

# A constraint programming approach to the analysis of Petri nets structural properties and application to biochemical networks

Faten Nabli

Supervised by François Fages and Sylvain Soliman

INRIA Paris-Rocquencourt  
CONTRAINTES

May 21, 2013

# Outline

- 1 Motivation from systems biology
- 2 Boolean model for siphons/traps
  - Complexity
  - Model & Strategy for minimal siphon
  - Comparison SAT/CSP/state-of-the-art algorithm
  - Linear time complexity result
- 3 Conclusion

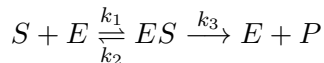
# Differential equations and structure

A classical way to describe and analyse biochemical reaction systems: [Differential equations](#).

# Differential equations and structure

A classical way to describe and analyse biochemical reaction systems: **Differential equations**.

## Reaction model



$$dS/dt = -k_1 \times S \times E + k_2 \times ES$$

$$dP/dt = k_3 \times ES$$

$$dE/dt = -k_1 \times S \times E + (k_2 + k_3) \times ES$$

$$dES/dt = k_1 \times S \times E - (k_2 + k_3) \times ES$$

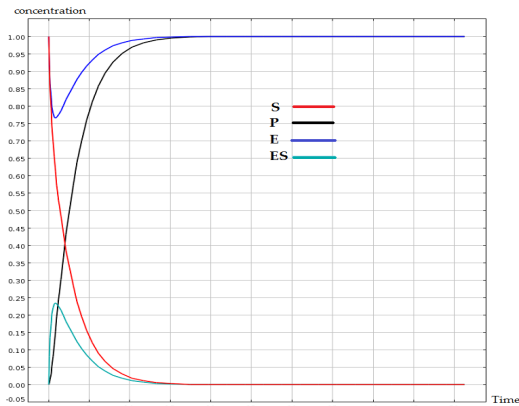
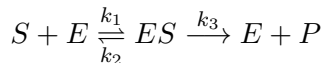
**1913** *Die Kinetik der Invertinwirkung.*

L. Menten, M.I. Michaelis. *Biochemistry Zeitung* 49.

# Differential equations and structure

A classical way to describe and analyse biochemical reaction systems: **Differential equations**.

## Reaction model



# Differential equations and structure

A classical way to describe and analyse biochemical reaction systems: [Differential equations](#).

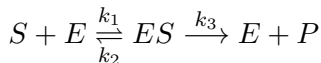
- ✎ Rate constants usually not known
- ✎ Analytically intractable even for small systems.

**Use of Petri nets [structural properties](#) to say something about the [system dynamics](#) without knowing the kinetics.**

# Differential equations and structure

A classical way to describe and analyse biochemical reaction systems: **Differential equations**.

## Reaction model



$$dS/dt = -k_1 \times S \times E + k_2 \times ES$$

$$dP/dt = k_3 \times ES$$

$$dE/dt = -k_1 \times S \times E + (k_2 + k_3) \times ES$$

$$dES/dt = k_1 \times S \times E - (k_2 + k_3) \times ES$$

Conservation laws

P-invariants:

$$E + ES = \text{cte}$$

$$P + S + ES = \text{cte}$$

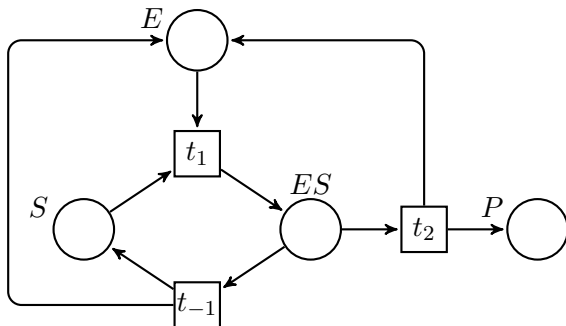
+



**Equivalent model:**  $dS/dt = k_2 \times ES - k_1 \times E \times S$   
 $dES/dt = k_1 \times E \times S - (k_2 + k_3) \times ES$

