

RAPPORT DE STAGE DE L3

ANALYSE DE LA DYNAMIQUE DES MODÈLES BIOLOGIQUES PAR PROGRAMMATION LOGIQUE

Léo-Paul DELSAUX[†]

Encadrant du stage
Maxime FOLSCHETTE[‡]

Juin-Août 2022

[†]. ENS de Lyon

[‡]. Équipe Bio-Computing, laboratoire CRIStAL, CNRS de Lille

Table des matières

1	Introduction	2
2	Answer Set Programming	2
2.1	Termes	2
2.2	Règles et modèles d'un programme	3
2.3	Variables	3
2.4	Agrégats	4
2.5	Premier exemple : le sudoku	4
2.6	Scripting Python	5
2.7	Second exemple : le sokoban	5
3	Formalismes d'automates et leur dynamique	7
3.1	AAN et leur traduction en ASP	7
3.2	Sémantiques	8
3.2.1	Asynchrone	8
3.2.2	Synchrone	9
3.2.3	Généralisée	10
3.3	Dynamique	10
4	Contribution personnelle	11
4.1	Code pré-existant	11
4.2	Nouveau code	13
5	Conclusion	13

1 Introduction

La biologie s'intéresse à l'étude des systèmes vivants ou composant le vivant (gènes, cellules, écosystèmes...). La bioinformatique est l'utilisation de méthodes informatiques pour aider la biologie. Un des domaines de la bioinformatique s'appelle la *biologie des systèmes* et consiste à proposer des modèles et des méthodes informatiques et mathématiques pour représenter et étudier de tels systèmes.

Ainsi, des études biologiques permettent de déterminer comment certains gènes interagissent entre eux, ce qui peut être représenté mathématiquement sous la forme d'un graphe. De plus, ces interactions peuvent être vues comme des transitions entre les sommets de notre graphe, ce qui se rapproche alors plus d'un automate. La représentation du modèle étudié lors de ce stage est explicité dans le chapitre Modèle.

L'étude de la dynamique de systèmes biologiques lève plusieurs problèmes tels que l'identification d'attracteurs, les bifurcations ou encore la connexité entre deux états globaux. Dans ce rapport je vais discuter de la recherche d'attracteurs dans des réseaux d'automates asynchrones : il s'agit d'un ensemble d'états duquel on ne peut pas s'échapper et minimal au sens de l'inclusion.

Answer Set Programming (= ASP) est un paradigme de programmation logique particulièrement efficace pour résoudre des problèmes combinatoires. Je n'ai travaillé qu'avec ce langage (et un peu de scripting python) durant mon stage afin de chercher les attracteurs parmi le graphe des états globaux de notre réseau.

Ce qui sera étudié dans ce rapport ne le sera que partiellement : on n'entrera pas en détail dans le code présenté, mais on se penchera simplement sur les grandes lignes et sur les points techniques qu'il arbore.

2 Answer Set Programming

L'Answer Set Programming (ASP) est un paradigme de programmation logique comparable à Prolog. Ces dernières décennies, ASP s'est trouvé être puissant pour traiter des modèles biologiques, permettant de parcourir un grand nombre de configurations rapidement. ASP peut énumérer facilement les ensembles solutions à un problème qu'on lui encode : c'est un paradigme très efficace pour la combinatoire. Nous allons ici présenter brièvement son fonctionnement, et plus précisément, les outils qui ont été utiles pour mon stage.

2.1 Termes

ASP fonctionne à l'aide de déclarations de faits, de règles, de prédicats contenant un ou plusieurs arguments tels que : *parentOf(jenny, charles)*, ce qui nécessite tout d'abord que j'introduise la notion de **terme** (ou **atome**). On définit un terme de la manière suivante :

1. terme simple :
 - un entier relatif (0, -35, 42)
 - une constante qui démarre par une lettre minuscule (vous, lis35, quatre, ex3mpl35)
 - une chaîne de caractères encadrée de guillemets ("ma_chaine_de_caracteres", "42")
 - une variable identique grammaticalement aux constantes, mais dont la première lettre est majuscule (N0u5, S0mm35, D35, V4r14bl35)

- un (ou plusieurs) underscore symbolisant une variable sans nom ($-$, $--$)
- 2. une fonction (qu'on appellera également prédicat) est de la forme : $\text{constante}(t_1, t_2, \dots, t_k)$ avec un nombre de termes (qui sont les arguments) fini qui peut être nul ($\text{f}(23)$, $\text{je_suis_une_fonction}(42, \text{true})$, $\text{et_moi_aussi}(\text{oui})$).
Ces fonctions n'ont pas vocation à calculer, mais uniquement à déclarer des objets avec les arguments souhaités comme étant vrais.
- 3. un tuple de la forme : (t_1, t_2, \dots, t_k) avec un nombre de termes fini qui ne peut être nul ($(37, \text{oui})$, $()$, $(1, \text{h}(-4), 0, 1)$).

