Teatime capsid:

Constrained Enumeration of Boolean Networks from Biological Data and Knowledge

A paper submission for OLA2021

Athénaïs Vaginay

2021 jan 19th (bonne année)

OLA2021

International Conference in
Optimization and Learning and their **Application**

Catania (Sicilia), Italy Jun 21-23 2021

Extended submission deadline : **Feb 05, 2020** (cross your fingers)

`https://ola2021.sciencesconf.org/`

Modeling of biological systems with Boolean Networks

Perfect for systems (even big ones) where the exact reaction times is not of central interest.

Generally built from experimental data and knowledge from the literature either manually or by using programs.

The automatic synthesis of BNs from biological data and knowledge is still a challenge.

BN of n component — formaly

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

BN of n component — formaly

in the Boolean world:
$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

1 components $\rightarrow$ 2 configurations:
0; 1

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

2 components $\rightarrow$ 4 configurations:
00; 01; 10; 11

BN of n component — formaly

in the Boolean world:
$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

3 components $\rightarrow$ 8 configurations:
000; 001; 010; 011; 100; 101; 110; 111

BN of n component — formaly

in the Boolean world:
$\mathbb{B} = \{0, 1\}$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

4 components $\rightarrow$ 16 configurations:
0000; 0001; 0010; 0011; 0100; 0101; 0110; 0111;
1000; 1001; 1010; 1011; 1100; 1101; 1110; 1111

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

A **local update function** is associated to each component

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

A **local update function** is associated to each component

$$\mathscr{B} = (f_i : \mathbb{B}^{k_i \leq n} \to \mathbb{B} \ \forall i \in V)$$

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

A **local update function** is associated to each component

$$\mathscr{B} = (f_i : \mathbb{B}^{k_i \leq n} \to \mathbb{B}\ \forall i \in V)$$

- "C activates A " $\to A_{t+1} = f_A(C_t)$
- "C inhibates E"$\to E_{t+1} = f_C(C_t)$
- "both A and B can activate D" $\to D_{t+1} = f_D(A_t, B_t)$

BN of n component — formaly

in the Boolean world:
$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

A **local update function** is associated to each component

$$\mathscr{B} = (f_i : \mathbb{B}^{k_i \leq n} \to \mathbb{B}\ \forall i \in V)$$

- "C activates A" $\to A_{t+1} = f_A(C_t)$
- "C inhibates E"$\to E_{t+1} = f_E(C_t)$
- "both A and B can activate D" $\to D_{t+1} = f_D(A_t, B_t)$

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

A **local update function** is associated to each component

$$\mathscr{B} = (f_i : \mathbb{B}^{k_i \leq n} \to \mathbb{B} \; \forall i \in V)$$

- "C activates A " $\to A_{t+1} = f_A(C_t)$
- "C inhibates E"$\to E_{t+1} = f_E(C_t)$
- "both A and B can activate D" $\to D_{t+1} = f_D(A_t, B_t)$

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

A **local update function** is associated to each component

$$\mathscr{B} = (f_i : \mathbb{B}^{k_i \leq n} \to \mathbb{B}\ \forall i \in V)$$

- "C activates A " $\to A_{t+1} = f_A(C_t)$
- "C inhibates E"$\to E_{t+1} = f_E(C_t)$
- "both A and B can activate D" $\to D_{t+1} = f_D(A_t, B_t)$

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

A **local update function** is associated to each component

$$\mathscr{B} = (f_i : \mathbb{B}^{k_i \leq n} \to \mathbb{B}\ \forall i \in V)$$

- "C activates A " $\to A_{t+1} = f_D(C_t)$
- "C inhibates E"$\to E_{t+1} = f_E(C_t)$
- "both A and B can activate D" $\to D_{t+1} = f_D(A_t, B_t)$

BN of n component — formaly

in the Boolean world:

$$\mathbb{B} = \{0, 1\}$$

A set V of n **components**

The **status** of the components are all in $\mathbb{B}$.

A system of n components can thus have 2^n **configurations**.

A **local update function** is associated to each component

$$\mathscr{B} = (f_i : \mathbb{B}^{k_i \leq n} \to \mathbb{B}\ \forall i \in V)$$

- "C activates A " $\to A_{t+1} = f_D(C_t)$
- "C inhibates E"$\to E_{t+1} = f_E(C_t)$
- "both A and B can activate D" $\to D_{t+1} = f_D(A_t, B_t)$

Expression of the local update functions

Expression of the local update functions

- negation $\neg a$: "not a" $\rightarrow$ a has to be False
- disjunction $a \vee b$: "a or b"
- conjunction $a \wedge b$: "a and b"

Expression of the local update functions

- negation $\neg a$: "not a" $\rightarrow$ a has to be False
- disjunction $a \vee b$: "a or b"
- conjunction $a \wedge b$: "a and b"

$\{\neg, \vee, \wedge\}$ is functionally complete

Expression of the local update functions

- negation $\neg a$: "not a" $\rightarrow$ a has to be False
- disjunction $a \vee b$: "a or b"
- conjunction $a \wedge b$: "a and b"

$\{\neg, \vee, \wedge\}$ is functionally complete

Disjunctive Normal Form (DNF)
$=$ disjunction of conjunctions.