# Michaelis–Menten enzymatic reactions

## Structural model: Reaction graph



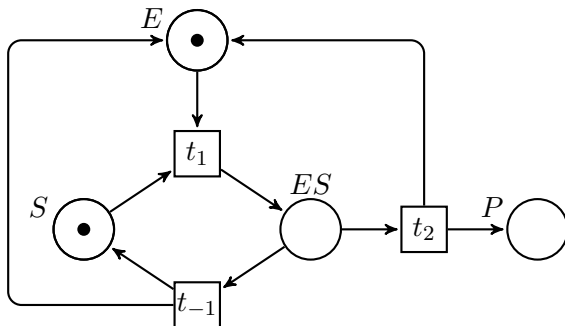
Petri-net = reaction graph + discrete dynamics

1962 *Kommunikation mit Automaten*. Carl Adam Petri.  
Ph. D. Thesis. University of Bonn.



# Michaelis–Menten enzymatic reactions

## Structural model: Reaction graph

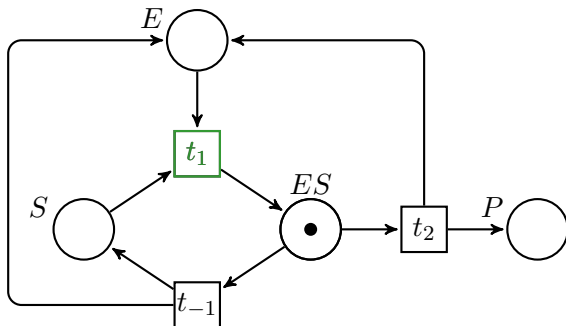


Petri-net = reaction graph + discrete dynamics

1962 *Kommunikation mit Automaten*. Carl Adam Petri.  
Ph. D. Thesis. University of Bonn.

# Michaelis–Menten enzymatic reactions

## Structural model: Reaction graph

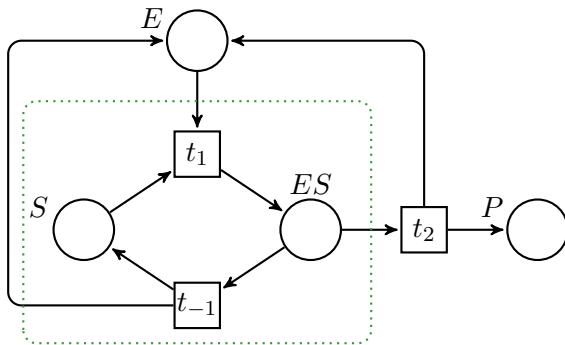


Petri-net = reaction graph + discrete dynamics

1962 *Kommunikation mit Automaten*. Carl Adam Petri.  
Ph. D. Thesis. University of Bonn.

# Siphon

A **siphon** is a non-empty set of chemical species that once all of them is absent, none of them will ever become present.

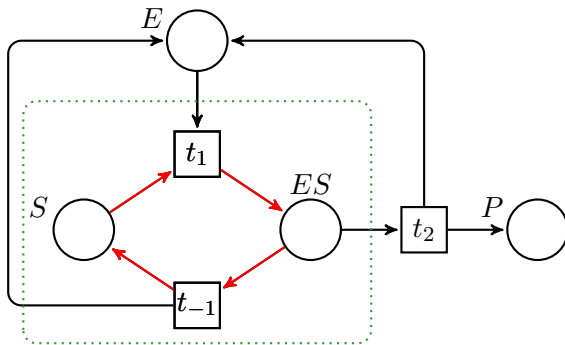


$\{S, ES\}$  is a siphon

2003 *Topological analysis of metabolic networks based on Petri net theory.* I.Zevedei-Oancea and S.Schuster. In *Silico Biology*.

# Siphon

A **siphon** is a non-empty set of chemical species that once all of them is absent, none of them will ever become present.