2.2 Règles et modèles d'un programme

Un *programme d'ensemble solutions* (= *answer set program*) est un nombre fini de règles de la forme :

$$a_0 \text{ :- } a_1, \dots, a_m, \text{not } a_{m+1}, \dots, \text{not } a_n.$$

avec $n \geq m \geq 0$. a_0 est un atome ou \perp (Bottom/le Faux) et représente ici la **tête** de la règle. Les a_1, \dots, a_n sont des atomes et représente quant à eux le **corps** de la règle, et le symbole "*not*" représente la négation par l'échec. Cette règle se lit intuitivement : si les atomes a_1, \dots, a_m sont tous vrais et qu'aucun des atomes a_{m+1}, \dots, a_n n'est vrai, alors a_0 est vrai.

Si $n = m = 0$, cela signifie que a_0 est vrai. Dans ce cas, on parle d'un fait, et on ne doit pas renseigner " :- ". D'une autre part, si $a_0 = \perp$, on parle de contrainte : comme \perp ne peut jamais être vrai, si le corps de la règle est vrai, cela invalide la solution actuelle. On ne renseigne pas \perp pour la tête, on laisse une tête vide au niveau du code.

Cette notion de "solution actuelle" se formalise de la façon suivante : on parle d'une **interprétation** I en tant qu'ensemble fini d'atomes propositionnels. Une règle r définie comme ci-dessus est *vraie dans* I si et seulement si :

$$\{a_1, \dots, a_m\} \subseteq I \wedge \{a_{m+1}, \dots, a_n\} \cap I = \emptyset \Rightarrow a_0 \in I$$

Si toutes les règles d'un programme P sont vraies dans une même interprétation I et que I est maximale, alors on dit que I est un **modèle** de P . Lorsque l'on déclare un problème P en ASP, le solveur nous renverra en sortie tout les modèles possibles pour P .

2.3 Variables

Si une variable apparaît dans un atome de la tête, elle doit également être dans le corps. Chaque instance d'une variable va être groundée par clingo afin de trouver tous les modèles à notre problème. Prenons un petit exemple, voici un problème encodé avec 2 règles :

parentOf(jenny, charles).
parentOf(mary, jenny).

Ce programme contient 2 faits : Charles est un parent de Jenny, et Jenny est un parent de Mary. C'est ainsi que l'on souhaite comprendre les termes "*parentOf(jenny, charles)*" et "*parentOf(mary, jenny)*". Si on souhaite désormais spécifier que Charles est un grand parent de Mary, on peut le faire en ajoutant le terme : "*grandparentOf(mary, charles)*". Cependant, ASP est capable de faire de la déduction. En effet, on peut généraliser la notion de grand parent comme étant le parent d'un parent de la façon suivante :

grandparentOf(X, Z) :- parentOf(X, Y), parentOf(Y, Z).

Pour chaque valeurs possible de X, Y, et Z, le grounding d'ASP va nous créer des règles associées. Ainsi, les 27 règles dont *grandparentOf(mary, mary) :- parentOf(mary, mary), parentOf(mary, mary)*, ou encore *grandparentOf(charles, mary) :- parentOf(charles, jenny), parentOf(jenny, mary)* vont être créées par cette façon de procéder. Ceci n'est pas du tout un problème : le solveur d'ASP va simplement regarder si les atomes des queues sont vraies ou pas. Si ce n'est pas le cas, comme le faux implique tout, les règles seront vraies. Si c'est le cas, la tête sera alors mise à vraie. Sachant qu'une seule de ces 27 possibilités mène à la véracité d'un nouvel atome, on obtiendra en sortie d'ASP les faits suivants :

SATISFAIBLE
parentOf(jenny, charles)
parentOf(mary, jenny)
grandparentOf(mary, charles)

2.4 Agrégats

Les agrégats forment le dernier outil d'ASP que j'ai utilisé dans le cadre de mon stage. Il s'agit d'un moyen de sélectionner un certain nombre d'atomes parmi un ensemble, et de les mettre à vrai. Ils se structurent suivant l'exemple suivant :

0 {coloration(X,Y,Teinte) : couleur(Teinte)} 1 :- abscisse(X), ordonnee(Y).

La **borne inférieure**, située à gauche des accolades (ici 0), et la **borne supérieure** (ici 1) définissent l'intervalle discret du nombre d'atomes que le solveur peut mettre à vrai. Si l'une de ces deux bornes est omise, ASP comprend qu'il s'agit de la borne maximale (ou minimale) possible (à savoir 0 pour la borne inf et cardinal de l'ensemble défini par les accolades pour la borne sup). Dans ce petit exemple, on choisira donc une ou 0 couleur parmi celles possibles (telles qu'il existe un atome *couleur(Teinte)* avec la Teinte souhaitée qui soit vrai) pour chaque abscisse et ordonnée définies.

Les agrégats permettent de faire des disjonctions de cas.

