

Un modèle booléen pour l'énumération des siphons et des pièges minimaux dans les réseaux de Petri

Faten Nabli

François Fages

Thierry Martinez

Sylvain

Soliman

INRIA Paris-Rocquencourt
CONTRAINTEs

JFPC'12

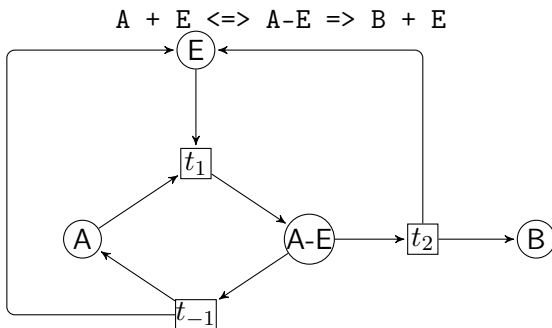
Plan

- 1 Motivation
- 2 Concepts de Base
 - Réseaux de Petri
 - Siphon Minimal
 - Complexité
- 3 Modèle Booléen
- 4 Algorithmes
 - Algorithme SAT itéré
 - Algorithme PLC(B)
- 5 Évaluation
 - Résultats et Comparaison
 - Pourquoi le calcul des siphons minimaux est-il si facile en pratique ?
- 6 Conclusion

Modèle réactionnel représenté par un réseau de Petri

espèces → places

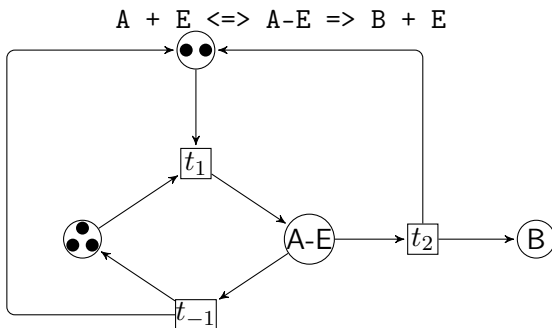
réactions → transitions



Modèle réactionnel représenté par un réseau de Petri

espèces → places

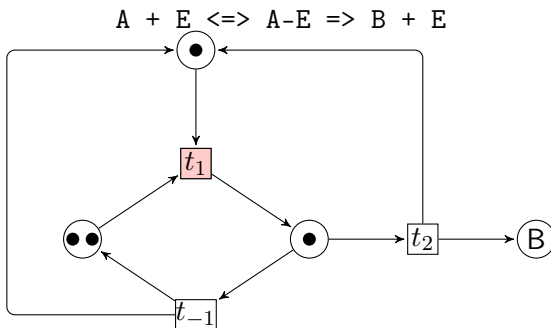
réactions → transitions



Modèle réactionnel représenté par un réseau de Petri

espèces → places

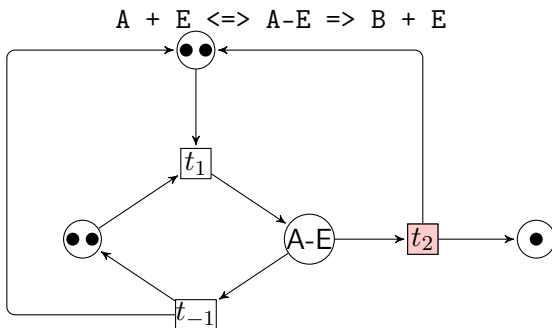
réactions → transitions



Modèle réactionnel représenté par un réseau de Petri

espèces → places

réactions → transitions



Équations différentielles et structure

- Conventionnellement les réseaux de réactions sont décrits et analysés avec des Equations différentielles.
- Manipuler des équations différentielles est inintéressant même pour les petits systèmes :

Exemple : $E + A \rightleftharpoons EA \Rightarrow E + B$

$$\frac{dA}{dt} = -k_1 E * A + k_1 EA$$

$$\frac{dEA}{dt} = k_1 E * A - (k_{-1} + k_2) EA$$

$$\frac{dE}{dt} = -k_1 E * A + (k_{-1} + k_2) ES$$

$$\frac{dB}{dt} = k_2 ES$$

Équations différentielles et structure

- Conventionnellement les réseaux de réactions sont décrits et analysés avec des Equations différentielles.
- Manipuler des équations différentielles est inintéressant même pour les petits systèmes :

Exemple : $E + A \rightleftharpoons EA \Rightarrow E + B$

$$\begin{aligned} \frac{dA}{dt} &= -k_1 E * A + k_{-1} EA & \frac{dE}{dt} &= -k_1 E * A + (k_{-1} + k_2) ES \\ \frac{dEA}{dt} &= k_1 E * A - (k_{-1} + k_2) EA & \frac{dB}{dt} &= k_2 ES \end{aligned}$$

- Ne peuvent pas être résolues analytiquement
- Constantes cinétiques k_i souvent inconnues (mal adapté pour les simulations !)