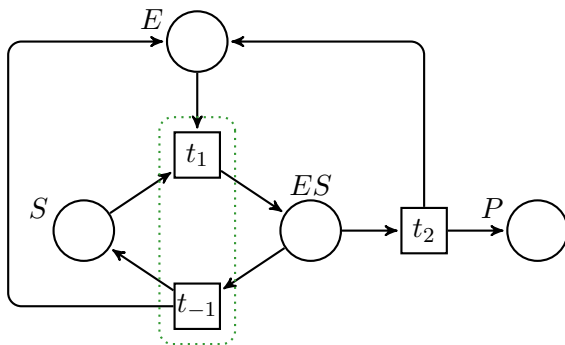


$\{S, ES\}$  is a siphon

2003 *Topological analysis of metabolic networks based on Petri net theory.* I.Zevedei-Oancea and S.Schuster. In *Silico Biology*.

# T-invariant

Minimal **T-invariants** correspond to elementary flux modes.



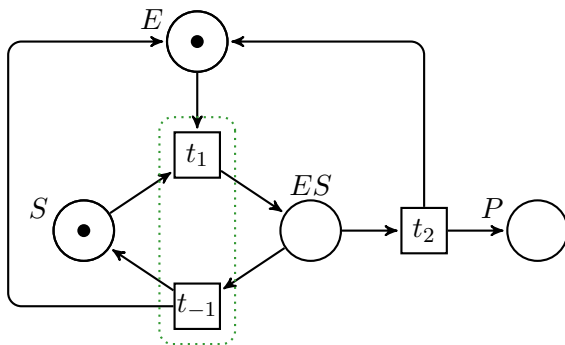
$t_1 = t_{-1} = k$  ( $k \geq 0$ ) and  $t_2 = 0$  is a T-invariant

1994 *On elementary flux modes in biochemical reaction systems at steady state.*

Schuster, S., Hilgetag, C., .J. Biol. Syst.

# T-invariant

Minimal **T-invariants** correspond to elementary flux modes.



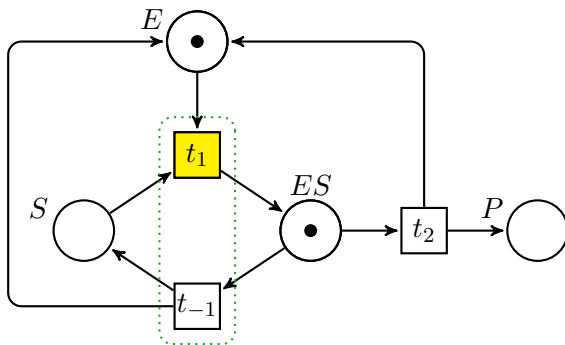
$t_1 = t_{-1} = k$  ( $k \geq 0$ ) and  $t_2 = 0$  is a T-invariant

1994 *On elementary flux modes in biochemical reaction systems at steady state.*

Schuster, S., Hilgetag, C., .J. Biol. Syst.

# T-invariant

Minimal **T-invariants** correspond to elementary flux modes.



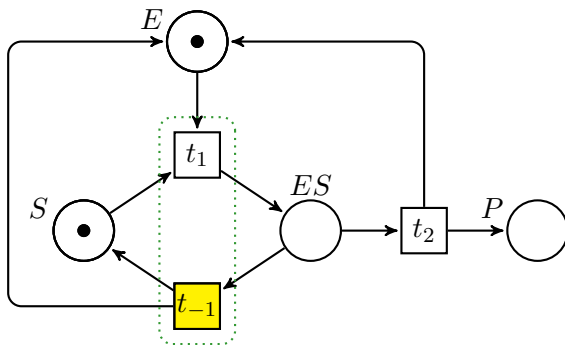
$t_1 = t_{-1} = k$  ( $k \geq 0$ ) and  $t_2 = 0$  is a T-invariant

1994 *On elementary flux modes in biochemical reaction systems at steady state.*

Schuster, S., Hilgetag, C., .J. Biol. Syst.

# T-invariant

Minimal **T-invariants** correspond to elementary flux modes.