2.5 Premier exemple : le sudoku

Durant les premières semaines de mon stage, j'ai encodé différents jeux de logique en ASP afin de me familiariser avec le langage. Le jeu le plus simple et connu que j'ai encodé a été le sudoku. Dans cette section, je vais brièvement détailler le fonctionnement de ce code.

On codera une grille de sudoku en ASP en utilisant *s*, une fonction d'arité 3 prenant R, C et V comme argument, où R est la ligne, C la colonne et V la valeur de la case (R,C) dans la grille.

val(1..9).

1..9 est un raccourci pour dire itérer pour les entiers allant de 1 à 9. Ici, on déclare qu'il y a 9 valeurs dans une fonction d'arité 1 : 'val'

border(1; 4; 7).

On déclare également 3 autres faits pour mémoriser où sont situés les bordures de notre grille car le point virgule permet de déclarer plusieurs atomes en un

1 {s(R, C, V) : val(V)} 1 :- val(R); val(C).

On ne prend qu'une seule valeur par carré

1 {s(R, C, V) : val(R)} 1 :- val(C); val(V).

Une valeur ne peut pas apparaître plusieurs fois dans la même colonne

$1 \{s(R, C, V) : val(C)\} 1 :- val(R); val(V).$

Ni dans la même ligne

$1 \{s(R, C, V) : val(R), val(C), R1 \leq R, R \leq (R1 + 2), C1 \leq C, C \leq (C1 + 2)\} 1 :- val(V); border(R1); border(C1).$

Une valeur ne peut pas apparaître plusieurs fois dans une sous-grille.

Une fois que l'on a fait cela, il ne nous reste plus qu'à instancier notre problème sur une grille. Pour cela, on renseigne $s(R, C, V)$ pour chaque case déjà pré-numérotée de la grille à notre code, et le solveur se chargera de nous renvoyer l'intégralité de la grille complétée.

2.6 Scripting Python

En ASP, on peut écrire des morceaux de script en Lua ou en Python. Pour cela, il suffit d'ajouter la commande '`#script (mon_langage)`', puis de taper son code dans le langage que l'on a choisi, et de finir le script par la commande '`#end.`'. Le scripting admet trois atouts majeurs :

- L'appel au fonction en ASP via la commande '`@ma_fonction(et, ses, arguments)`', permettant d'effectuer des calculs sous Python et de mettre des variables à vrai en fonction de ce qui a été trouvé dans ces calculs.
- Un contrôle de la résolution permettant le filtrage des ensembles solutions, l'ajout de faits/règles, la demande de grounding et de solving. C'est ce dont je me suis le plus servi durant mon stage.
- Une méthode incrémentale efficace lorsque l'on a besoin d'avoir une notion de temporalité ou bien un calcul qui se fait dans un certain ordre : en fonction des résultats trouvés on peut partir sur une autre piste. La notion d'ordre étant totalement inexistante en ASP (les lignes de code peuvent être écrites dans n'importe quel ordre), cela est pratique.

2.7 Second exemple : le sokoban

Le sokoban a été le dernier jeu de logique que j'ai encodé en ASP, et le plus complexe à cause de la duplication très couteuse du nombre de coups nécessaires pour la résolution d'une grille.

Le jeu est composé d'une grille possédant des cases qui sont du sol ou du mur. Sur une case représentant le sol, il ya le joueur. Sur plusieurs autres cases de sol il y a des caisses. Sur autant de cases de sol qu'il y a de caisses, on trouve des cases d'arrivée (qui sont également du sol). Le but du jeu est de pousser toutes les caisses sur les cases d'arrivée sachant que le joueur ne peut se déplacer qu'en haut, à gauche, en bas ou à droite, et ne peut pas traverser les murs ni pousser une caisse s'il y a un mur ou une autre caisse derrière celle-ci.

Une grille de sokoban est encodée en spécifiant quelles cases sont des murs, et quelles cases sont des sols avec les prédicats $mur(X, Y)$ et $sol(X, Y)$. On renseigne de plus un prédicat $init(X, Y)$, ce qui nous donne la position initiale du joueur, et $caisse_init(Numero_caisse, X, Y)$ nous renseignant sur les positions initiales des caisses (le premier argument varie de 1 à n lorsque l'on a n de ces caisses). Enfin, le prédicat $arrivee(X, Y)$ symbolisera une case d'arrivée en X, Y .

J'ai fait différentes versions pour ce jeu. L'une d'entre elles utilise une méthode incrémentale. Pour l'importer, il nous suffit d'ajouter la commande '`#include <incmode>.`'. On doit alors définir les sous-programmes $base$, $step(k)$ et $check(k)$.

