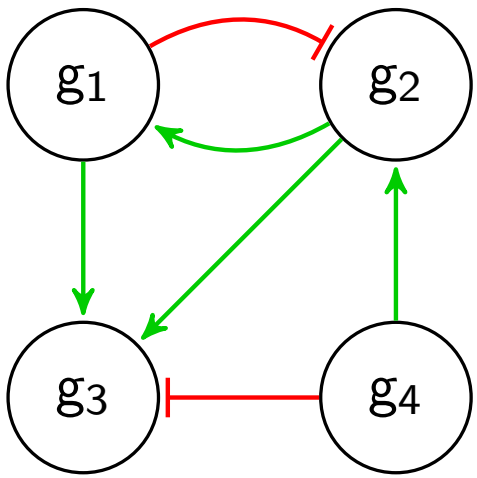
INESC-ID / Instituto Superior Técnico
Universidade de Lisboa
Lisbon, Portugal

ISBRA 2019

Table of Contents

Regulatory Networks

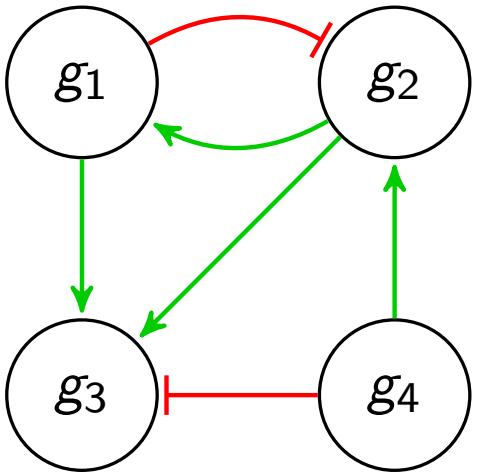
- Biological processes arise at the cellular level, governed by complex regulatory networks
- Regulatory network
  - Collection of molecular compounds (e.g. proteins, genes)
  - Compounds interact with each other
- Computational modelling allows
  - Functional understanding of the network
  - Test hypotheses
  - Identify predictions *in silico*
  - ...



- Different formalisms can be used [KS08]
  - We consider the Boolean logical formalism [Tho73].

Logical Model

- Compounds represented by a Boolean variable
  - active/inactive
  - 2^n possible states with n variables
- Interactions defined as positive (activation) or negative (inhibition)
- Regulations defined as Boolean functions
  - 2^{2^k} possible functions for a variable with k regulators



$$f_{g_1} = g_2$$

$$f_{g_2} = \neg g_1 \wedge g_4$$

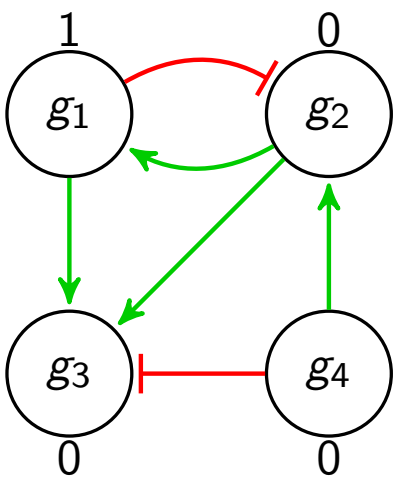
$$f_{g_3} = g_1 \vee (g_2 \wedge \neg g_4)$$

Motivation

- Constructing such biological models is still mainly a manual task
  - In particular the definition of regulatory effects
- As new data is acquired, models need to be revised or updated
  - Model revision

Motivation

- Constructing such biological models is still mainly a manual task
  - In particular the definition of regulatory effects
- As new data is acquired, models need to be revised or updated
  - Model revision



$f_{g_1} = g_2$

$f_{g_2} = \neg g_1 \wedge g_4$

$f_{g_3} = g_1 \vee (g_2 \wedge \neg g_4)$

How can the model be repaired?

- Changing a regulatory function?
  - 2^{2^k} possibilities for each node
- Changing the sign of an edge?
- Adding or removing a regulator?