Équations différentielles et structure

- Conventionnellement les réseaux de réactions sont décrits et analysés avec des Equations différentielles.
- Manipuler des équations différentielles est inintéressant même pour les petits systèmes :

Exemple : $E + A \rightleftharpoons EA \Rightarrow E + B$

$$\begin{aligned} \frac{dA}{dt} &= -k_1 E * A + k_{-1} EA & \frac{dE}{dt} &= -k_1 E * A + (k_{-1} + k_2) ES \\ \frac{dEA}{dt} &= k_1 E * A - (k_{-1} + k_2) EA & \frac{dB}{dt} &= k_2 ES \end{aligned}$$

- Ne peuvent pas être résolues analytiquement
- Constantes cinétiques k_i souvent inconnues (mal adapté pour les simulations !)

➤ Utilisation des **propriétés structurelles** des réseaux de Pétri pour décrire la dynamique des systèmes

Banc d'essai

1 Base de donnée **Biomodels.net**

Entrepôt de 409 modèles quantitatifs biochimiques et cellulaires.

Publication de référence pour chaque modèle.

Résultats numériques corrects.

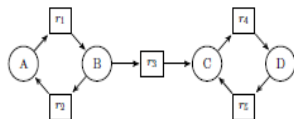
2 Base de donnée **Petriweb**

Entrepôt de 80 modèles modélisant des processus industriels réels.

Généralités sur les réseaux de Petri

Un réseau de Pétri est un 4-tuple $PN=(P, T, Pre, Post)$ où :

- $P=\{p_1, p_2, \dots, p_n\}$ est un ensemble fini de **places** \Rightarrow **espèces chimiques**
- $T=\{t_1, t_2, \dots, t_m\}$ est un ensemble fini de **transitions** \Rightarrow **réactions chimiques**
- $Pre : (P \times T) \rightarrow \mathbb{N}$ est **la matrice d'entrée**
- $Post : (P \times T) \rightarrow \mathbb{N}$ est **La matrice de sortie**



Prédécesseurs d'une place

$$\bullet p = \{t \in T \mid Post(p, t) > 0\}$$

Prédécesseurs d'un ensemble de places

$$\bullet S = \cup_{p \in S} \{ \bullet p \}$$

Successeurs d'une place

$$p^\bullet = \{t \in T \mid Pre(p, t) > 0\}$$

Successeurs d'un ensemble de places

$$S^\bullet = \cup_{p \in S} \{ p^\bullet \}$$

Définition : Siphon

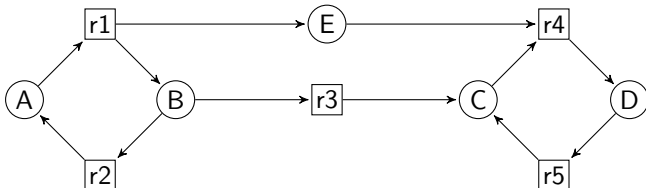
caractérisation structurelle : Ensemble de places S tel que $\bullet S \subseteq S^\bullet$.

caractérisation dynamique : Un ensemble de places tel que une fois non marqué, le reste.

interprétation biologique : Un ensemble d'espèces telles que une fois toutes absentes, aucune d'entre elles ne peut redevenir présente.

- Un siphon est minimal s'il ne contient pas d'autres siphons.

Exemple :



$\{A, B\}$ est un siphon minimal

$(\bullet\{A, B\} = \{r1, r2\} \subset \{A, B\}^\bullet = \{r1, r2, r3\})$.

Complexité

- Décider si un réseau de Petri contient un siphon ou un piège et en donner un s'il existe est polynomial. (Cordone et al, 2003)
- Le problème de décision de l'existence d'un siphon minimal contenant une place donnée est NP-difficile. (Tanimoto et al, 1996)
- Il peut y avoir un nombre exponentiel de siphons et de pièges minimaux.

Modèle Booléen

S peut être représenté par un vecteur \vec{V} de $\{0, 1\}^n$ tel que pour tout $i \in \{1, 2, \dots, n\}$, $V_i = 1$ si et seulement si $p_i \in S$.