- **base :**
 1. *perso*(0, *R*, *C*) :- *init*(*R*, *C*).
 - % prédicats pour mémoriser la position du joueur à l'étape 0
 2. *caisse_a_instant*(0, *R*, *C*) :- *caisse_init*(-, *R*, *C*). % ainsi que celles des caisses
- **step(k) :**

Pour des raisons de lisibilité, je ne mettrai qu'une instruction pour une seule direction

 3. *acces_a_instant*(*k*-1, *R*, *C*) :- *perso*(*k*-1, *R*, *C*).
 - % prédicats pour mémoriser toutes les cases accessibles
 4. *acces_a_instant*(*k*-1, *R*+1, *C*) :- *sol*(*R*+1, *C*), *acces_a_instant*(*k*-1, *R*, *C*), *not* *caisse_a_instant*(*k*-1, *R*+1, *C*).
 - % on ajoute les cases à droite de celles accessibles (idem pour les 3 autres directions)
 5. 1 { *coup_a_instant*(*k*, (*R*, *C*), (*R*+1, *C*)) : % on peut faire un coup à l'instant *k*
 6. *sol*(*R*, *C*), *sol*(*R*+2, *C*), % si l'on a deux cases de *sol* *c*₁ et *c*₂
 7. *acces_a_instant*(*k*-1, *R*, *C*), % tels que l'on ait accès à *c*₁
 8. *caisse_a_instant*(*k*-1, *R*+1, *C*), % qu'il y ait une caisse entre *c*₁ et *c*₂
 9. *not* *caisse_a_instant*(*k*-1, *R*+2, *C*); % et qu'il n'y a pas de caisse en *c*₂
 10. [...] % idem pour les 3 autres directions
 11. } 1.
 12. *perso*(*k*, *R*2, *C*2) :- *coup_a_instant*(*k*, -, (*R*2, *C*2)). % Déplacement du personnage
 13. *caisse_a_instant*(*k*, 2**R*2-*R*, 2**C*2-*C*) :- % On pousse une seule caisse : elle bouge
 14. *sol*(*R*2, *C*2), *sol*(*R*, *C*), *sol*(2**R*2-*R*, 2**C*2-*C*),
 15. *caisse_a_instant*(*k*-1, *R*2, *C*2), *coup_a_instant*(*k*, (*R*, *C*), (*R*2, *C*2)).
 16. *caisse_a_instant*(*k*, *R*, *C*) :- % On ne pousse pas une caisse : elle ne bouge pas
 17. *sol*(*R*, *C*), *sol*(*R*2, *C*2),
 18. *caisse_a_instant*(*k*-1, *R*, *C*), *perso*(*k*, *R*2, *C*2), *R*2 != *R*.
 19. *caisse_a_instant*(*k*, *R*, *C*) :- % Idem, mais pour les colonnes
 20. *sol*(*R*, *C*), *sol*(*R*2, *C*2),
 21. *caisse_a_instant*(*k*-1, *R*, *C*), *perso*(*k*, *R*2, *C*2), *C*2 != *C*.
 22. *caisse_pas_place*(*k*) :- *caisse_a_instant*(*k*, *R*, *C*), *not* *arrivee*(*R*, *C*).
 - % On détermine s'il existe encore des caisses qui ne sont pas placées pour l'étape *k*
 23. *fini*(*k*) :- *not* *caisse_pas_place*(*k*).
 - % Si toutes les caisses sont placées, on a fini la grille
- **check(k) :**
 24. *instant*(*k*) :- *k* != 0. % si on est à un instant *k* ≥ 1
 25. :- *not* *fini*(*k*), *query*(*k*). % que l'on a fini à cette étape
 26. :- *query*(*k*), *k* = 0. % et qu'on n'a pas *k* = 0

% Alors $check(k)$ est vrai, ce qui a pour conséquence d'arrêter les appels successifs à clingo (le grounder et solveur d'ASP), ce qui fait arrêter le programme

On récupère alors la liste des coups de caisses à faire, et on peut résoudre notre grille en comblant les coups qui déplaçaient les caisses par ceux nécessaires pour atteindre la prochaine caisse à déplacer.

3 Formalismes d'automates et leur dynamique

Différents modèles existent pour représenter efficacement un système biologique et manipuler sa dynamique : les deux principaux sont les réseaux booléens synchrones de Stuart Kauffman, et les réseaux asynchrones de René Thomas. Dans le cadre de mon stage je ne parlerai que des réseaux asynchrones, et plus particulièrement des réseaux d'automates asynchrone (AAN) [Folschette u. a. (2015)] et [Paulevé (2016)], qui forment une extension d'une précédente structure appelé "Process Hitting" [Paulevé u. a. (2014)].

3.1 AAN et leur traduction en ASP

Un automate A , dans le contexte de mon stage, sera défini comme étant un ensemble d'états $q_0, q_1, \dots, q_{|A|-1}$ avec des transitions dont les étiquettes seront un ou plusieurs états (qu'on appellera également *niveaux*) d'automates externes. Il n'y a donc ni état final, ni état initial. On ne lui donnera pas non plus de mot à lire en entrée puisqu'on va s'intéresser à la dynamique de nos ensembles d'automates. On parlera alors de transition locale : $q_i \xrightarrow{l} q_j$ symbolisera le fait que l'on peut passer de l'état q_i à l'état q_j si toutes les conditions de l sont vérifiées.