There are $\approx 10^{24}$ possible combinations!
(65536 Boolean functions with 4 regulators)

Table of Contents

Model Revision

- A model is *consistent* if all of its nodes are consistent
  - Value of each node given by its regulatory function is equal to the observed value
- A model is *inconsistent* otherwise
  - Needs to be revised

Model Revision

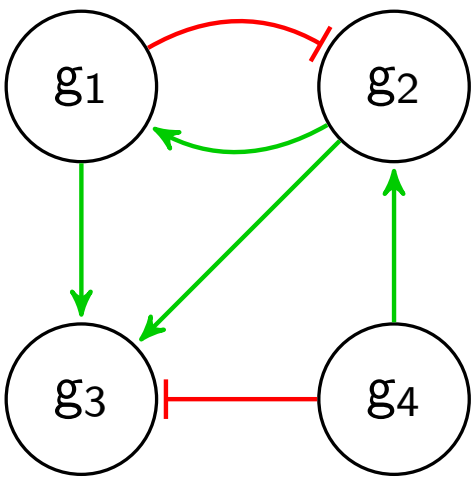
- A model is *consistent* if all of its nodes are consistent
  - Value of each node given by its regulatory function is equal to the observed value
- A model is *inconsistent* otherwise
  - Needs to be revised

Cause	Repair Operation	Class
Wrong Regulatory Function	Function change	F
Wrong Interaction Type	Edge sign flip	T
Wrong Regulator	Edge removal	T
Missing Regulator	Edge addition	T

- **F**unction repair
- **T**opology repair

Boolean Functions

- Only consider monotone nondegenerate Boolean functions
  - **Monotone**: each regulator only has one role: positive/activation or negative/inhibition
    - ★ each variable appears with the same sign in the function (DNF)

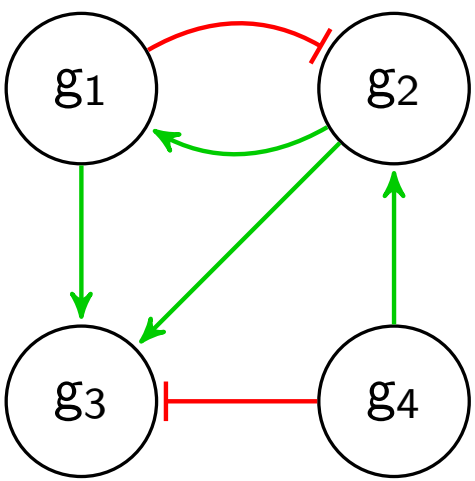


$f_{g_3} = g_1 \vee (g_2 \wedge \neg g_4)$ Monotone

$f_{g_3} = (g_1 \wedge \neg g_2) \vee (g_2 \wedge \neg g_4)$ Non-monotone

Boolean Functions

- Only consider monotone nondegenerate Boolean functions
  - **Monotone**: each regulator only has one role: positive/activation or negative/inhibition
    - ★ each variable appears with the same sign in the function (DNF)
  - **Nondegenerate**: each regulator is essential in the regulatory function
    - ★ each variable has an impact on the truth table



$f_{g_3} = g_1 \vee (g_2 \wedge \neg g_4)$ Nondegenerate

$f_{g_3} = (g_2 \wedge \neg g_4)$ (missing g_1) Degenerate

Boolean Functions

- Relation $\preceq$ between monotone nondegenerate Boolean functions [CMC19]

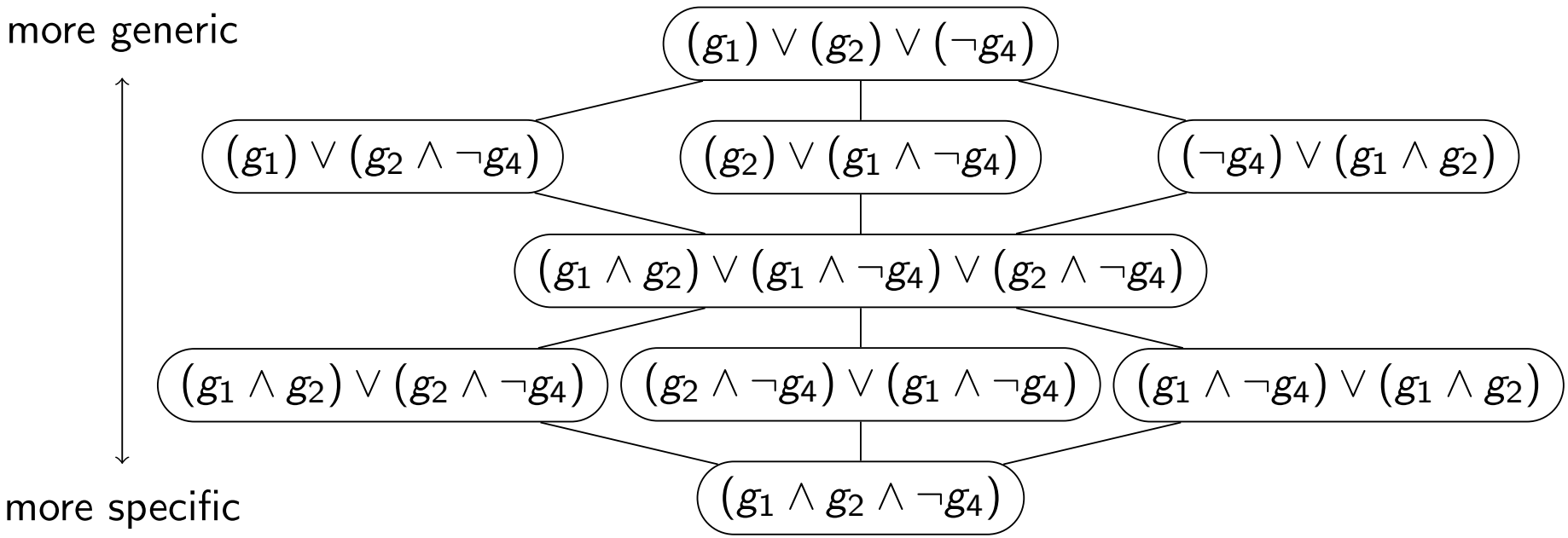
$$f \preceq f' \iff f(X) \Rightarrow f'(X).$$