Expression of the local update functions

- negation $\neg a$: "not a" $\rightarrow$ a has to be False
- disjunction $a \vee b$: "a or b"
- conjunction $a \wedge b$: "a and b"

$\{\neg, \vee, \wedge\}$ is functionally complete

Disjunctive Normal Form (DNF)
= disjunction of conjunctions. $(a \wedge b \wedge c) \vee (a \wedge \neg a \wedge b) \vee a$

Expression of the local update functions

- negation $\neg a$: "not a" $\rightarrow$ a has to be False
- disjunction $a \vee b$: "a or b"
- conjunction $a \wedge b$: "a and b"

$\{\neg, \vee, \wedge\}$ is functionally complete

Disjunctive Normal Form (DNF)
$=$ disjunction of (satisfiable) conjunctions.

satisfiable conjunction $\rightarrow$ does not contain a variable and its negation.

Expression of the local update functions

- negation $\neg a$: "not a" $\rightarrow$ a has to be False
- disjunction $a \vee b$: "a or b"
- conjunction $a \wedge b$: "a and b"

$\{\neg, \vee, \wedge\}$ is functionally complete

Disjunctive Normal Form (DNF)
$=$ disjunction of (satisfiable) conjunctions.

satisfiable conjunction $\rightarrow$ does not contain a variable and its negation.

$$a\text{: OK}$$
$$a \wedge \neg a\text{: not OK}$$
$$(a \wedge \neg a) \vee b\text{: satisfiable } (\equiv b\ ) \text{ but still not OK.}$$

Examples of local update functions

- "C activates A" $\rightarrow A_{t+1} = f_A(C_t)$
- "C inhibates E" $\rightarrow E_{t+1} = f_E(C_t)$
- "both A and B can activate D"
  $\rightarrow D_{t+1} = f_D(A_t, B_t)$

Examples of local update functions

- "C activates A" $\rightarrow A_{t+1} = f_A(C_t) = C_t$
- "C inhibates E" $\rightarrow E_{t+1} = f_E(C_t)$
- "both A and B can activate D"
  $\rightarrow D_{t+1} = f_D(A_t, B_t)$

Examples of local update functions

- "C activates A" $\rightarrow A_{t+1} = f_A(C_t) = C_t$
- "C inhibates E" $\rightarrow E_{t+1} = f_E(C_t) = \neg C_t$
- "both A and B can activate D"
  $\rightarrow D_{t+1} = f_D(A_t, B_t)$

Examples of local update functions

- "C activates A" $\rightarrow A_{t+1} = f_A(C_t) = C_t$
- "C inhibates E" $\rightarrow E_{t+1} = f_E(C_t) = \neg C_t$
- "both A and B can activate D"
  $\rightarrow D_{t+1} = f_D(A_t, B_t) = A_t \vee B_t$

Examples of local update functions

- "C activates A" $\rightarrow A_{t+1} = f_A(C_t) = C_t$
- "C inhibates E" $\rightarrow E_{t+1} = f_E(C_t) = \neg C_t$
- "both A and B can activate D"
  $\rightarrow D_{t+1} = f_D(A_t, B_t) = A_t \vee B_t$

Examples of local update functions

- "C activates A" $\rightarrow A = f_A(C) = C$
- "C inhibates E" $\rightarrow E = f_E(C) = \neg C$
- "both A and B can activate D"
  $\rightarrow D = f_D(A, B) = A \vee B$

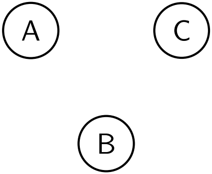
Note: from now on, I will omit the time specifications

Examples of BNs

$$\mathscr{B}_1 = \begin{cases} f_A : A = C \\ f_B : B = B \wedge \neg C \\ f_C : C = \neg C \end{cases}$$

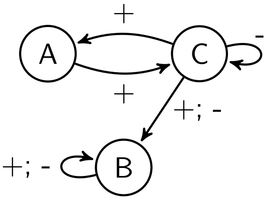
$$\mathscr{B}_2 = \begin{cases} f_A : A = 0 \\ f_B : B = (B \wedge \neg C) \vee (\neg B \wedge C) \\ f_C : C = A \end{cases}$$

Automatic synthesis of BN



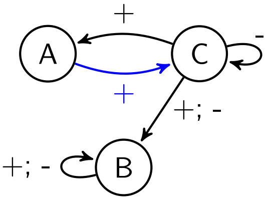
Automatic synthesis of BN

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



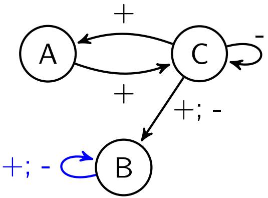
Automatic synthesis of BN

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



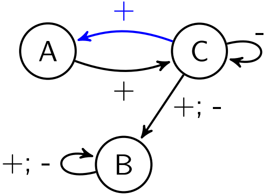
Automatic synthesis of BN

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



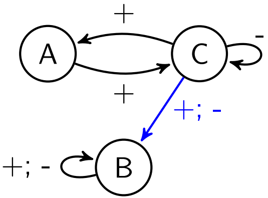
Automatic synthesis of BN

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



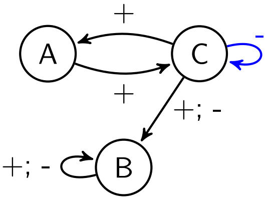
Automatic synthesis of BN

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



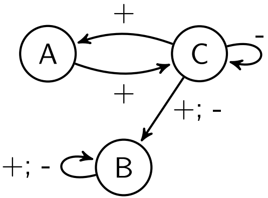
Automatic synthesis of BN

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



Automatic synthesis of BN

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"