Un réseau d'automates asynchrone est un triplet (Σ, S, T) avec :

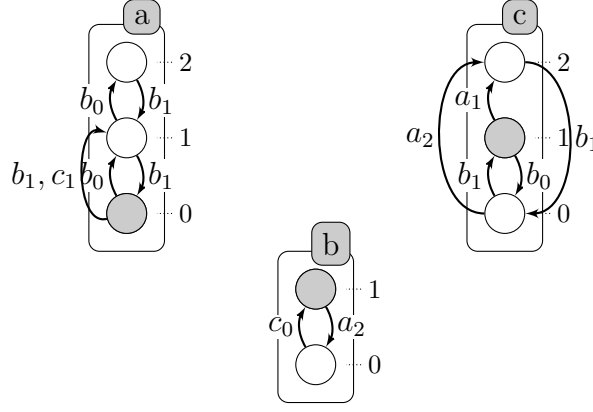
- $\Sigma = \{a, b, \dots\}$ est un ensemble fini d'automates non vides.
- Si C_a est le nombre d'états d'un automate a , alors $S_a = \{a_0, a_1, \dots, a_{C_a-1}\}$ est l'ensemble des **états locaux** de l'automate a . $S = \prod_{a \in \Sigma} S_a$ est l'ensemble fini des **états globaux**, et

$LS = \bigcup_{a \in \Sigma} S_a$ représente l'ensemble de tous les états locaux.

- Pour chaque $a \in \Sigma$, $T_a \subseteq \left\{ a_i \xrightarrow{l} a_j \in S_a \times \rho(LS/S_a) \times S_a \mid a_i \neq a_j \right\}$ est l'ensemble des **transitions locales** d'un automate a , avec ρ qui désigne la puissance ensembliste. $T = \bigcup_{a \in \Sigma} T_a$ est l'ensemble des transitions locales du modèle.

Exemple : On représente l'AAN suivant de cette manière :

- $\Sigma = \{a, b, c\}$
- $S_a = \{a_0, a_1, a_2\}$, $S_b = \{b_0, b_1\}$ et $S_c = \{c_0, c_1, c_2\}$
- $T_a = \left\{ a_0 \xrightarrow{b_0} a_1, a_0 \xrightarrow{b_1, c_1} a_1, a_1 \xrightarrow{b_1} a_0, a_1 \xrightarrow{b_0} a_2, a_2 \xrightarrow{b_1} a_1 \right\}$
 $T_b = \left\{ b_0 \xrightarrow{c_0} b_1, b_1 \xrightarrow{a_2} b_0 \right\}$
 $T_c = \left\{ c_0 \xrightarrow{b_1} c_1, c_0 \xrightarrow{a_2} c_2, c_1 \xrightarrow{b_0} c_0, c_1 \xrightarrow{a_1} c_2, c_2 \xrightarrow{b_1} c_0 \right\}$



En ASP, on définira un AAN en deux temps :

- On commencera par déclarer chacun de nos automates avec les niveaux (= états) qu'il contient :

automaton_level("a", 0..2).

automaton_level("b", 0..1).

automaton_level("c", 0..2).

- Enfin, les transitions seront encodées via des labels (t1, t2, ...), et on donnera chacune des conditions, ainsi que l'état d'arrivée, via un fait :

condition(t1, "a", 0). target(t1, "a", 1). condition(t1, "b", 0).

condition(t2, "a", 1). target(t2, "a", 2). condition(t2, "b", 0).

condition(t3, "a", 2). target(t3, "a", 1). condition(t3, "b", 1).

condition(t4, "a", 1). target(t4, "a", 0). condition(t4, "b", 1).

condition(t5, "b", 0). target(t5, "b", 1). condition(t5, "c", 0).

condition(t6, "b", 1). target(t6, "b", 0). condition(t6, "a", 2).

condition(t7, "c", 0). target(t7, "c", 1). condition(t7, "b", 1).

condition(t8, "c", 1). target(t8, "c", 0). condition(t8, "b", 0).

condition(t9, "c", 1). target(t9, "c", 2). condition(t9, "a", 1).

condition(t10, "c", 0). target(t10, "c", 2). condition(t10, "a", 2).

condition(t11, "c", 2). target(t11, "c", 0). condition(t11, "b", 1).

condition(t12, "a", 0). target(t12, "a", 1). condition(t12, "b", 1). condition(t12, "c", 1).

Nous avons ainsi défini l'automate de notre **exemple** en ASP.