  - where $f(X)$ denotes the entries where the function is **true** (or 1)

- f' is a **father** of f iff $f \preceq f'$ and $\nexists f''$ such that $f \preceq f''$ and $f'' \preceq f'$
  - f is a **child** of f'

Boolean Functions

Hasse Diagram



Function Repair

If a function is inconsistent

- Determine if is necessary to generalize or specify the function
- Compute set of fathers (children) to go up (down) the diagram
- Continue to do so until a consistent function is found
  - or no function is found

If function found, is closest to original function

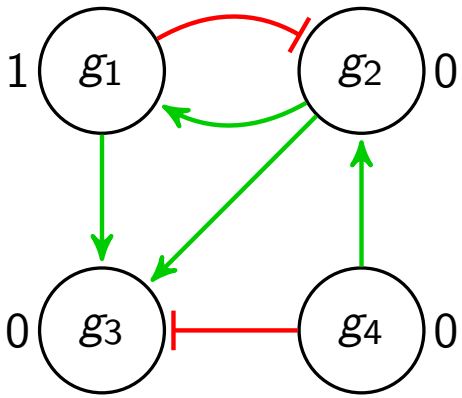
Topology Repair

Changing a regulatory function may not be sufficient to render a model consistent

- Flip sign of edge
  - Change the role of a regulator
- Remove edge
  - Remove a regulator from the regulatory function
- Add edge
  - Add a new regulator in the regulatory function

Model Revision

Example



$f_{g_1} = g_2$

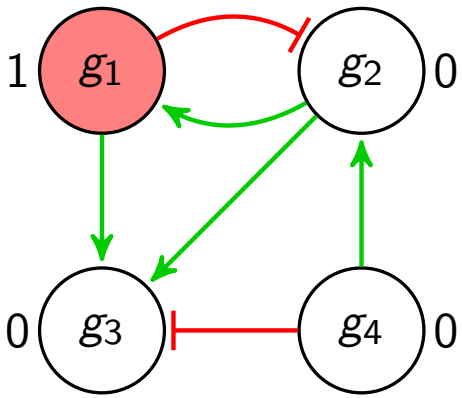
$f_{g_2} = \neg g_1 \wedge g_4$

$f_{g_3} = g_1 \vee (g_2 \wedge \neg g_4)$

- Inconsistent Model
- Node g_1 is inconsistent
- Change the interaction between g_2 and g_1
- Node g_3 is inconsistent
- Consistent Model

Model Revision

Example



$f_{g_1} = g_2$

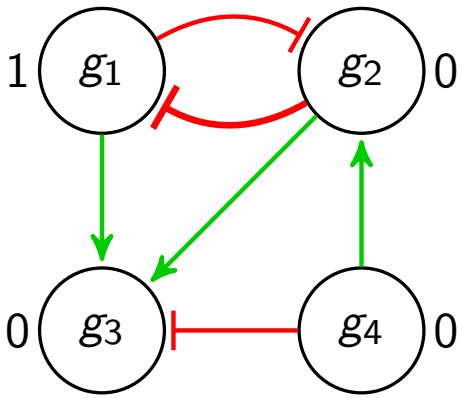
$f_{g_2} = \neg g_1 \wedge g_4$

$f_{g_3} = g_1 \vee (g_2 \wedge \neg g_4)$

- Inconsistent Model
- Node g_1 is inconsistent
- Change the interaction between g_2 and g_1
- Node g_3 is inconsistent
- Consistent Model