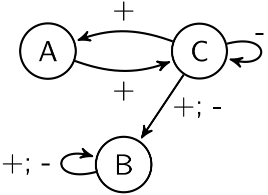


Automatic synthesis of BN

Prior Knowledge Network (PKN)

Super-set of influences allowed in the synthesized BNs

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"

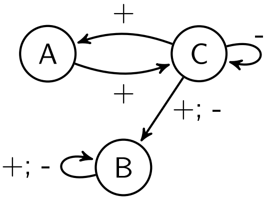


Automatic synthesis of BN

Prior Knowledge Network (PKN)

Super-set of influences allowed in the synthesized BNs

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



Multivariate Time Series (TS)

Concentrations / activities of the components over time

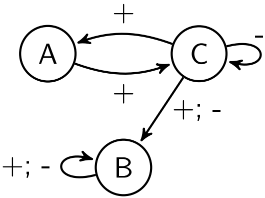
t	1	2	3	4	5	6	7	8	9	10	13	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Automatic synthesis of BN

Prior Knowledge Network (PKN)

Super-set of influences allowed in the synthesized BNs

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



Multivariate Time Series (TS)

Concentrations / activities of the components over time

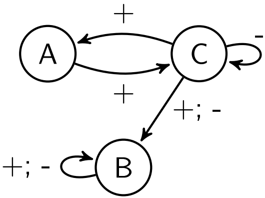
t	1	2	3	4	5	6	7	8	9	10	13	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Automatic synthesis of BN

Prior Knowledge Network (PKN)

Super-set of influences allowed in the synthesized BNs

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



Multivariate Time Series (TS)

Concentrations / activities of the components over time

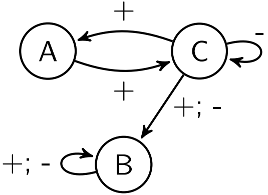
t	1	2	3	4	5	6	7	8	9	10	13	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Automatic synthesis of BN

Prior Knowledge Network (PKN)

Super-set of influences allowed in the synthesized BNs

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



Multivariate Time Series (TS)

Concentrations / activities of the components over time

t	1	2	3	4	5	6	7	8	9	10	13	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Automatic approaches

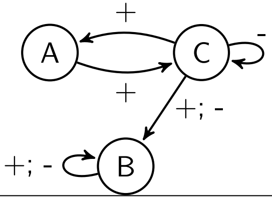
Automatic enumeration of BNs respecting a PKN and a TS.

- `REVEAL`
- `Best-Fit`
- `caspots`
- `our`

Hard constraint: use PKN to delimit the search space

Super-set of influences allowed in the synthesized BNs

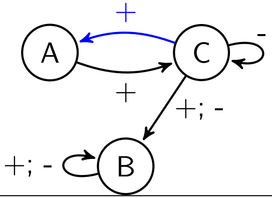
"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



Hard constraint: use PKN to delimit the search space

Super-set of influences allowed in the synthesized BNs

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



For A:

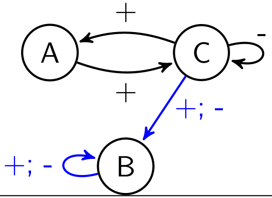
3 choices:
$A = C$;
$A = 0$;
$A = 1$;

but not:
$A = B$;
$A = \neg C$

Hard constraint: use PKN to delimit the search space

Super-set of influences allowed in the synthesized BNs

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



For A:

3 choices:
$A = C$;
$A = 0$;
$A = 1$;

but not:
$A = B$;
$A = \neg C$

For B:

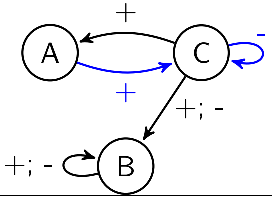
16 choices:
$B = B \wedge \neg C$;
$B = (B \wedge \neg C) \vee (\neg B \wedge C)$;
...
$B = 0$;
$B = 1$

but not:
$B = A$

Hard constraint: use PKN to delimit the search space

Super-set of influences allowed in the synthesized BNs

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



For A:

3 choices:
$A = C$;
$A = 0$;
$A = 1$;

but not:
$A = B$;
$A = \neg C$

For B:

16 choices:
$B = B \wedge \neg C$;
$B = (B \wedge \neg C) \vee (\neg B \wedge C)$;
...
$B = 0$;
$B = 1$

but not:
$B = A$

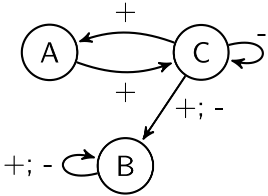
For C:

6 choices:
$C = \neg C$;
$C = A$;
...
$C = 0$;
$C = 1$

but not:
$C = A \wedge B$

How did the PKN help reducing the search space?