3.2 Sémantiques

Soit $R = (\Sigma, S, T)$ un AAN (défini ci-dessus). On définit la notion de transition jouable :

- On dit qu'une transition locale est **jouable** si toutes les conditions de celle-ci sont vérifiées

On notera P_ζ l'ensemble des transitions locales jouables depuis un état global ζ

L'état global (a_0, b_1, c_1) de l'**exemple** admet une seule transition jouable : $a_0 \xrightarrow{b_1, c_1} a_1$

La dynamique d'un AAN se définit à l'aide de sa **sémantique**. La sémantique forme l'ensemble des propriétés définissant les **transitions globales jouables** : dans un AAN on s'intéresse à l'évolution globale du réseau.

3.2.1 Asynchrone

Soit $R = (\Sigma, S, T)$ un AAN, et $\zeta \in S$ un état global. L'ensemble des transitions globales jouables depuis ζ pour la sémantique *asynchrone* est donnée par :

$$U^{syn}(\zeta) = \left\{ \left\{ a_i \xrightarrow{l} a_j \right\} \mid a_i \xrightarrow{l} a_j \in P_\zeta \right\}$$

De manière informelle : chaque transition locale jouable est une transition globale.

En ASP, on commence donc par déclarer que deux états globaux sont différents sur un automate à l'aide de la règle suivante :

1. *different_on(Gs1, Gs2, Automaton) :- % Cet atome est vrai*
2. *global_state(Gs1), global_state(Gs2), Gs1 != Gs2, % si deux états globaux diffèrent*
3. *automaton(Automaton), % qu'il existe un automate*
4. *active_in_g(level(Automaton, LevelI), Gs1), % sur lequel Gs1 vaut I*
5. *active_in_g(level(Automaton, LevelJ), Gs2), % et sur lequel Gs2 vaut J*
6. *LevelI != LevelJ. % avec I ≠ J*

Ensuite, on fait comprendre à ASP qu'une transition locale n'est pas jouable si l'une (au moins) de ses conditions n'est pas vérifiée de la façon suivante :

1. *unplayable(Transition, Gs) :- % Cet atome est vrai*
2. *local_transition(Transition), % s'il existe une transition locale*
3. *global_state(Gs), % un état global*
4. *automaton(Automaton), % et un automate*
5. *active_in_g(level(Automaton, LevelI), Gs), % sur lequel Gs vaut I*
6. *condition(Transition, Automaton, LevelJ), % alors qu'il devrait valoir J ≠ I*
7. *LevelI != LevelJ. % pour que la transition soit jouable*

On déclare ensuite que deux états globaux diffèrent sur au moins un autre automate qu'*Automaton* avec le terme *not_equal_except(Automaton, Gs1, Gs2)* comme ceci :

1. *not_equal_except(Automaton, Gs1, Gs2) :- % Cet atome est vrai*
2. *automaton(Automaton), automaton(Automaton2), % s'il existe deux automates*
3. *Automaton != Automaton2, % qui diffèrent*
4. *global_state(Gs1), global_state(Gs2), % deux états globaux*
5. *different_on(Gs1, Gs2, Automaton2). % tels que Gs1 ≠ Gs2 sur Automaton2*

On peut enfin définir ce qu'est une transition globale jouable à l'aide de nos trois précédentes règles :

1. *playable(Gs1, Gs2) :- % Cet atome est vrai*
2. *global_state(Gs1), global_state(Gs2), % s'il existe deux états globaux Gs1 et Gs2*
3. *automaton(Automaton), local_transition(Transition), % un automate et une transition*
4. *target(Transition, Automaton, LevelJ), % qui fait changer le niveau de l'automate*
5. *not unplayable(Transition, Gs1), % et avec la transition (locale) qui est jouable*
6. *not not_equal_except(Automaton, Gs1, Gs2), % et Gs1 = Gs2 excepté sur Automaton*
7. *active_in_g(level(Automaton, LevelI), Gs1), % tel que le niveau d'Automaton dans Gs1*
8. *active_in_g(level(Automaton, LevelJ), Gs2), % et celui dans Gs2*
9. *LevelI != LevelJ. % sont différents*

3.2.2 Synchrone

Soit $R = (\Sigma, S, T)$ un AAN, et $\zeta \in S$ un état global. L'ensemble des transitions globales jouables depuis ζ pour la sémantique *synchrone* est donnée par :

$$U^{syn}(\zeta) = \{u \subseteq T \mid u \neq \emptyset \wedge \forall a \in \Sigma, (P_\zeta \cap T_a = \emptyset \Rightarrow u \cap T_a = \emptyset) \wedge (P_\zeta \cap T_a \neq \emptyset \Rightarrow |u \cap T_a| = 1)\}$$

De manière informelle : tous les automates possédant au moins une transition locale jouable doivent changer de niveau.

Pour coder cela en ASP, on peut se servir des atomes *different_on* et *unplayable*.

On ajoute de plus une autre règle : *has_playable(Automaton, Gs, LevelI, LevelJ)* spécifiant que l'on peut faire changer *Automaton* dans *Gs* du niveau *I* vers le niveau *J*.

