

Construire et Exploiter des Convictions sur le Monde à Partir de Régularités d'Interactions Observées

Florian J. Bernard

LIRIS - Université Claude Bernard, Lyon 1
florian.bernard3@etu.univ-lyon1.fr

Résumé Beaucoup de modèles d'intelligence artificielle se représentent sous la forme d'une boucle perception/cognition/action. Cette boucle donne un flux d'interactions qui sont liées aux actions/perceptions que l'agent réalise dans son environnement.

Une interaction est une action ou perception couplé à un résultat que l'agent peut réaliser à travers son « corps ». L'agent ne connaît pas les liens entre les différentes interactions et doit pouvoir les construire pour créer des « objets » qui pourront le renseigner sur l'état du monde. Ces objets peuvent correspondre à une conviction pour l'agent et représente une manière de voir le monde à travers ses interactions pour pouvoir anticiper les futures interactions qui satisferont sa motivation intrinsèque.

En partant de ce constat, comment construire et exploiter des convictions sur le monde à partir de régularités d'interactions observées ?

Table des matières

1	Introduction.....	3
2	L'Apprentissage Développemental.....	4
2.1	Radical Interactionnisme.....	5
2.2	Interaction.....	6
2.3	Motivation.....	8
2.4	Programmation de l'agent.....	8
2.5	L'espace.....	8
3	Énoncé du problème.....	9
4	Les phénomènes et les tables d'usage.....	12
4.1	Tables d'usages.....	12
	Mise à jour des tables d'usages.....	14
	Table d'usage pour les interactions composites.....	14
4.2	Phénomènes.....	14
5	Environnement de travail.....	17
6	Travail à venir.....	19
7	Conclusion.....	19

1 Introduction

Ce stage se place dans le contexte de l'apprentissage développemental. Ce domaine se base sur un agent agnostique, c'est-à-dire, un agent qui ne possède pas d'information sur l'environnement dans lequel il évolue et qui n'a pas de comportement pré enregistré. Il a des méthodes d'apprentissage pour apprendre ses mécanismes sensorimoteurs en interagissant avec son environnement. L'agent est incarné dans un corps et développe des mécanismes cognitifs à travers le couplage du son corps physique et de l'environnement [And03,Bro91].

L'apprentissage développemental s'appuie sur le modèle de l'interactionnisme radical qui permet de construire des arbres d'interactions représentés par la figure n°1. Cette hiérarchie est utilisée pour trouver une suite d'interactions permettant de satisfaire l'agent. Le but de cette approche est de construire des comportements augmentant la motivation interactionnelle de l'agent[GMG12]. Néanmoins, l'agent utilisant cette approche ne possède pas de mécanisme lui permettant de représenter les objets du monde.

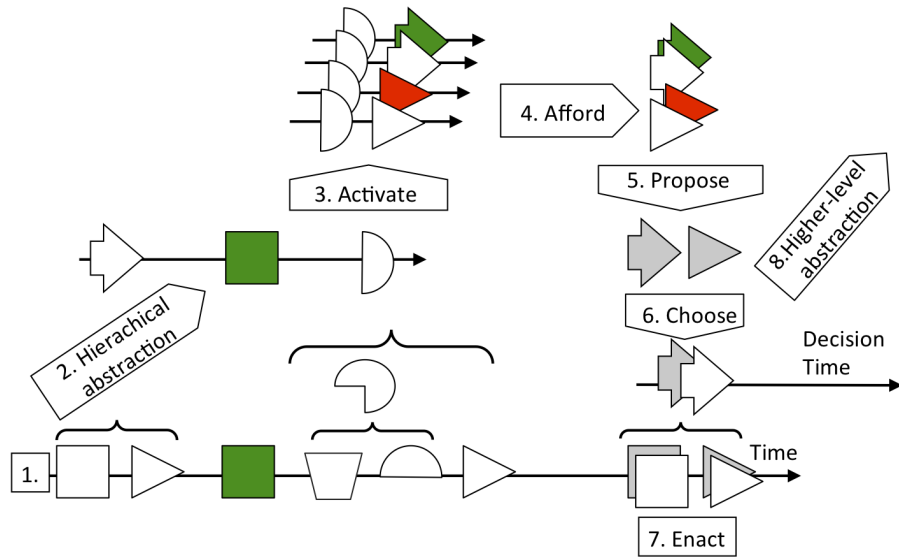


FIGURE 1. Apprentissage hiérarchique des interactions

Comment, à partir d'une suite d'interactions, ne possédant pas de sémantique, un agent agnostique peut construire une représentation des « choses » de son environnement ?