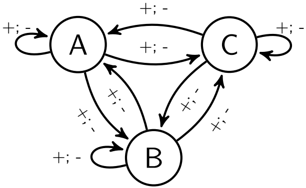
With the PKN



A	B	C
3	16	6

$\rightarrow 3 \times 16 \times 6 = 288$
potential BNs

Without informations



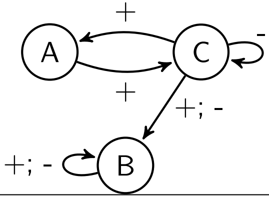
$2^{2^3} = 256$ candidates for each components

$\rightarrow 256 \times 256 \times 256 = 16777216$
potential BNs

Hard constraint: use PKN to delimit the search space

Super-set of influences allowed in the synthesized BNs

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



For A:

3 choices:
$A = C$;
$A = 0$;
$A = 1$;

but not:
$A = B$;
$A = \neg C$

For B:

16 choices:
$B = B \wedge \neg C$;
$B = (B \wedge \neg C) \vee (\neg B \wedge C)$;
...
$B = 0$;
$B = 1$

but not:
$B = A$

For C:

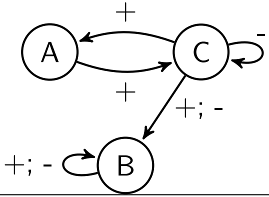
6 choices:
$C = \neg C$;
$C = A$;
...
$C = 0$;
$C = 1$

but not:
$C = A \wedge B$

Hard constraint: use PKN to delimit the search space

Super-set of influences allowed in the synthesized BNs

"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



For A:

3 choices:
$\boxed{A = C;}$
$\boxed{A = 0;}$
$A = 1;$

but not:
$A = B;$
$A = \neg C$

For B:

16 choices:
$\boxed{B = B \wedge \neg C;}$
$\boxed{B = (B \wedge \neg C) \vee (\neg B \wedge C);}$
$\ldots$
$B = 0;$
$B = 1$

but not:
$B = A$

For C:

6 choices:
$\boxed{C = \neg C;}$
$\boxed{C = A;}$
$\ldots$
$C = 0;$
$C = 1$

but not:
$C = A \wedge B$

Soft constraints: adequation with the TS

Soft constraints: adequation with the TS

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Soft constraints: adequation with the TS

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Binarization of the TS:

$$o'_t = \begin{cases} 1 \text{ if } o_t > \theta_o \\ 0 \text{ otherwise} \end{cases}$$

$$\theta_o = \frac{o_{max} - o_{min}}{2}$$

Soft constraints: adequation with the TS

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Binarization of the TS:

$$o'_t = \begin{cases} 1 \text{ if } o_t > \theta_o \\ 0 \text{ otherwise} \end{cases}$$

$$\theta_o = \frac{o_{max} - o_{min}}{2}$$

Here :

$\theta_A = \frac{100-0}{2} = 50$ $\theta_B = \frac{100-0}{2} = 50$ $\theta_C = \frac{100-0}{2} = 50$

Soft constraints: adequation with the TS

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Binarization of the TS:

$$o'_t = \begin{cases} 1 \text{ if } o_t > \theta_o \\ 0 \text{ otherwise} \end{cases}$$

$$\theta_o = \frac{o_{max} - o_{min}}{2}$$

Here :

$\theta_A = \frac{100-0}{2} = 50$ $\theta_B = \frac{100-0}{2} = 50$ $\theta_C = \frac{100-0}{2} = 50$

Soft constraints: adequation with the TS

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
		010																		
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Soft constraints: adequation with the TS

			010		$\rightarrow$		011													
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Soft constraints: adequation with the TS

			010		$\rightarrow$	011		$\rightarrow$	100											
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Soft constraints: adequation with the TS

	010			$\rightarrow$	011		$\rightarrow$	100		$\rightarrow$	001									
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Which function is best? (for A)

	010			$\rightarrow$	011		$\rightarrow$	100	$\rightarrow$	001									
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Which function is best? (for A)

			010	$\rightarrow$		011		$\rightarrow$	100		$\rightarrow$ 001									
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$

Which function is best? (for A)

		010		$\rightarrow$	011		$\rightarrow$	100		$\rightarrow$	001									
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$

Mean Absolute Error

Which function is best? (for A)

			010		→	011		→	100		→	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$

Mean Absolute Error $\quad \mathrm{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

Which function is best? (for A)

		010		$\rightarrow$		011		$\rightarrow$		100		$\rightarrow$		001						
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

Which function is best? (for A)

			010		$\rightarrow$	011		$\rightarrow$	100		$\rightarrow$	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$

Mean Absolute Error $\quad \mathrm{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

Which function is best? (for A)

		010		$\rightarrow$		011		$\rightarrow$	100		$\rightarrow$	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

Which function is best? (for A)

	010			→	011			→	100			→	001							
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$

Mean Absolute Error $\quad \mathrm{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for A)