Enfin, il ne nous reste plus qu'à déclarer (par la négation) quelles sont les transitions globales qui ne sont pas jouables. Pour cela, on essaie pour toutes les paires d'états globaux si on peut faire la transition de l'un à l'autre en créant un terme *not_playable(Gs1, Gs2)* si la transition ne peut pas avoir lieu. On énumère alors les différents cas de figure :

1. si *Automaton* n'admet aucune transition locale jouable dans *Gs1* et que $Gs1 \neq Gs2$ sur *Automaton*
2. si *Automaton* admet une (ou plusieurs) transition locale jouable dans *Gs1* et qu'il n'existe pas de transition pour faire passer le niveau d'*Automaton* sur *Gs1* à celui de *Gs2*
3. si aucune transition locale n'est jouable depuis *Gs1*

alors la transition $Gs1 \rightarrow Gs2$ n'est pas jouable.

Enfin, il nous suffit de tester pour toutes les paires d'états globaux distincts si on a *not_playable(Gs1, Gs2)* ou pas. Dans le cas contraire, alors on crée l'atome *playable(Gs1, Gs2)*.

3.2.3 Généralisée

Soit $R = (\Sigma, S, T)$ un AAN, et $\zeta \in S$ un état global. L'ensemble des transitions globales jouables depuis ζ pour la sémantique *généralisée* est donnée par :

$$U^{gen}(\zeta) = \{u \subseteq T \mid u \neq \emptyset \wedge \forall a \in \Sigma, (P_\zeta \cap T_a = \emptyset \Rightarrow u \cap T_a = \emptyset) \wedge (P_\zeta \cap T_a \neq \emptyset \Rightarrow |u \cap T_a| \leq 1)\}$$

De manière informelle : tous les automates possédant au moins une transition locale jouable peuvent changer de niveau, et au moins l'un d'entre eux doit changer de niveau.

Pour coder cela en ASP, on prend exactement le code de la sémantique synchrone, en remplaçant le second cas de figure par ceci :

2. si *Gs1* diffère de *Gs2* sur *Automaton* et qu'il n'existe pas de transition pour faire passer le niveau d'*Automaton* sur *Gs1* à celui de *Gs2*

Cela permet bien de ne pas prendre une transition par automate qui est jouable puisqu'on va tester des paires d'états globaux qui ont peut-être 3 automates jouables, et si l'on ne fait qu'une seule transition locale cela ne sera pas filtré par notre atome *not_playable(Gs1, Gs2)*.

3.3 Dynamique

Soit U une sémantique. On introduit les dernières notions nécessaires pour la définition d'un attracteur :

- On appelle **état stable** un état global ne possédant aucune transition globale jouable pour U .

Dans l'exemple, aucun état global n'est stable pour la sémantique asynchrone (et donc cela est également vrai pour la sémantique synchrone et généralisée)

- Pour $Gs_1, Gs_2 \in S$, on notera $Gs_1 \rightarrow_U Gs_2$ pour symboliser le fait qu'il existe $u \in U(Gs_1)$ tel que si l'on joue toutes les transitions locales de u , alors Gs_1 devient Gs_2 .

- Un **chemin** est une famille finie $(A_k)_{1 \leq k \leq n}$ et ordonnée d'états globaux tels que pour tout $1 \leq k \leq n-1$, $A_k \rightarrow_U A_{k+1}$.
- Un **cycle** est un chemin $(A_k)_{1 \leq k \leq n}$ tel que $A_1 = A_n$.
- Un **domaine de piège** est un ensemble non vide d'états globaux $S_E \in S$ tel que toute transition globale jouable depuis S_E arrive dans un état global de S_E . Plus formellement :

$$\forall \zeta_1 \in S_E \wedge \forall \zeta_2 \in S, \zeta_1 \rightarrow_U \zeta_2 \Rightarrow \zeta_2 \in S_E$$

Enfin, on peut définir ce qu'est un attracteur.

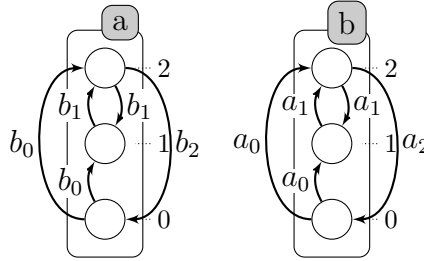
Un ensemble $A \subseteq S$ d'états globaux, avec $|A| \geq 2$ est appelé **attracteur** si et seulement si c'est un domaine de piège minimal en terme d'inclusion ensembliste.

Un travail précédent a démontré le résultat mathématique suivant dont on va se servir pour trouver les attracteurs au sein d'un AAN :

Lemme : Les attracteurs d'un AAN sont exactement les domaines de piège cycliques.