- La contrainte de siphon peut être formulée comme suit :

$$\forall i, V_i = 1 \Rightarrow (\forall t \in T, t \in \bullet p_i \Rightarrow t \in (\cup_{V_j=1} \{p_j\})^\bullet)$$

Ce qui peut être écrit comme :

$$\forall i, V_i = 1 \Rightarrow \bigwedge_{t \in \bullet p_i} \left(\bigvee_{p_j \in \bullet t} V_j = 1 \right)$$

- Exclure l'ensemble vide, la contrainte suivante est ajoutée :

$$\bigvee_i V_i = 1$$

Stratégie de recherche

- Enumérer uniquement les siphons minimaux (vis-à-vis de l'inclusion ensembliste) peut être assuré par la stratégie de recherche et l'ajout de nouvelles contraintes : étiqueter sur les variables booléennes, par valeur croissante (d'abord 0, puis 1) et parcourir l'arbre en profondeur d'abord.
- A chaque fois un siphon S est trouvé, la contrainte $\bigvee_{i \in S} V_i = 0$ est ajoutée au modèle afin d'interdire les siphons le contenant dans la suite de la recherche.

Exemple : Codage SAT de l'énumération des siphons minimaux du modèle de la réaction enzymatique

Les variables $X1$, $X2$, $X3$ et $X4$ correspondent respectivement à E , A , AE et B .

Dans la première itération, le problème consiste à résoudre les clauses suivantes :

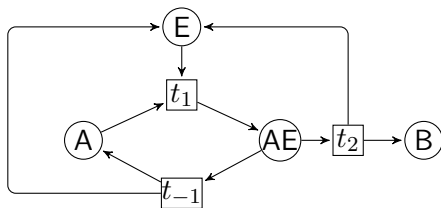
$$(\neg X2 \vee X3) \wedge$$

$$(\neg X3 \vee X1 \vee X2) \wedge$$

$$(\neg X1 \vee X3) \wedge$$

$$(\neg X1 \vee X3) \wedge$$

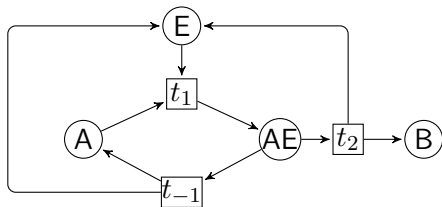
$$(\neg X4 \vee X3)$$



Exemple : Codage SAT de l'énumération des siphons minimaux du modèle de la réaction enzymatique

Les variables $X1$, $X2$, $X3$ et $X4$ correspondent respectivement à E , A , AE et B .

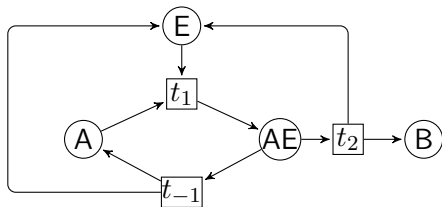
- Le problème est satisfiable :
 $X2$ et $X3$ à *vrai* et $X1$ et $X4$ à *faux*.
 $\Leftrightarrow \{A, AE\}$ est un siphon minimal.
- Une nouvelle clause est ajoutée : soit A soit AE n'appartient pas au siphon.



Exemple : Codage SAT de l'énumération des siphons minimaux du modèle de la réaction enzymatique

Les variables X_1 , X_2 , X_3 et X_4 correspondent respectivement à E , A , AE et B .

- Le programme itère une autre fois.
affectation qui satisfait le problème :
 X_1 et X_3 à *vrai*, X_2 et X_4 à *faux*
 $\Rightarrow \{E, AE\}$ est aussi un siphon minimal.
- Une nouvelle clause est ajoutée : soit E soit AE n'appartient pas au siphon.

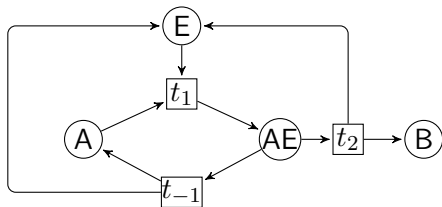


Exemple : Codage SAT de l'énumération des siphons minimaux du modèle de la réaction enzymatique

Les variables X_1 , X_2 , X_3 et X_4 correspondent respectivement à E , A , AE et B .

- Aucune affectation des variables ne peut satisfaire le problème.

Ainsi, ce modèle contient 2 siphons minimaux : $\{A, AE\}$ et $\{E, AE\}$.