$t_1 = t_{-1} = k$  ( $k \geq 0$ ) and  $t_2 = 0$  is a T-invariant

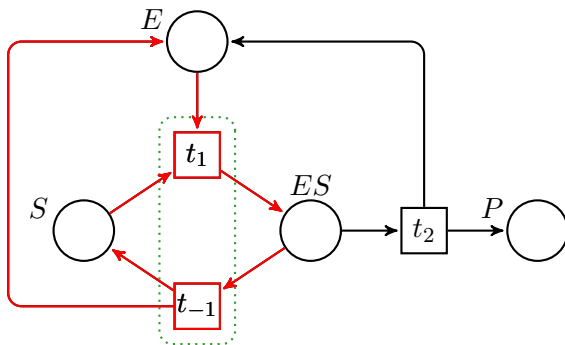
1994 *On elementary flux modes in biochemical reaction systems at steady state.*

Schuster, S., Hilgetag, C., .J. Biol. Syst.



# T-invariant

Minimal **T-invariants** correspond to elementary flux modes.



$t_1 = t_{-1} = k$  ( $k \geq 0$ ) and  $t_2 = 0$  is a T-invariant

1994 *On elementary flux modes in biochemical reaction systems at steady state.*

Schuster, S., Hilgetag, C., .J. Biol. Syst.

# Benchmark for evaluation

## Database **Biomodels.net** (version march 2012)

404 manually curated quantitative biochemical models.

Average  $\sim 50$  species,  $\sim 90$  reactions.

Biggest model has 194 species, 313 reactions.

Reference publication for each model.

**2006** *BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems.*  
le Novère et al. Nucleic Acid Research.

## Benchmark for evaluation (2)

### Database **Petriweb**

Repository of 80 models modelling real industrial processes.

Average  $\sim 10$  places,  $\sim 8$  transitions.

Biggest model has 68 places, 51 transitions.

2006 *Petriweb: A Repository for Petri Nets*,  
R. Goud et al. *Petri Nets and Other Models of  
Concurrency - ICATPN*.

# Combinatorial problems

- There can be an **exponential** number of minimal siphons and minimal T-invariants.
- Finding a (minimal) siphon of a given cardinality in a general Petri net is **NP-complete**.

1996 *Finding minimal siphons in general Petri nets.*  
S. Tanimoto, M. Yamauchi, and T. Watanabe. IEICE.

- Finding a minimal siphon containing a given set of places in a general Petri net is **NP-complete**.

1999 *Time complexity analysis of the minimal siphon extraction problem of petri nets.* S. Tanimoto, M. Yamauchi, and T. Watanabe. IEICE.

**Our goal: enumerating efficiently all of them in practical benchmark!**

# Thesis contribution

- Minimal siphons and minimal T-invariants as constraint solving problems.
- Customized branch & bound implementation for minimal siphons efficient enumeration.
- Enumerating all minimal siphons on all Biomodels.net and Petriweb models.
- Outperforming state-of-the-art algorithms on Biomodels.net and Petriweb.
- Understanding why our techniques are efficient:
  - New complexity result: linear time complexity on classes of Petri nets of bounded tree-widths.

# Boolean model for siphons

# Minimal siphon

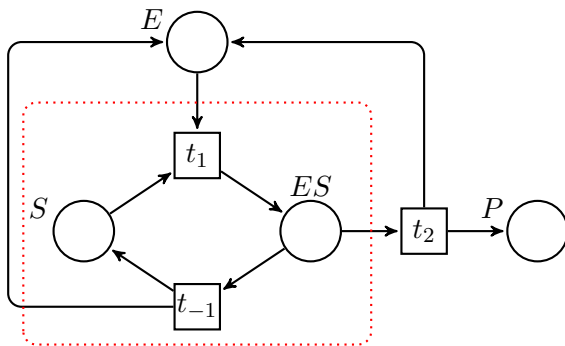
**Biological interpretation:** A siphon refers to a non-empty set of chemical species that once all of them is absent, none of them will ever become present.

**Dynamical characterisation:** A siphon is a non-empty set of places that, once it is unmarked, remains so.

**Structural characterisation:** A **siphon** is a non-empty set of places  $S$  such that  $\bullet S \subseteq S^\bullet$ .