Model Revision

Example



$f_{g_1} = \neg g_2$

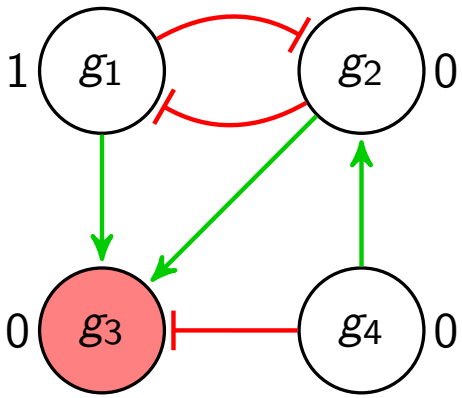
$f_{g_2} = \neg g_1 \wedge g_4$

$f_{g_3} = g_1 \vee (g_2 \wedge \neg g_4)$

- Inconsistent Model
- Node g_1 is inconsistent
- Change the interaction between g_2 and g_1
- Node g_3 is inconsistent
- Consistent Model

Model Revision

Example



$f_{g_1} = \neg g_2$

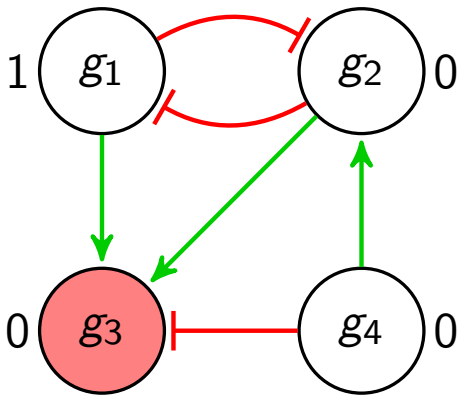
$f_{g_2} = \neg g_1 \wedge g_4$

$f_{g_3} = g_1 \vee (g_2 \wedge \neg g_4)$

- Inconsistent Model
- Node g_1 is inconsistent
- Change the interaction between g_2 and g_1
- Node g_3 is inconsistent
- Consistent Model