La phénoménologie se présente comme l'une des solutions envisageable pour permettre à un agent d'acquérir des connaissances à partir de l'expérience réali-

sée dans un environnement. La phénoménologie est un courant philosophique qui se concentre sur l'étude des phénomènes, de l'expérience vécue et des contenus de conscience.

La phénoménologie a été introduite par Edmund Husserl en 1913. Elle permet, d'un point de vue philosophique, de définir la conscience des choses par la phrase : toute conscience doit être conçue comme conscience de quelque chose.

« *On ne trouve dans la donnée immédiate [de la conscience] rien de ce qui, dans la psychologie traditionnelle, entre en jeu, comme si cela allait de soi, à savoir : des data-de-couleur, des data-de-son et autres data de sensation ; des data-de-sentiment, des data-de-volonté, etc. Mais on trouve ce que trouvait déjà René Descartes, le cogito, l'intentionnalité, dans les formes familières qui ont reçu, comme tout le réel du monde ambiant, l'empreinte de la langue : le « je vois un arbre, qui est vert ; j'entends le bruissement de ses feuilles, je sens le parfum de ses fleurs, etc. » ; ou bien « je me souviens de l'époque où j'allais à l'école », « je suis inquiet de la maladie de mon ami », etc. Nous ne trouvons là, en fait de conscience, qu'une conscience de...» [Hus76]*

Selon Emmanuel Kant, pour avoir une conscience de quelque chose, l'individu utilise une représentation interne des « choses » du monde qu'on appellera phénomène par la suite. Ces « choses » peuvent représenter un objet physique, une relation entre plusieurs objets ou une entité abstraite. A contrario, le noumène représente l'objet tel qu'il est dans le monde. Le noumène en lui-même est inaccessible à l'individu qui devra utiliser des phénomènes pour se le représenter.

Dans ce stage, nous avons réalisé un agent implémentant les 4 principes ci-dessus, à savoir :

- le concept de l'apprentissage développemental ;
- le modèle de l'interactionnisme radical ;
- la phénoménologie d'Edmund Husserl ;
- la distinction entre les nouèmes et les phénomènes.

L'agent découvre des régularités de son environnement, à travers ses interactions, et les utilise pour construire des phénomènes. Pour visualiser l'état de la mémoire de l'agent, j'ai développé un logiciel gérant la simulation d'agent avec différents algorithmes, un système de visualisation d'interactions et des phénomènes avec la possibilité de tester les agents dans différents environnements.

Expliquer la contribution logiciel, visualiser. en une petite phrase. La suite du rapport comportera LE PLAN ICI.

2 L'Apprentissage Développemental

L'apprentissage développemental a été, dans un certain sens, introduit par Jean Piaget dans les années 50, lorsqu'il travaillait sur la psychologie développementale [Pia59]. En dépit du fait, que les travaux de Piaget ont été menés dans les années 50, l'apprentissage développemental est un nouveau courant de l'intelligence artificielle. Il sollicite des connaissances de plusieurs disciplines comme la robotique, les neurosciences, la psychologie et les sciences cognitives. En apprentissage développemental, les agents :

- ne possèdent pas d'ontologie présumée (i.e : pas de connaissances présumées sur le monde) ;
- connaissent que les possibilités d'interactions avec leur environnement ;
- apprennent leurs propres capacités sensorimotrices de façon autonome ;
- construisent leur propre représentation de l'environnement et qui par définition sera différente de celle que nous pouvons avoir ;
- utilisant le concept d'auto-programmation ¹.

Le but de l'agent est d'apprendre, découvrir, organiser et exploiter des régularités d'interactions dans le temps et l'espace pour favoriser certains comportements et d'adaptabilité.