A siphon is **minimal** if it does not contain any other siphon.

# Minimal siphon example

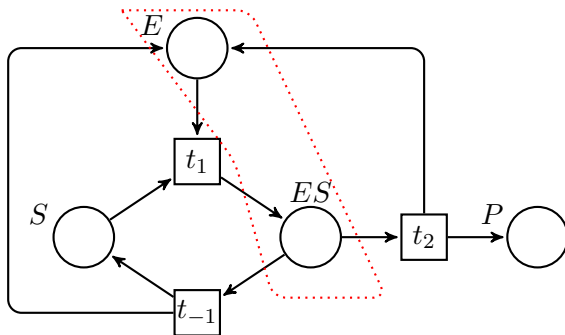


$$\bullet\{S, ES\} = \{t_1, t_{-1}\} \quad \{S, ES\}^\bullet = \{t_1, t_{-1}, t_2\}$$

$S$  siphon iff  $\bullet S \subseteq S^\bullet$



# Minimal trap example



$$\{E, ES\}^\bullet = \{t_1, t_{-1}, t_2\} \quad \bullet\{E, ES\} = \{t_1, t_{-1}, t_2\}$$

$D$  trap iff  $D^\bullet \subseteq \bullet D$

# State-of-the-art algorithms for enumerating minimal siphons

2005 *Enumeration algorithms for minimal siphons in Petri nets based on place constraints.*

R. Cordone, L. Ferrarini, and L. Piroddi.

IEEE TSC.

PN size	#minimal siphons (avg)	time (in s.) state-of-the-art algorithm
5	2	0.05
10	10	0.07
15	60	0.39
20	302	6.84

# Boolean Model of Siphons

variables

$$(\forall p) X_p = 1 \iff p \in S$$

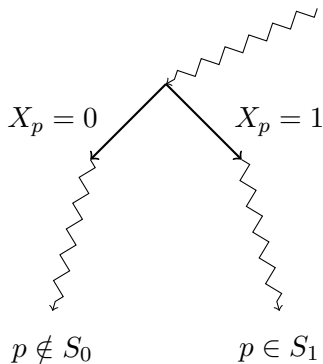
constraints

$$(\forall p) X_p = 1 \Rightarrow \bigwedge_{t \in \bullet p} \bigvee_{p' \in \bullet t} X_{p'} = 1$$

Finding siphons is reduced to finding  
**Boolean assignments satisfying these formulas.**

# Minimality in the boolean model

Value selection strategy: 0 then 1



0 before 1 and  $S_0$  before  $S_1$  in the search tree  $\implies S_1 \not\subseteq S_0$

# Enumerating all minimal siphons in the boolean model

## Branch & Bound

Once a (minimal) siphon  $S$  is found

① add the constraint  $\bigvee_{p \in S} X_p = 0$

② **restart** the search

# Enumerating minimal siphons with SAT and CSP( $\mathcal{B}$ )

database	# models	avg #P,#T	max #P+#T	avg# siphons	total time (in ms.)		
					st.art algo	mini SAT	GNU Prolog
Petriweb	80	10,8	119	2.85	2325	156	6
Biomodels.net	403	50,90	507	4.21	19734	611	195

model	#P,#T	# siphons	time (in ms.)		
			state-of-the-art algorithm	miniSAT	GNU Prolog
Biomodel #175	118, 194	3042	$\infty$	137000	$\infty$

# Enumerating minimal siphons with SAT and CSP( $\mathcal{B}$ )

database	# models	avg #P,#T	max #P+#T	avg# siphons	total time (in ms.)		
					st.art algo	mini SAT	GNU Prolog
Petriweb	80	10,8	119	2.85	2325	156	6
Biomodels.net	403	50,90	507	4.21	19734	611	195

model	#P,#T	# siphons	time (in ms.)		
			state-of-the-art algorithm	miniSAT	GNU Prolog
Biomodel #175	118, 194	3042	$\infty$	137000	$\infty$

but why are we so efficient?