		010		$\rightarrow$		011		$\rightarrow$		100		$\rightarrow$		001						
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$

Mean Absolute Error $\quad \mathrm{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for A)

		010		$\rightarrow$	011		$\rightarrow$	100		$\rightarrow$	001									
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$

Mean Absolute Error $\quad \mathrm{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained
θ_X: binarization threashold of X
$X_{t'}$: observed value of X at time t'
T: number of time steps in the TS

Which function is best? (for A)

			010		$\rightarrow$	011		$\rightarrow$	100		$\rightarrow$	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for A)

		010		$\rightarrow$	011		$\rightarrow$	100		$\rightarrow$	001									
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained
θ_X: binarization threashold of X
$X_{t'}$: observed value of X at time t'
T: number of time steps in the TS

Which function is best? (for A)

			010		→	011	→		100	→		001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$
$\mathscr{U}$		

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for A)

			010		→	011		→	100		→	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$
$\mathscr{U}$		

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for A)

		010	→	011		→	100	→	001										
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$
$\mathscr{U}$		

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained
θ_X: binarization threashold of X
$X_{t'}$: observed value of X at time t'
T: number of time steps in the TS

Which function is best? (for A)

	010				→	011		→	100		→	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$
$\mathscr{U}$		

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained
θ_X: binarization threashold of X
$X_{t'}$: observed value of X at time t'
T: number of time steps in the TS

Which function is best? (for A)

		010		→	011		→	100		→	001									
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$
$\mathscr{U}$		

Mean Absolute Error $\quad \mathrm{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for A)

		010		$\rightarrow$	011		$\rightarrow$	100		$\rightarrow$	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$
$\mathscr{U}$		

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained
θ_X: binarization threashold of X
$X_{t'}$: observed value of X at time t'
T: number of time steps in the TS

Which function is best? (for A)

	010				→	011			→	100		→	001							
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$
$\mathscr{U}$		

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for A)

		010		$\rightarrow$		011	$\rightarrow$		100	$\rightarrow$		001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$
$\mathscr{U}$	$\emptyset$	

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained
θ_X: binarization threashold of X
$X_{t'}$: observed value of X at time t'
T: number of time steps in the TS

Which function is best? (for A)

		010		$\rightarrow$	011		$\rightarrow$	100		$\rightarrow$	001									
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$
$\mathscr{U}$	$\emptyset$	

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained
θ_X: binarization threashold of X
$X_{t'}$: observed value of X at time t'
T: number of time steps in the TS

Which function is best? (for A)

		010		$\rightarrow$		011		$\rightarrow$		100		$\rightarrow$	001							
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$
$\mathscr{U}$	$\emptyset$	$\{8\}$

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained
θ_X: binarization threashold of X
$X_{t'}$: observed value of X at time t'
T: number of time steps in the TS

Which function is best? (for A)

	010				→	011		→	100		→	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$	$A = 0$
$\mathscr{U}$	$\emptyset$	$\{8\}$
MAE	0	0.55

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for A)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
		010		$\to$		011		$\to$	100		$\to$	001								
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$		$A = 0$
$\mathscr{U}$	$\emptyset$		$\{8\}$
MAE	0	✓	0.55

Mean Absolute Error $\quad \mathrm{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for A)

	010				→	011		→	100		→	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for A	$A = C$ ✓	$A = 0$
$\mathscr{U}$	$\emptyset$	$\{8\}$
MAE	0 ✓	0.55

Mean Absolute Error $\quad \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained
θ_X: binarization threashold of X
$X_{t'}$: observed value of X at time t'
T: number of time steps in the TS

Which function is best? (for B)

			010	$\rightarrow$		011		$\rightarrow$	100		$\rightarrow$	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

$$\text{Mean Absolute Error MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for B)

	010			$\rightarrow$	011		$\rightarrow$	100		$\rightarrow$	001									
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for B	$B = B \wedge \neg C$	$B = (B \wedge \neg C) \vee (\neg B \wedge C)$

$$\text{Mean Absolute Error } \mathrm{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained
θ_X: binarization threashold of X
$X_{t'}$: observed value of X at time t'
T: number of time steps in the TS

Which function is best? (for B)

			010		$\rightarrow$	011		$\rightarrow$	100		$\rightarrow$	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for B	$B = B \wedge \neg C$	$B = (B \wedge \neg C) \vee (\neg B \wedge C)$
$\mathscr{U}$	$\emptyset$	$\emptyset$
MAE	0 ✓	0 ✓

Mean Absolute Error $\text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for B)

	010			$\rightarrow$	011		$\rightarrow$	100		$\rightarrow$	001									
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for B	$B = B \wedge \neg C$		$B = (B \wedge \neg C) \vee (\neg B \wedge C)$	
$\mathscr{U}$	$\emptyset$		$\emptyset$	
MAE	0	✓	0	✓
# of influences	2	✓	4	

$$\text{Mean Absolute Error } \mathrm{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for B)

				010		→		011		→		100		→		001				
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for B	$B = B \wedge \neg C$ ✓		$B = (B \wedge \neg C) \vee (\neg B \wedge C)$	
$\mathscr{U}$	$\emptyset$		$\emptyset$	
MAE	0	✓	0	✓
# of influences	2	✓	4	

$$\text{Mean Absolute Error } \mathrm{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for C)