Un point important de l'apprentissage développemental est que l'agent n'est pas un observateur passif de son environnement, il ne réagit pas à un stimuli donné par son environnement, mais choisit volontairement ce qu'il va observer ou réaliser au travers de ses interactions. Pour développer un comportement cognitif, l'agent a besoin d'être incarné dans un corps couplé à un environnement. On parle de l'*embodied paradigm* [GC14,VT] qui place l'agent comme étant une partie de l'environnement. Les interactions sont les éléments centraux de l'agent, elles lui permettent d'interagir avec son environnement et de construire ses connaissances. La modélisation des interactions est un point important pour la conception de l'agent. Nous avons choisi d'utiliser le modèle de l'interactionnisme radical.

Le modèle de l'interactionnisme radical permet de représenter les actions et les résultats sous forme d'interaction.

2.1 Radical Interactionnisme

Le radical interactionnisme [GA13] cherche à se rapprocher au mieux des modèles cognitifs de l'apprentissage développemental [Pia59,WF13,O'R11]. Dans les modèles cognitifs, l'agent interagit et construit ses connaissances avec son environnement par le biais de ses propres expériences et de leurs résultats.

Le modèle est décrit avec la figure n°2. À l'initialisation, l'agent possède une liste prédéfinie d'interactions sensorimotrices I , appelées **interactions primitives**. À chaque cycle de décisions, l'agent choisit une interaction appelée **interaction intended** qu'il tentera de réaliser dans l'environnement. L'interaction effectivement réalisée dans l'environnement, appelée **interaction enacted** est retournée à l'agent. Les interactions intended et enacted appartiennent à la liste d'interactions sensorimotrices I .

Il est à noter que l'agent ne connaît pas les conséquences des différentes interactions ni sur l'environnement ni sur lui-même et des liens qui peut exister entre chaque interaction et nous ne souhaitons donner aucune heuristique qui pourrait aider l'agent à comprendre ces interactions.

L'environnement représente le couplage entre le corps de l'agent et le monde. Les mécanismes mis en place dans la décision permettent :

1. L'auto-programmation est définie comme étant la possibilité de créer un code exécutable qui peut être exécuté au besoin cf. : <http://liris.cnrs.fr/ideal/mooc/lesson.php?n=041>

- d'apprendre des interactions sur plusieurs niveaux comme dans la figure n°1 tout en évitant l'explosion combinatoire ;
- de sélectionner la prochaine interaction en fonction du contexte d'interaction précédemment réalisée.

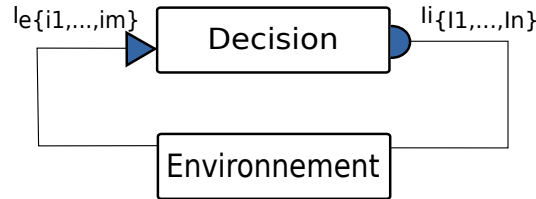


FIGURE 2. Modèle du radical interactionnisme

L'agent étant agnostique, c'est-à-dire, qu'il ne connaît pas du tout l'environnement [GS12], il interagit avec celui-ci uniquement aux travers des interactions qu'il possède. L'agent découvre des suites d'interactions qui sont avantageuses de son point de vue.

Le mécanisme de décision de l'agent ne commence pas par une observation donnée par l'environnement mais par une interaction sélectionnée par l'agent. L'agent est donc pro-actif, il sélectionne la prochaine interaction qu'il souhaite énoncer en prenant en compte l'expérience passée de l'agent et non une fonction de l'environnement.

2.2 Interaction

Dans le contexte des interactions on ne peut pas dissocier une perception d'une action. Elle est composée d'un couple action/résultat défini *a priori* qui est indépendant de l'environnement, du système sensoriel et des actionneurs que possède l'agent. Elle n'a pas de sémantique et ne caractérise pas un état du monde. Les algorithmes mis en place dans la *décision* de l'agent doivent prendre en compte le fait que la connaissance de l'agent se construit uniquement à travers les interactions.

Pour que l'agent puisse apprendre et exploiter ses connaissances, l'algorithme de décision permet d'apprendre des interactions composites sur plusieurs niveaux. Ces interactions composites donnent à l'agent la possibilité de développer des comportements de plus haut niveau.

Une interaction composite est composée d'une pré-interaction et d'une post-interaction. La pré-interaction renseigne sur le contexte dans lequel l'agent se trouve. La figure n°3 indique le processus d'apprentissage et de sélection des interactions composites.