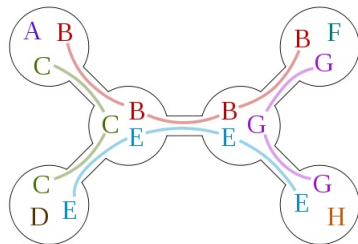
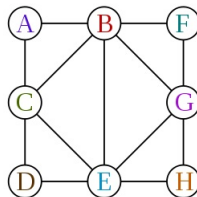
# Tree-width of non-oriented graph

## Tree-decomposition of a graph:

- 1 Each graph vertex is associated with at least one tree node.
- 2 Each graph edge connects two vertices that are listed together at some tree node.
- 3 Each graph vertex is listed at the nodes of a contiguous sub-tree of the tree.

The **width of a tree-decomposition** is the size of its largest node minus one.

The **tree-width of a graph** is the minimum **width** among all its possible **tree-decompositions**.





## Bounded tree-widths

**Theorem.** Finding minimal siphon containing a given set of places  $Q$  is **linear** for Petri-nets of **bounded tree-width**.

**Proof.** Definable in Monadic Second Order logic  $\implies$  linear-time resolution.

1990 *Definable monadic second-order logic of graphs I. Recognizable sets of Finite Graphs.* Bruno Courcelle. Information and Computation.

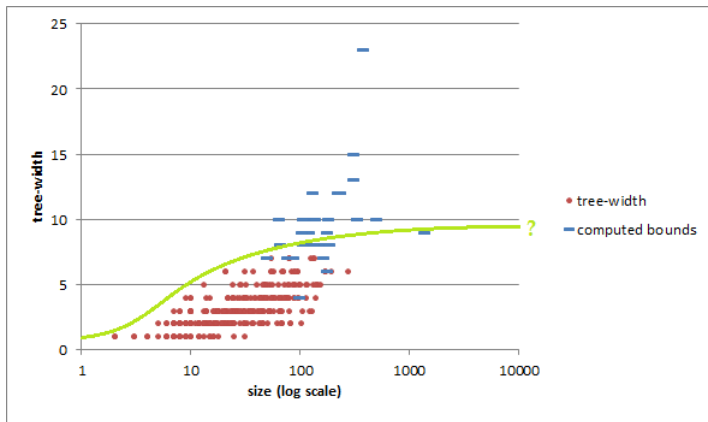
SIPHON( $S$ )

$$\begin{aligned} & \forall v(v \in S \Rightarrow \text{place}(v)) \\ & \wedge \exists v(v \in S) \\ & \wedge \forall t(\exists v(v \in S \wedge \text{edge}(t, v)) \Rightarrow \exists v(v \in S \wedge \text{edge}(v, t))) \end{aligned}$$

Minimal siphon containing  $Q$ :

$$\begin{aligned} & \exists S(\text{SIPHON}(S) \\ & \wedge \forall S'(\text{SIPHON}(S') \wedge \forall v(v \in S' \Rightarrow v \in S) \Rightarrow \forall v(v \in S \Rightarrow v \in S')) \\ & \wedge \forall v(v \in Q \Rightarrow v \in S)) \end{aligned}$$

# Biochemicals networks seem to have a **bounded** tree-width



Biomodels.net tree-width as a function of the size (places and transitions) of the Petri net  
 Computed tree-width  $\leq 10$

# Conclusion

- Petri nets structural properties can give us some information about the biochemical network dynamics when the kinetics are missing.
- Constraint programming over finite domain can be successfully applied to Petri nets structural problems.
- CSP and miniSAT solvers outperform state-of-the-art algorithms for enumerating minimal siphons.
- Surprisingly good performance on real-size models.
- Linear-time complexity result for Petri nets with bounded tree width.

# Perspectives

- Add biological constraints to the computation of T-invariants (compute only those of interest).
- Provide a generic tool for the verification of graphs structural properties having a biological meaning, with an underlying CSP engine. (Biocham)
- Identify parameters for certain structural properties computation that ensure no performances issues.
- Understand the links between structural measurements of graphs and the complexity of solving the CSP.