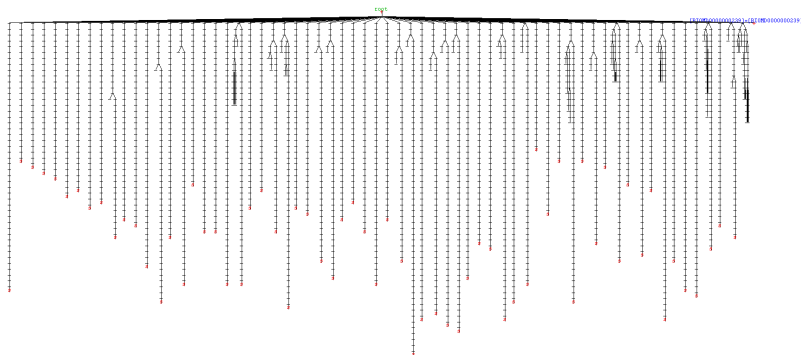
Algorithme PLC(B)

La stratégie utilisée est un Branch-and-Bound sans relance implémentée comme suit :

- 1 A chaque fois qu'un siphon est trouvé, le chemin menant à cette solution est mémorisé.
- 2 La recherche est entièrement défaite dans le but d'ajouter au modèle une nouvelle contrainte de non sur-ensemble.
- 3 Le chemin mémorisé est rejoué pour poursuivre la recherche au point où elle a été arrêtée.

Arbre de recherche

Énumération de 64 siphons minimaux du modèle 239 de biomodels.net contenant 51 espèces et 72 réactions.



Peu de backtracks sont nécessaires pour trouver toutes les solutions.

Performance sur tout le banc d'essai

Database	# model	# pl, # tr min-max	# siphons min-max	siphons size min-max	total time (ms)		
					dedicated algorithm	SAT	GNU Prolog
Biomodels.net	403	1,1-194,313	0-64	1-413	19734	611	195
Petriweb	80	1,1-68,51	0-11	0-7	2325	156	6

✎ Les programmes SAT et PLC(B) résolvent le problème de l'énumération d'un siphon minimal en moins d'une milliseconde en moyenne.

✎ Le programme PLC(B) a les meilleurs performances et dépasse l'algorithme dédié par 2 ordres de grandeur.

Performances sur les instances les plus difficiles

model	# siphons	# places	# transitions	dedicated algorithm	sat	GNU Prolog
Kohn's map of cell cycle	81	509	775	28 ms	1 ms	221 ms
BIOMD000000205	32	194	313	21 ms	1 ms	34 ms
BIOMD000000239	64	51	72	2980 ms	1 ms	22 ms

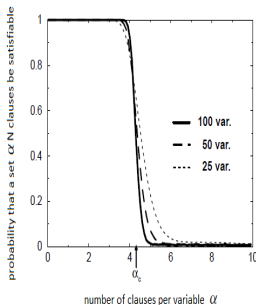
- ✎ Le codage SAT est le plus efficace : il convient parfaitement au nombre de siphons et à la taille du réseau.
- ✎ Le programme PLC(B) a de meilleurs performance que l'algorithme dédié mais semble manipuler moins facilement les réseaux de grande taille.

Pourquoi le calcul des siphons minimaux est-il si facile en pratique ?

- MiniSAT et PLC(B) dépassent en rapidité l'algorithme spécialisé d'au moins un ordre de grandeur et le temps de calcul est étonnamment court sur nos exemples pratiques.
- Pourtant, cette énumération de tous les siphons minimaux résout le problème de décision de l'existence d'un siphon minimal contenant un ensemble donné de places qui a été prouvé NP-difficile.

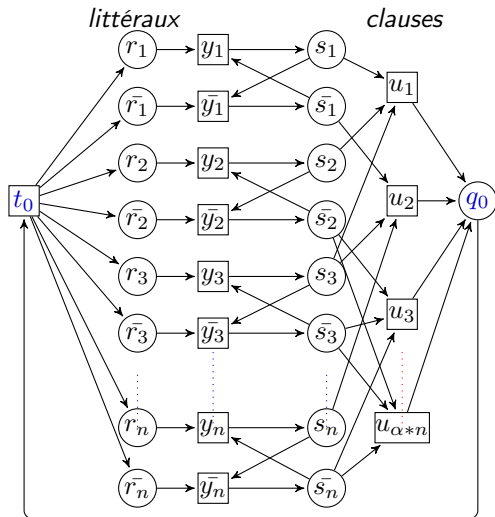
Génération d'instances difficiles

- Phénomène de transition de phase dans 3-SAT. (B. Selman et al 1996)



- Preuve de NP-difficulté par réduction de 3-SAT au problème de l'existence d'un siphon minimal contenant une place donnée.