Model Revision

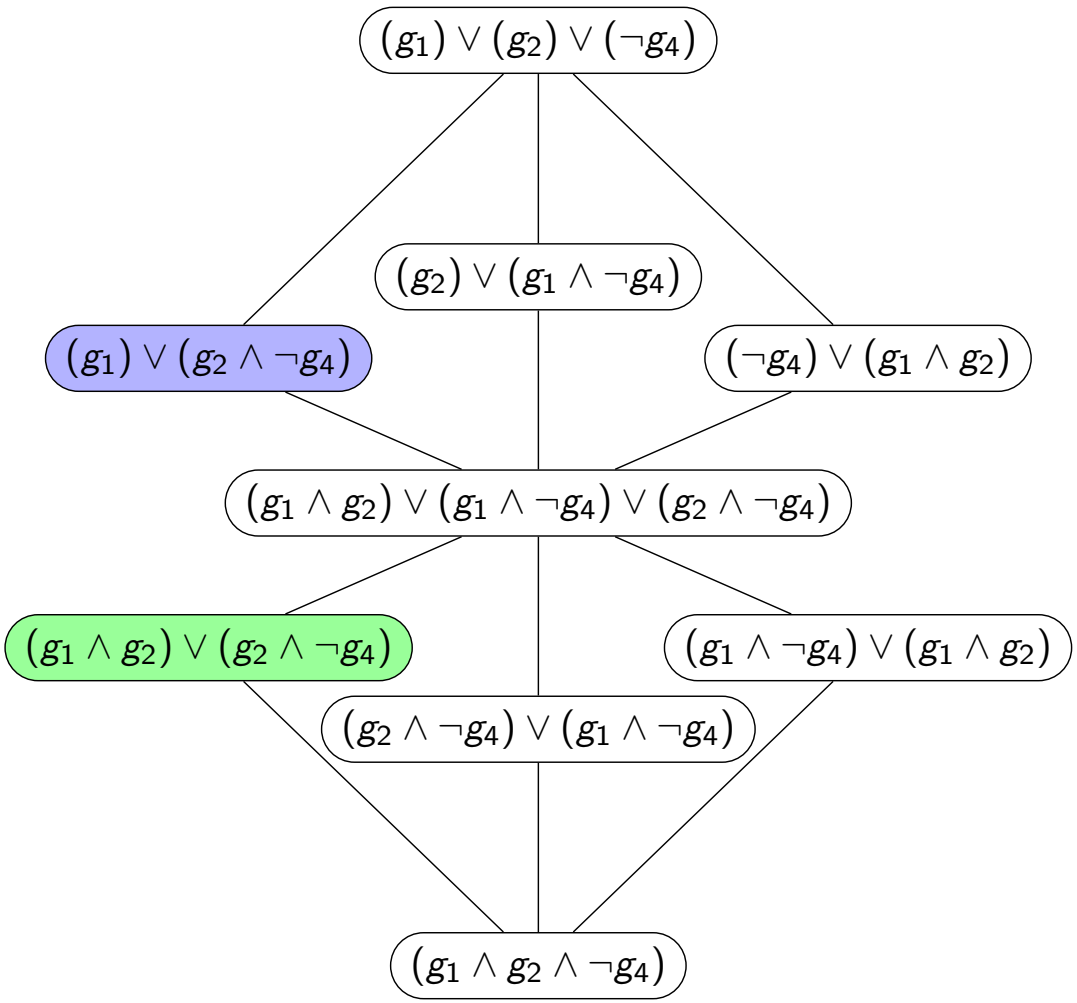
Example



$f_{g_1} = \neg g_2$

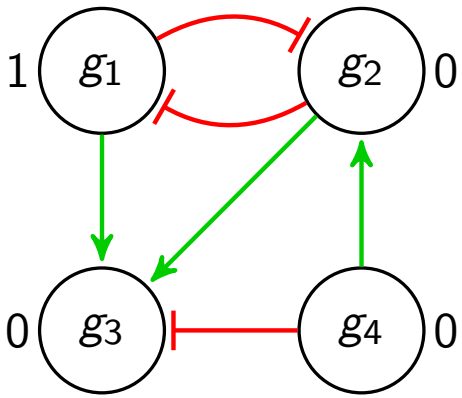
$f_{g_2} = \neg g_1 \wedge g_4$

$f_{g_3} = g_1 \vee (g_2 \wedge \neg g_4)$



Model Revision

Example



$f_{g_1} = \neg g_2$

$f_{g_2} = \neg g_1 \wedge g_4$

$f_{g_3} = (g_1 \wedge g_2) \vee (g_2 \wedge \neg g_4)$

- Inconsistent Model
- Node g_1 is inconsistent
- Change the interaction between g_2 and g_1
- Node g_3 is inconsistent
- Consistent Model

Table of Contents

Model Revision Approach

Assumptions

- Regulatory functions as monotone nondegenerate Boolean functions
- Only consider Stable State observations
- Higher level of confidence in the correctness of the network topology than in the regulatory functions of the model

Model Revision Approach

Consistency Check

Given a model and a set of observations we verify the consistency of the model

- Answer Set Program developed [GLM18]
- If the model is consistent no model revision necessary
- Otherwise, return the minimum number of inconsistent nodes
  - Observational data may not be complete
- Information regarding the reason of inconsistency is also retrieved
  - Need a more generic or specific node

Model Revision Approach

Optimization Criteria

Lexicographic optimization criteria defined

1. Minimize the number of add/remove edge operations
2. Minimize the number of flip sign of an edge operations
3. Minimize the number of change regulatory function operations

Model Revision Approach

An iterative approach was developed to repair an inconsistent node

- Try to change the function
- Try to flip the sign of 1 edge
  - Consider changing the function again
  - Repeat this step for 2 edges, and so on
- Try to add or remove 1 edge
  - Consider changing the function and/or flip the sign of edges as previously
  - Repeat this step for adding or removing 2 edges and so on