4 Contribution personnelle

La version de laquelle j'ai commencé à travailler est assez proche du travail que j'ai effectué. Cette version était fonctionnelle et efficace pour la sémantique synchrone, mais ne fonctionnait pour la sémantique synchrone que lorsque les attracteurs étaient exactement des cycles simples (chaque état global du graphe produit doit avoir un degré sortant égal à 1), et ne fonctionnait pas pour les autres AAN comme celui-ci :



En effet, l'ensemble $\{(0,0), (1,1), (1,2), (2,1), (2,2)\}$ est un attracteur de cet AAN pour la sémantique synchrone. De plus, il existe deux (même quatre ici) transitions globales jouables depuis l'état global $(0,0)$: on peut aller en $(1,1)$ ou encore en $(2,2)$.

Après avoir identifié le problème au sein du code, j'ai trouvé deux alternatives pour corriger cela : la première consiste à reprendre tout le code, enlever le problème ciblé et corriger cela avec un script Python, et la seconde consiste à considérer les états globaux et à définir les transitions globales jouables dans des fichiers dédiés aux sémantiques.

Lors de mon exploration de la seconde alternative, je me suis confronté à de nombreux problèmes, et la complexité finale de mon algorithme étant piteuse, je ne préfère parler dans ce rapport que de la première alternative.

Je vais donc détailler comment nous arrivons à trouver les attracteurs au sein des AAN en ASP avec la sémantique synchrone.

4.1 Code pré-existant

On commence par renseigner quelques faits utiles pour pouvoir manipuler aisément les informations que l'on a à disposition.

```

%%% Initialisation (faits étendus)
% Noms des automates
1. automaton(Automaton) :- automaton_level(Automaton, _).
% Noms des transitions locales
2. local_transition(Transition, Automaton) :- target(Transition, Automaton, _).
3. local_transition(Transition) :- target(Transition, _, _).

%%% Etapes
% Etapes dans tout le chemin
4. step(0..n).
% Longueur du cycle principal (i.e., un sous-chemin cyclique)
5. 1 { main_cycle_length(N) : step(N), N > 0 } 1.
% Etapes dans le cycle principal
6. cycle_step(0..N) :- main_cycle_length(N).
% Etapes après le cycle principal
7. after_cycle_step(N+1..n) :- main_cycle_length(N).

```

On choisit ensuite un état initial, on définit les transitions non jouables, et la sémantique.

```

%%% Etat initial
% On sélectionne aléatoirement un état initial (step "0")
8. 1 { active(level(Automaton, Level), 0) : automaton_level(Automaton, Level) } 1 :-
9.      automaton(Automaton).

```

```

%%% Préparation pour le calcul des transitions jouables par la sémantique
% Calcule les transitions locales non jouables pour chaque étape
10. unplayable(Transition, Step) :-
11.      active(level(Automaton, LevelI), Step),
12.      condition(Transition, Automaton, LevelJ),
13.      LevelI != LevelJ,
14.      step(Step).

```

```

%%% Sémantique synchrone
% On renseigne les automates qui ont au moins une transition locale jouable
15. has_playable(Automaton, Step) :-
16.      not unplayable(Transition, Step),
17.      local_transition(Transition, Automaton),
18.      step(Step).

```

```

% On sélectionne une transition à jouer pour chaque automate si possible
19. 1 { played(Transition, Step) :
20.      not unplayable(Transition, Step),
21.      local_transition(Transition, Automaton)
22. } 1 :- has_playable(Automaton, Step).

```

```

% Contrainte : on doit jouer au moins une transition locale

```

23. $\text{:- } 0 \{ \textit{played}(-, \textit{Step}) \} 0, \textit{step}(\textit{Step}).$

4.2 Nouveau code

5 Conclusion

Références

- [Ben Abdallah u. a. 2015] BEN ABDALLAH, Emna ; FOLSCHETTE, Maxime ; ROUX, Olivier ; MAGNIN, Morgan : Exhaustive analysis of dynamical properties of Biological Regulatory Networks with Answer Set Programming. In : *Bioinformatics and Biomedicine (BIBM), 2015 IEEE International Conference on IEEE* (Veranst.), 2015, S. 281–285
- [Folschette u. a. 2015] FOLSCHETTE, Maxime ; PAULEVÉ, Loïc ; MAGNIN, Morgan ; ROUX, Olivier : Sufficient conditions for reachability in automata networks with priorities. In : *Theoretical Computer Science* 608 (2015), S. 66–83
- [Paulevé 2016] PAULEVÉ, Loïc : Pint, a static analyzer for dynamics of Automata Networks. In : *14th International Conference on Computational Methods in Systems Biology (CMSB 2016)*, 2016
- [Paulevé u. a. 2014] PAULEVÉ, Loïc ; CHANCELLOR, Courtney ; FOLSCHETTE, Maxime ; MAGNIN, Morgan ; ROUX, Olivier : Analyzing Large Network Dynamics with Process Hitting. In : *Logical Modeling of Biological Systems* (2014), S. 125 – 166. ISBN 978-1-4020-4222-5