Réduction de 3-SAT au problème de l'existence d'un siphon minimal contenant q_0



- degré entrant maximal linéaire en nombre de clauses (q_0)
- degré maximal sortant linéaire en nombre de variables pour (t_0).

Temps d'énumération des siphons minimaux dans les réseaux de Petri codant 3-SAT

model	# min siphons conten. q0	Petri net view			3-SAT view			time (ms)
		# places	# transitions	density	# variables	# clauses	α	
pn0.2.xml	>129567	801	441	0,56	200	40	0,2	60000
pn0.6.xml	>32392	801	521	0,65	200	120	0,6	60000
pn1.xml	>87376	801	601	0,751	200	200	1	60000
pn2.xml	>74462	801	801	1	200	400	2	60000
pn3.xml	>63816	801	1001	1,24	200	600	3	60000
pn4.xml	>59827	801	1201	1,49	200	800	4	60000
pn4.2.xml	>41415	801	1241	1,54	200	840	4,2	60000
pn4.4.xml	0	801	1281	1,59	200	880	4,4	1596
pn4.6.xml	0	801	1321	1,64	200	920	4,6	1411
pn5.xml	0	801	1401	1,74	200	1000	5	370
pn6.xml	0	801	1601	1,99	200	1200	6	175
pn7.xml	0	801	1801	2,24	200	1400	7	157
pn8.xml	0	801	2001	2,49	200	1600	8	157
pn9.xml	0	801	2201	2,74	200	1800	9	133
pn10.xml	0	801	2401	2,99	200	2000	10	137

Réseaux de Petri aléatoires avec des degrés linéaires et des degrés bornés

	nb places	nb transitions	density	nb siphons	time
random-degree-pn1	801	80	0,099875156	> 77918	60000
random-degree-pn2	801	320	0,399500624	> 69246	60000
random-degree-pn3	801	560	0,699126092	> 45782	60000
random-degree-pn4	801	801	1	> 28285	60000
random-degree-pn5	801	1041	1,299625468	0	7473
random-degree-pn6	801	1281	1,599250936	0	11233
random-degree-pn7	801	1521	1,898876404	0	15040
random-degree-pn8	801	1762	2,199750312	0	9548
random-degree-pn9	801	2242	2,799001248	0	13807
bounded-degree-pn1	801	80	0,099875	377	120
bounded-degree-pn2	801	320	0,399501	250	55
bounded-degree-pn3	801	560	0,699126	146	32
bounded-degree-pn4	801	801	1	66	14
bounded-degree-pn5	801	1041	1,299625	29	11
bounded-degree-pn6	801	1281	1,599251	13	3
bounded-degree-pn7	801	1521	1,898876	4	5
bounded-degree-pn8	801	1762	2,199750312	1	1
bounded-degree-pn9	801	2242	2,799001248	0	2

Degrés dans les réseaux de Petri de biomodels.net

Les réseaux de Petri associés aux modèles de Biomodels.net ont un degré entrant et sortant borné pour les places et les transitions même si le modèle est large.

	average
number of arcs	92
Avg-indegree-places	1,89
Avg-indegree-transitions	1.06
Max-indegree-places	5,94
Max-indegree-transitions	2,72
Avg-outdegree-places	1,93
Avg-outdegree-transitions	0,99
Max-outdegree-places	5,53
Max-outdegree-transitions	2,87

Degrés dans les réseaux de Petri de biomodels.net

Le modèle 175 représente un cas exceptionnel.

model	# siphons	# places	# transitions	dedicated algorithm	sat	GNU Prolog
BIOMD000000175	3042	118	194	∞	137000 ms	∞

Forte connectivité sur les places avec des espèces qui sont les réactants de 32 réactions et le produit de 31 réactions et de grands degrés.

	average	model 175
number of arcs	92	1125
Avg-indegree-places	1,89	5
Avg-indegree-transitions	1.06	3
Max-indegree-places	5,94	31
Max-indegree-transitions	2,72	11
Avg-outdegree-places	1,93	5
Avg-outdegree-transitions	0,99	3
Max-outdegree-places	5,53	32
Max-outdegree-transitions	2,87	11

Conclusion

- Nous avons décrit un modèle booléen pour le problème de l'énumération des siphons et des pièges minimaux d'un réseau de Petri.
- Les solveurs SAT et PLC(B) sont plus efficaces que l'algorithme dédié de deux ordres de grandeur.
- Nous avons proposé une explication de l'efficacité étonnante de miniSAT et de PLC sur les modèles biologiques réels : les degrés des réseaux de Petri associés sont bornés par de petites valeurs.

Merci pour votre attention