Model Revision Approach

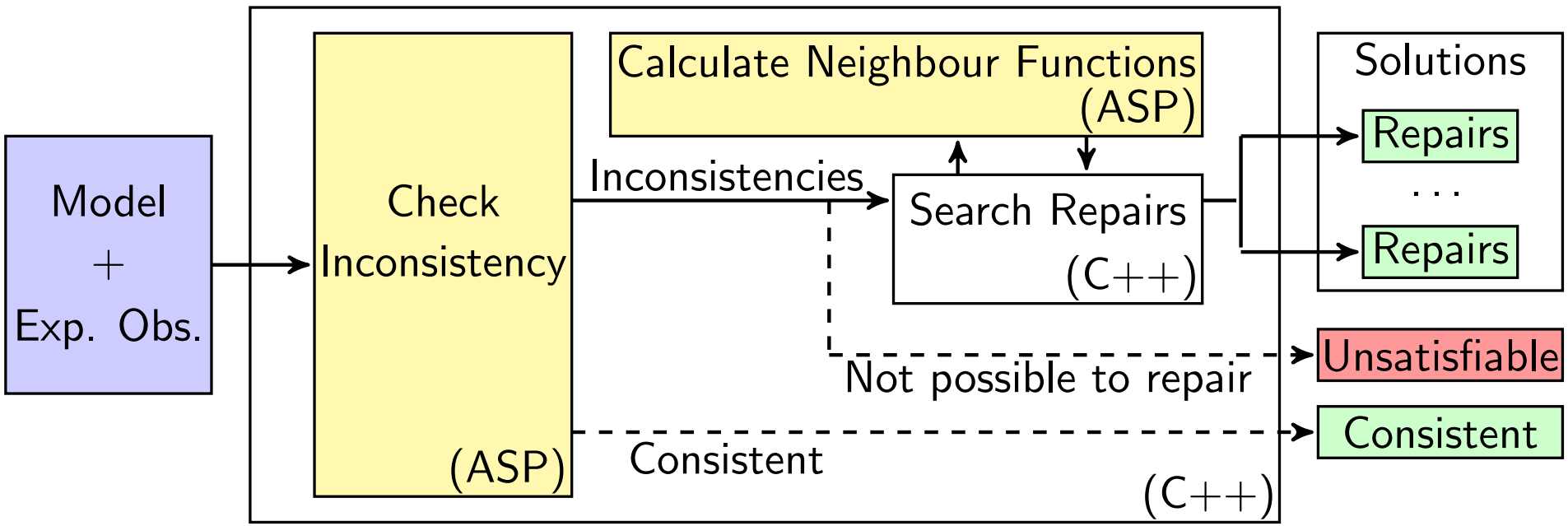


Table of Contents

Evaluation

- 5 well-known biological models considered

Abbr.	**Model**	**#N**	**#E**	**#SS**	**Reg.**	**Ref.**
FY	Fission Yeast	10	27	12	5	[DB08]
SP	Segment Polarity	19	57	7	8	[SCT02]
TCR	TCR Signalisation	40	57	7	5	[KSRL$^+$06]
MCC	Mammalian Cell Cycle	10	35	1	6	[FNCT06]
Th	Th Cell Differentiation	23	35	3	5	[MX06]

Evaluation

- Random changes were made according to probabilistic parameters
  - F% : Change a **F**unction
  - E% : Flip the sign of an **E**dge
  - R% : **R**emove an existing edge
  - A% : **A**dd a missing edge
- Several configurations of these parameters were considered
- 10 instances were generated for each configuration for each model
- Timeout of 600 seconds was considered

Evaluation

(%)				FY		SP		TCR		MCC		Th	
F	E	R	A	T (s)	#TO	T (s)	#TO	T (s)	#TO	T (s)	#TO	T (s)	#TO
5	0	0	0	0,034	0	0,036	0	0,047	0	0,021	0	0,028	0
25	0	0	0	0,059	0	4,734	0	0,063	0	0,021	0	0,061	0
50	0	0	0	0,060	0	14,003	2	0,097	0	0,033	0	0,677	0
100	0	0	0	0,072	0	18,937	2	0,129	0	0,046	0	0,751	0
0	5	0	0	0,070	0	0,105	0	0,050	0	0,033	0	0,061	1
0	10	0	0	0,070	0	1,566	1	0,050	0	0,101	0	0,044	0
0	15	0	0	0,035	1	0,168	3	0,050	0	0,039	0	0,051	1
0	20	0	0	0,071	2	0,284	4	0,050	0	0,136	0	0,062	1
0	0	1	0	0,034	0	0,635	0	0,045	0	0,020	0	0,025	1
0	0	5	0	0,069	0	5,021	1	0,046	0	0,020	0	0,026	1
0	0	10	0	0,095	2	24,481	4	0,060	0	0,019	0	0,589	2
0	0	15	0	0,083	2	32,896	3	7,106	0	0,029	0	1,613	2
0	0	0	1	0,874	0	0,130	2	0,152	0	0,020	0	0,028	3
0	0	0	5	0,096	0	42,684	7	2,518	3	0,219	0	0,497	8
0	0	0	10	0,842	1	-	10	-	10	0,234	1	-	10
0	0	0	15	6,003	4	-	10	-	10	0,622	0	258,022	9
25	5	0	0	0,062	0	5,358	0	0,063	0	0,032	0	0,108	0
50	25	0	0	0,127	2	13,989	4	0,187	0	0,570	0	0,724	1
5	25	5	5	0,453	4	-	10	3,979	8	0,549	1	0,781	9
10	10	5	5	0,601	2	24,637	8	50,662	6	0,142	1	0,745	7