			010		$\to$	011		$\to$	100		$\to$	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Mean Absolute Error $\text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for C)

			010		→	011		→	100		→	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for C	$C = \neg C$	$C = A$

$$\text{Mean Absolute Error } \text{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for C)

010 $\rightarrow$ 011 $\rightarrow$ 100 $\rightarrow$ 001

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for C	$C = \neg C$		$C = A$
$\mathscr{U}$	$\emptyset$		$\{5\}$
MAE	0	✓	0.5

$$\text{Mean Absolute Error } \mathrm{MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained

θ_X: binarization threashold of X

$X_{t'}$: observed value of X at time t'

T: number of time steps in the TS

Which function is best? (for C)

		010		→		011		→	100		→	001								
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

candidates for C	$C = \neg C$ ✓	$C = A$
$\mathscr{U}$	$\emptyset$	$\{5\}$
MAE	0 ✓	0.5

$$\text{Mean Absolute Error MAE}_X = \frac{\sum_{t' \in \mathscr{U}} |\theta_X - X_{t'}|}{T}$$

$\mathscr{U}$: the set of t' such that the transition at t–t' is unexplained
θ_X: binarization threashold of X
$X_{t'}$: observed value of X at time t'
T: number of time steps in the TS

Generated BN

Only one best candidate of each component $\rightarrow$1 BN generated:

$$\mathscr{B}_1 = \begin{cases} f_A : A = C \\ f_B : B = B \wedge \neg C \\ f_C : C = \neg C \end{cases}$$

Summary of our approach

For each component:

- generates the candidate functions
- removes the ones that do not respect the PKN
- acts like an exhaustive evaluation of all the candidates
- returns the ones with the smallest MAE and smallest number of inlfuences

Summary of our approach

For each component:

- generates the candidate functions
- removes the ones that do not respect the PKN
- **acts like** an exhaustive evaluation of all the candidates
- returns the ones with the smallest MAE and smallest number of inlfuences

The ASP solver is actually able to cut into the search space, and does not need to really evaluate all the candidates to find the best ones.

Other automatic approaches

Other automatic approaches

- `REVEAL`
- `Best-Fit`

Other automatic approaches

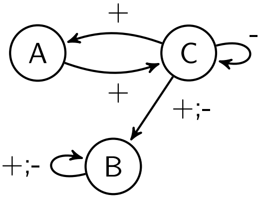
- `REVEAL`
- `Best-Fit`

unsigned PKN

Other automatic approaches

- `REVEAL`
  unsigned PKN
- `Best-Fit`

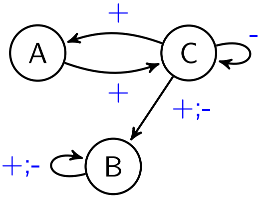
"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



Other automatic approaches

- `REVEAL`
  unsigned PKN
- `Best-Fit`

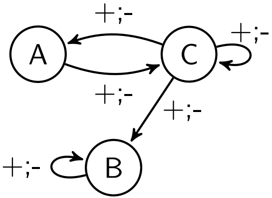
"A activates C"
"B interacts with itself"
"C activates A"
"C interacts with B"
"C inhibits itself"



Other automatic approaches

- `REVEAL`
  unsigned PKN
- `Best-Fit`

"A interacts with C"
"B interacts with itself"
"C interacts with A"
"C interacts with B"
"C interacts with itself"



Other automatic approaches

- `REVEAL`
- `Best-Fit`

unsigned PKN $+$ binarized TS.

Other automatic approaches

- `REVEAL`
- `Best-Fit`

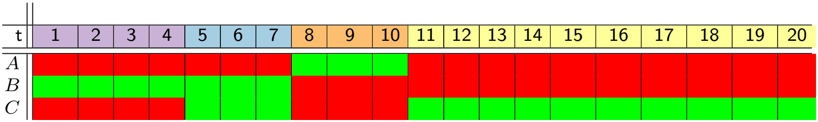
unsigned PKN + binarized TS.

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

Other automatic approaches

- `REVEAL`
- `Best-Fit`

unsigned PKN + binarized TS.



Other automatic approaches

- `REVEAL`
- `Best-Fit`

unsigned PKN + binarized TS.

Tries to explain all / the max # of **time steps**.

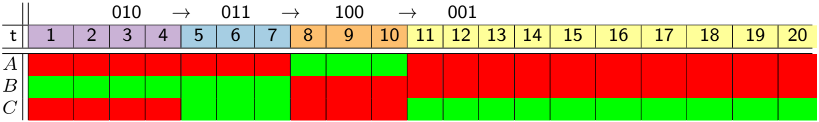
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A																				
B																				
C																				

Other automatic approaches

- `REVEAL`
- `Best-Fit`

unsigned PKN + binarized TS.

Tries to explain all / the max # of **time steps**.