Pour pouvoir réaliser cet apprentissage, l'agent doit pouvoir :

- décomposer une interaction composite intend ;
- reconstruire une interaction composite enact.

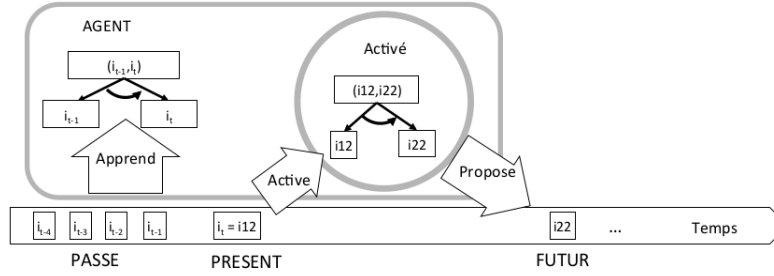


FIGURE 3. Principe d'apprentissage et de sélection d'une interaction composite. Pour sélectionner la prochaine interaction, la décision sélectionne toutes les interactions composites ayant comme pré-interaction la précédente interaction énoncée $\{i12\}$. Puis, regarde l'interaction composite qui a le plus de chance d'être énoncée avec la plus forte motivation interactionnelle. L'interaction composite ainsi sélectionnée a comme pré-interaction l'interaction énoncée $\{i12\}$ et comme post-interaction l'interaction à réaliser.

Une interaction composite est énoncée en une seule fois. Elle peut réussir ou échouer.

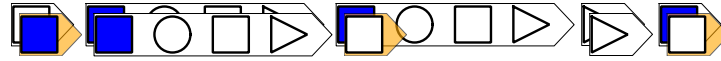


FIGURE 4. Interactions composites intended et enacted

La figure n°4 est une sous partie d'un flux d'interactions réalisées par un agent utilisant les interactions composites. La première interaction composite a été correctement prévue par l'agent. *A contrario*, dans la deuxième interaction, l'agent a essayé l'interaction {carré bleu, rond blanc, carré blanc, triangle blanc} or il a énoncé {carré blanc}. L'interaction composite intended a échoué et l'interaction enacted est {carré blanc}.

Une interaction alternative est une interaction qui peut-être énoncé à la place d'une interaction intended. Dans la figure n°4, l'interaction {carré blanc} est une interaction alternative à l'interaction {carré bleu, rond blanc, carré blanc, triangle blanc}. Cette propriété n'est pas symétrique, par exemple, l'interaction {carré blanc, triangle blanc} n'est pas une interaction alternative de l'interaction {carré bleu}.

Lorsque l'alternative est symétrique c'est-à-dire que deux interactions sont alternatives l'une de l'autre alors on parle d'interaction opposée et de régularité immédiate. Les interactions {carré blanc} et {carré bleu} sont des interactions opposées l'une à l'autre.

Afin que l'agent puisse sélectionner une interaction plutôt qu'une autre, les interactions reposent sur une motivation qui permet de les différencier les unes aux autres.

2.3 Motivation

Pour que l'agent puisse déterminer les interactions intéressantes, le concept de motivation interactionnelle [GMG12] est mis en place dans l'apprentissage développemental. Ce concept permet de définir des interactions *plaisantes* et *déplaisantes* pour l'agent. Grâce à cette motivation l'agent sélectionne une interaction sans autre but que la satisfaction qu'elle peut lui procurer [Ste04]. Elle est implémentée grâce à une valence définie {a priori}. Si la valence est positive (resp. négative), l'interaction est plaisante (resp. déplaisante). Si la valence est nulle, l'interaction en elle-même n'apporte rien à l'agent, par contre couplée avec d'autres interactions, celle-ci peut amener aux interactions plaisantes.

D'autres formes de motivations sont également prises en compte et permettent de construire des comportements différents.

- Il aime prévoir les prochaines interactions à énoncer ;
- Il peut être « curieux » : l'agent va préférer tester les liens entre les différentes interactions pour avoir le plus de choix possibles, et potentiellement avoir la meilleure approche ;
- Il peut être « joueur » : l'agent va répéter des interactions dans le but de les réutiliser par la suite dans un autre contexte.