Evaluation

(%)				FY		SP		TCR		MCC		Th	
F	E	R	A	T (s)	#TO	T (s)	#TO	T (s)	#TO	T (s)	#TO	T (s)	#TO
5	0	0	0	0,034	0	0,036	0	0,047	0	0,021	0	0,028	0
25	0	0	0	0,059	0	4,734	0	0,063	0	0,021	0	0,061	0
50	0	0	0	0,060	0	14,003	2	0,097	0	0,033	0	0,677	0
100	0	0	0	0,072	0	18,937	2	0,129	0	0,046	0	0,751	0
0	5	0	0	0,070	0	0,105	0	0,050	0	0,033	0	0,061	1
0	10	0	0	0,070	0	1,566	1	0,050	0	0,101	0	0,044	0
0	15	0	0	0,035	1	0,168	3	0,050	0	0,039	0	0,051	1
0	20	0	0	0,071	2	0,284	4	0,050	0	0,136	0	0,062	1
0	0	1	0	0,034	0	0,635	0	0,045	0	0,020	0	0,025	1
0	0	5	0	0,069	0	5,021	1	0,046	0	0,020	0	0,026	1
0	0	10	0	0,095	2	24,481	4	0,060	0	0,019	0	0,589	2
0	0	15	0	0,083	2	32,896	3	7,106	0	0,029	0	1,613	2
0	0	0	1	0,874	0	0,130	2	0,152	0	0,020	0	0,028	3
0	0	0	5	0,096	0	42,684	7	2,518	3	0,219	0	0,497	8
0	0	0	10	0,842	1	-	10	-	10	0,234	1	-	10
0	0	0	15	6,003	4	-	10	-	10	0,622	0	258,022	9
25	5	0	0	0,062	0	5,358	0	0,063	0	0,032	0	0,108	0
50	25	0	0	0,127	2	13,989	4	0,187	0	0,570	0	0,724	1
5	25	5	5	0,453	4	-	10	3,979	8	0,549	1	0,781	9
10	10	5	5	0,601	2	24,637	8	50,662	6	0,142	1	0,745	7

Evaluation

(%)				FY		SP		TCR		MCC		Th	
F	E	R	A	T (s)	#TO	T (s)	#TO	T (s)	#TO	T (s)	#TO	T (s)	#TO
5	0	0	0	0,034	0	0,036	0	0,047	0	0,021	0	0,028	0
25	0	0	0	0,059	0	4,734	0	0,063	0	0,021	0	0,061	0
50	0	0	0	0,060	0	14,003	2	0,097	0	0,033	0	0,677	0
100	0	0	0	0,072	0	18,937	2	0,129	0	0,046	0	0,751	0
0	5	0	0	0,070	0	0,105	0	0,050	0	0,033	0	0,061	1
0	10	0	0	0,070	0	1,566	1	0,050	0	0,101	0	0,044	0
0	15	0	0	0,035	1	0,168	3	0,050	0	0,039	0	0,051	1
0	20	0	0	0,071	2	0,284	4	0,050	0	0,136	0	0,062	1
0	0	1	0	0,034	0	0,635	0	0,045	0	0,020	0	0,025	1
0	0	5	0	0,069	0	5,021	1	0,046	0	0,020	0	0,026	1
0	0	10	0	0,095	2	24,481	4	0,060	0	0,019	0	0,589	2
0	0	15	0	0,083	2	32,896	3	7,106	0	0,029	0	1,613	2
0	0	0	1	0,874	0	0,130	2	0,152	0	0,020	0	0,028	3
0	0	0	5	0,096	0	42,684	7	2,518	3	0,219	0	0,497	8
0	0	0	10	0,842	1	-	10	-	10	0,234	1	-	10
0	0	0	15	6,003	4	-	10	-	10	0,622	0	258,022	9
25	5	0	0	0,062	0	5,358	0	0,063	0	0,032	0	0,108	0
50	25	0	0	0,127	2	13,989	4	0,187	0	0,570	0	0,724	1
5	25	5	5	0,453	4	-	10	3,979	8	0,549	1	0,781	9
10	10	5	5	0,601	2	24,637	8	50,662	6	0,142	1	0,745	7