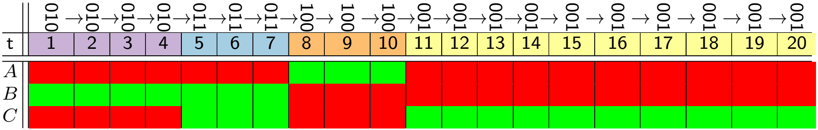


Other automatic approaches

- `REVEAL`
- `Best-Fit`

unsigned PKN + binarized TS.

Tries to explain all / the max # of **time steps**.



Other automatic approaches

- `REVEAL`
- `Best-Fit`

unsigned PKN + binarized TS.
Tries to explain all / the max # of **time steps**.

- `caspots`

PKN + TS.
Assert **reachability** of the configurations and fit the BN to the TS using MAE (slightly differently than our approach).

Other automatic approaches

- `REVEAL`
- `Best-Fit`

unsigned PKN + binarized TS.

Tries to explain all / the max # of **time steps**.

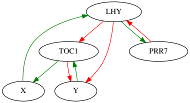
- `caspots`

PKN + TS.

Assert **reachability** of the configurations and fit the BN to the TS using MAE (slightly differently than our approach).

	010	→	011	→	100	→	001													
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	3	7	13	20	30	49	61	100	63	36	25	2	3	1	1	3	0	0	0
B	100	86	64	57	54	53	51	49	45	37	33	28	22	19	14	12	9	5	2	0
C	0	27	36	42	60	75	54	44	38	48	60	72	88	90	100	100	100	100	100	100

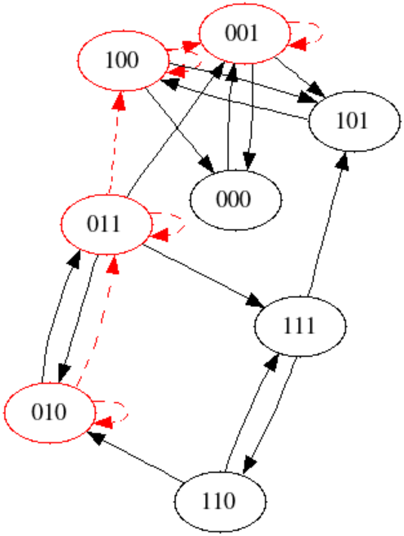
Evaluation on real datasets

System	Genes	PKN	TS	Source
yeast (cell cycle)	Fkh2, Swi5, Sic1 & Clb1	Sic1 does not influence itself nor Fkh2	14 time steps 6 transitions	Spellman 1998
A. thaliana (circadian clock)	LHY, PRR7, TOC1, X & Y		50 time steps 11 transitions	Locke 2006

Evaluation on real datasets

- Give the PKN and TS to REVEAL, Best-Fit, caspots and our method
- For each BN returned, compute its generalized asynchronous STG
- Count the number of configurations transition (extracted from the TS) retrieved.

asynchronous subset $(\{A\},\{B\},\{C\})$ of a generalized asynchronous STG:



Evaluation on real datasets

yeast

6 transitions to retrieve

\# BNs:

`Best-Fit`: 16

`caspots`: 61

`our`: 16

A. thaliana

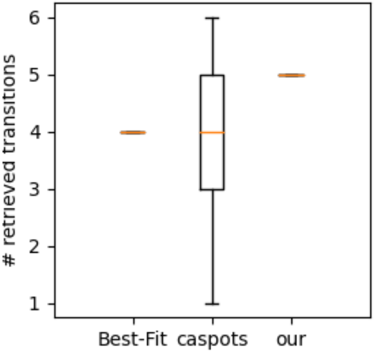
11 transitions to retrieve

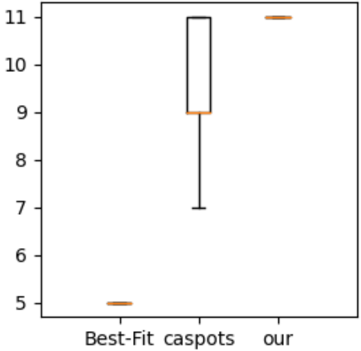
\# BNs:

`Best-Fit`: 1

`caspots`: 5

`our`: 8





Note: `REVEAL` failed because of inconsistencies

Conclusion

The approach we propose is quite simple. Yet, it gaves us interesting results so far.

We are now applying it on more datasets, and then we will use it on PKN and TS directly extracted from SBML models.

The end. Any questions ?

Annexe

About minimal DNF

Minimal DNF: smallest DNF among all the equivalent DNFs.

About minimal DNF

Minimal DNF: **smallest** DNF among all the equivalent DNFs.

Length of a formula: the number of litteral
Example:
$(a \wedge \neg b) \vee (a \wedge c)$: length = 4
$a \wedge a \wedge a \wedge a \wedge a$: length = 5

About minimal DNF

Minimal DNF: **smallest** DNF among all the equivalent DNFs.

Length of a formula: the number of litteral
Example:
$(a \wedge \neg b) \vee (a \wedge c)$: length $= 4$
$a \wedge a \wedge a \wedge a \wedge a$: length $= 5$

Example of min DNF:
a is a minimal DNF, but $a \wedge a$ is not.