2.4 Programmation de l'agent

La figure n°5 explicite l'implémentation d'un agent avec le modèle de l'interactionnisme radical. La partie pro-active possède les mécanismes d'apprentissage et de sélection des interactions. Elle sélectionne une interaction (primitive ou composite) I_i qui sera réalisée à l'aide de la partie réactive. L'*enacter* décompose, si besoin, l'interaction intended I_i en interaction primitive $Iip_1 \dots Iip_n$ et l'exécute à travers les différentes interfaces sensorimotrices. Les résultats des capteurs se présentent sous la forme d'interaction primitives énoncées $Iep_1 \dots Iep_m$. La taille de l'interaction composite énoncée peut être plus petite ou égale à la taille de l'interaction intended. L'*enacter* reconstruit l'interaction composite énoncée Ie et la retourne à la décision qui va pouvoir apprendre de nouvelles interactions. Dans la suite du mémoire, nous nous concentrons sur la partie pro-active de l'agent. Nous considérons que le programme réactif et l'environnement sont fonctionnels.

2.5 L'espace

Mes travaux se basent sur la thèse de Simon Gay [Gay14] et sur le MOOC réalisé par Olivier Georgeon². Tous deux ont travaillé sur le modèle de l'interactionnisme radical. S. Gay s'est orienté sur la représentation de l'espace et

2. <http://liris.cnrs.fr/ideal/mooc/>

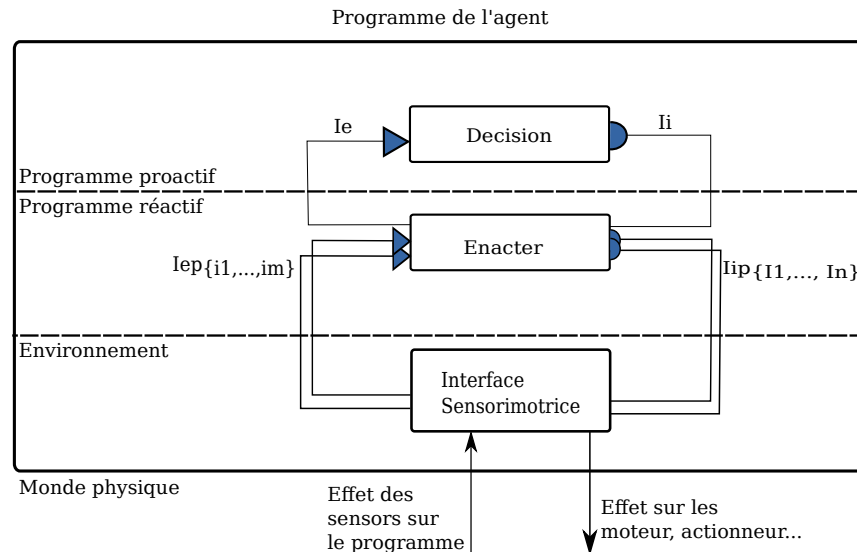


FIGURE 5. Schema de l'implémentation d'un agent

des objets que l'environnement permet d'afforder [Gib77]. « Une affordance est définie comme une possibilité d'interaction proposée par l'environnement à un agent. En effet, dans ces expériences, chaque objet révèle qu'il peut être saisi par un certain mouvement, et, de ce fait, afforde ce mouvement. ». Durant sa thèse, S. Gay a plus travaillé sur la représentation de l'espace dans un agent implémentant l'apprentissage développemental et la perception d'objet dans l'espace observable et de l'espace global.

Travail de Simon Gay. Source d'inspiration.

3 Énoncé du problème

Le problème est défini en se mettant à la place de l'agent. Nous ne connaissons rien de l'environnement et nous devons créer une structure permettant de définir les différents *objets* du monde. Les interactions de la figure n°6 représentent les perceptions et les actions de l'agent. À partir du flux d'interactions de la figure n°7, nous souhaitons que l'agent découvre des régularités qu'il puisse exploiter.