Evaluation

(%)				FY		SP		TCR		MCC		Th	
F	E	R	A	T (s)	#TO	T (s)	#TO	T (s)	#TO	T (s)	#TO	T (s)	#TO
5	0	0	0	0,034	0	0,036	0	0,047	0	0,021	0	0,028	0
25	0	0	0	0,059	0	4,734	0	0,063	0	0,021	0	0,061	0
50	0	0	0	0,060	0	14,003	2	0,097	0	0,033	0	0,677	0
100	0	0	0	0,072	0	18,937	2	0,129	0	0,046	0	0,751	0
0	5	0	0	0,070	0	0,105	0	0,050	0	0,033	0	0,061	1
0	10	0	0	0,070	0	1,566	1	0,050	0	0,101	0	0,044	0
0	15	0	0	0,035	1	0,168	3	0,050	0	0,039	0	0,051	1
0	20	0	0	0,071	2	0,284	4	0,050	0	0,136	0	0,062	1
0	0	1	0	0,034	0	0,635	0	0,045	0	0,020	0	0,025	1
0	0	5	0	0,069	0	5,021	1	0,046	0	0,020	0	0,026	1
0	0	10	0	0,095	2	24,481	4	0,060	0	0,019	0	0,589	2
0	0	15	0	0,083	2	32,896	3	7,106	0	0,029	0	1,613	2
0	0	0	1	0,874	0	0,130	2	0,152	0	0,020	0	0,028	3
0	0	0	5	0,096	0	42,684	7	2,518	3	0,219	0	0,497	8
0	0	0	10	0,842	1	-	10	-	10	0,234	1	-	10
0	0	0	15	6,003	4	-	10	-	10	0,622	0	258,022	9
25	5	0	0	0,062	0	5,358	0	0,063	0	0,032	0	0,108	0
50	25	0	0	0,127	2	13,989	4	0,187	0	0,570	0	0,724	1
5	25	5	5	0,453	4	-	10	3,979	8	0,549	1	0,781	9
10	10	5	5	0,601	2	24,637	8	50,662	6	0,142	1	0,745	7

Evaluation

Results

- Tool repaired the model with less or equal number of operations
- Models repaired mostly under 60 seconds
- Changing the topology of the network has greater impact
  - Number of timeouts increased with the number of topological changes
- Addition of new edges has bigger impact than removal of edges
  - Increase the search space for possible function repairs

Table of Contents

Conclusions

- Proposed a Model Revision tool considering models stable states
- Use logic-based tool (Answer Set Programming)
  - Verify the consistency of a model
  - Compute neighbour functions
- C++ procedure to search for set of repair operations
- Tool produces all the optimum sets repair operations under defined optimization criteria

Conclusions

- Tool successfully tested using well-known biological models
- Most instances repaired under 60 seconds
- The dimension of the regulatory functions has the biggest impact on the tool performance
  - Number of monotone nondegenerate Boolean functions **increases double exponentially**

Future Work

- Consider the model dynamics using time-series data
- Repair inconsistent nodes with multiple causes for inconsistency
  - Need to be more specific and generic at the same time
  - Possible non-comparable function
- Heuristics could be used to reduce the number of solutions presented
- Study the solutions produced to determine common repair operations

Thank you!

Questions

Acknowledgments:

inesc id lisboa

TÉCNICO LISBOA

FCT Fundação para a Ciência e a Tecnologia

References I


José ER Cury, Pedro T Monteiro, and Claudine Chaouiya, *Partial Order on the set of Boolean Regulatory Functions*, arXiv preprint arXiv:1901.07623 (2019).


Maria I. Davidich and Stefan Bornholdt, *Boolean network model predicts cell cycle sequence of fission yeast*, PLoS ONE **3** (2008), no. 2, e1672.


A. Fauré, A. Naldi, C. Chaouiya, and D. Thieffry, *Dynamical analysis of a generic boolean model for the control of the mammalian cell cycle*, Bioinformatics **22** (2006), no. 14, e124–e131.


Filipe Gouveia, Inês Lynce, and Pedro T Monteiro, *Model revision of logical regulatory networks using logic-based tools*, ICLP 2018 (Technical Communications), Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.


Guy Karlebach and Ron Shamir, *Modelling and analysis of gene regulatory networks*, Nature Reviews Molecular Cell Biology **9** (2008), no. 10, 770.


Steffen Klamt, Julio Saez-Rodriguez, Jonathan A Lindquist, Luca Simeoni, and Ernst D Gilles, *A methodology for the structural and functional analysis of signaling and regulatory networks*, BMC Bioinformatics **7** (2006), no. 1, 56.


Luis Mendoza and Ioannis Xenarios, *A method for the generation of standardized qualitative dynamical systems of regulatory networks*, Theor. Biol. Med. Model. **3** (2006), no. 1, 13.


Lucas Sánchez, Claudine Chaouiya, and Denis Thieffry, *Segmenting the fly embryo: logical analysis of the role of the segment polarity cross-regulatory module*, Int. J. Dev. Biol. **52** (2002), no. 8, 1059–1075.


René Thomas, *Boolean formalization of genetic control circuits*, J. Theor. Biol. **42** (1973), no. 3, 563–585.