Logical operators and their notation

- $\neg a$: "not a" $\rightarrow$ a has to be False
- $a \vee b$: "a or b"
- $a \wedge b$: "a and b"

- $a \Rightarrow b$: "a implies b"
- $a \Leftrightarrow b$: "equivalent"
- True
- False
- $a \oplus b$: "exclusive or", "xor"
- $a \uparrow b$ "nand"
- $a \downarrow b$: "nor"
- ...

Logical operators and their notation

- $\neg a$: "not a" $\rightarrow$ a has to be False
- $a \vee b$: "a or b"
- $a \wedge b$: "a and b"

- $a \Rightarrow b \equiv \neg a \vee b$: "$a$ implies b"
- $a \Leftrightarrow b \equiv (\neg a \vee b) \wedge (\neg b \vee a)$: "equivalent"
- True $\equiv a \vee \neg a$
- False $\equiv a \wedge \neg a$
- $a \oplus b \equiv (a \wedge \neg b) \vee (\neg a \wedge b)$: "exclusive or", "xor"
- $a \uparrow b \equiv \neg(a \wedge b)$ "nand"
- $a \downarrow b \equiv \neg(x \vee b)$: "nor"
- ...

Logical operators and their notation

- $\neg a$: "not a" $\rightarrow$ a has to be False
- $a \vee b$: "a or b"
- $a \wedge b$: "a and b"

abbreviations $\rightarrow$ not absolutely necessary:

- $a \Rightarrow b \equiv \neg a \vee b$: "$a$ implies b"
- $a \Leftrightarrow b \equiv (\neg a \vee b) \wedge (\neg b \vee a)$: "equivalent"
- True $\equiv a \vee \neg a$
- False $\equiv a \wedge \neg a$
- $a \oplus b \equiv (a \wedge \neg b) \vee (\neg a \wedge b)$: "exclusive or", "xor"
- $a \uparrow b \equiv \neg(a \wedge b)$ "nand"
- $a \downarrow b \equiv \neg(x \vee b)$: "nor"
- $\ldots$

Logical operators and their notation

functionally complete:

- $\neg a$: "not a" $\rightarrow$ a has to be False
- $a \vee b$: "a or b"
- $a \wedge b$: "a and b"

abbreviations $\rightarrow$ not absolutely necessary:

- $a \Rightarrow b \equiv \neg a \vee b$: "$a$ implies b"
- $a \Leftrightarrow b \equiv (\neg a \vee b) \wedge (\neg b \vee a)$: "equivalent"
- True $\equiv a \vee \neg a$
- False $\equiv a \wedge \neg a$
- $a \oplus b \equiv (a \wedge \neg b) \vee (\neg a \wedge b)$: "exclusive or", "xor"
- $a \uparrow b \equiv \neg(a \wedge b)$ "nand"
- $a \downarrow b \equiv \neg(x \vee b)$: "nor"
- ...

Logical operators and their notation

functionally complete:

- $\neg a$: "not a" $\rightarrow$ a has to be False
- $a \vee b$: "a or b"
- $a \wedge b$: "a and b"

Automatic approaches

Automatic approaches

- `REVEAL`

unsigned PKN + binarized TS.
Tries to explain all the time steps.

Automatic approaches

- `REVEAL`

unsigned PKN + binarized TS.
Tries to explain all the time steps.

- `Best-Fit`

unsigned PKN + binarized TS.
Tries to explain the max # of time steps.

Automatic approaches

- `REVEAL`

unsigned PKN + binarized TS.
Tries to explain all the time steps.

- `Best-Fit`

unsigned PKN + binarized TS.
Tries to explain the max # of time steps.

- `caspots`

PKN + TS.
Assert reachability of the configurations and fit the BN to the TS using MAE.

Automatic approaches

- `REVEAL`

unsigned PKN + binarized TS.
Tries to explain all the time steps.

- `Best-Fit`

unsigned PKN + binarized TS.
Tries to explain the max # of time steps.

- `caspots`

PKN + TS.
Assert reachability of the configurations and fit the BN to the TS using MAE.

- `our`

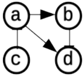
PKN + TS.
Fit the BN to the TS using MAE (but a bit differently than `caspots` ;)).

Boolean Networks (BN): vocabulary

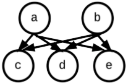
- component
- component's status
- BN's configuration
- influence
- local update functions
- updating scheme
- state transition graph
- reachability

Formalisms for modeling biological systems

Boolean network



Bayesian network

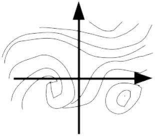


Process algebras

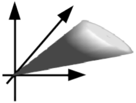
((b(x,de)[E]) || (B(y, dI)[I]))

bh(x, dE) bh(y, dI) (E || I)

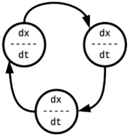
Differential equation



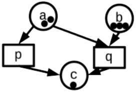
Constraint based model



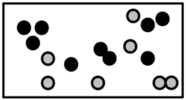
Hybrid systems



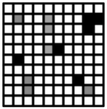
Petri Nets



Agent-based model



Cellular automata



Interacting state machine

Compartment based

Rule based

...

Choice of the formalism

the more abstract the better (do not bother with details) as long as it is sufficient to answer the question