À partir du flux d'interactions de la figure n°7 nous pouvons trouver :

- les régularités immédiates ;
- des régularités séquentielles ;

Les régularités immédiates sont données grâce aux lignes n°2 et 3. Au début du flux d'interaction, l'agent intended des interactions sans savoir si elles sont réalisables. Régulièrement, l'interaction enacted ne correspond pas, ce qui implique que cette interaction est une interaction alternative. Au pas 1, l'interaction

4. <http://liris.cnrs.fr/ideal/mooc/lesson.php?n=058>

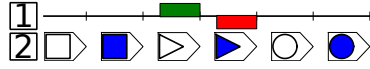


FIGURE 6. Liste des interactions que possède l'agent dans l'environnement inconnu. La ligne n°1 représente la valence qui caractérise la motivation interactionnelle. En vert, la valence est positive, l'interaction est plaisante. En rouge l'interaction est déplaisante, la valence est négative. L'absence de symbole indique que l'interaction n'apporte rien. La ligne n°2 montre les interactions sous forme de symbole.

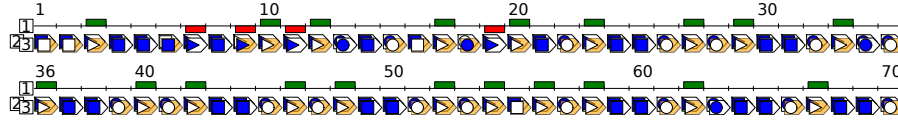


FIGURE 7. Flux d'interactions réalisé par l'algorithme d'O. Georgeon disponible sur le MOOC⁴. La ligne n°2 est l'interaction intended et la ligne n°3 est l'enacted. Les interactions en jaune ont été mal prévues par l'agent qui est alors insatisfait. L'agent a bien prédit les interactions énoncées blanche par conséquent son humeur est satisfaite.

{carré bleu} a pour alternative {carré blanc}. Au pas 6, les interactions {carré bleu} et {carré blanc} deviennent des interactions opposées l'une à l'autre.

Les régularités séquentielles de niveau 2 que l'on peut trouver via ce flux sont représentées par la figure n°8. Ces régularités sont définies par l'environnement, et nous posons l'hypothèse qu'avec ces régularités, l'agent peut construire des connaissances sur l'environnement. La régularité d'interactions {carré blanc}, {carré blanc} informe sur la présence d'une observation. Cette interaction est supposée répétable indéfiniment. Elle est considérée comme étant une interaction persistante qui renseigne sur la présence ou l'absence d'un objet. Les interactions {rond bleu}, {triangle blanc} et {triangle bleu} ne se répètent pas de manière systématique, on considère que ce sont des interactions qui modifient l'environnement et sont appelées interactions sporadiques. L'interaction {rond blanc} ne donne pas d'informations particulières sur son état, par contre elle renseigne des interactions qui peuvent être effectuées juste après elle.

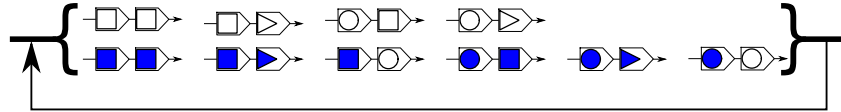


FIGURE 8. Régularités observées dans le flux d'interaction

L'agent est motivé par la réalisation de l'interaction {triangle blanc}, de ce fait, nous espérons pouvoir construire les régularités séquentielles de la figure n°9. Ces régularités sont déjà apprises par l'algorithme via les interactions composites

mais ces interactions ne permettent pas d'avoir une représentation des *objets* de l'environnement.

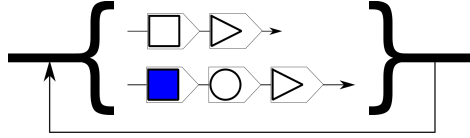


FIGURE 9. Régularités séquentielles que l'on souhaiterait que l'agent trouve et utilise pour satisfaire sa motivation

Les interactions sporadiques {rond blanc} et {rond bleu} permettent une fois énoncées de réaliser respectivement les interactions persistantes {carré blanc} et {carré bleu}. Nous appelons état de croyance l'information qui permet à une interaction sporadique de prédire une interaction persistante. L'interaction {carré blanc} permet d'enact l'interaction {triangle blanc}. L'interaction {rond blanc} permet de changer l'environnement de sorte que l'interaction persistante {carré blanc} soit toujours enactable. De ce fait, l'interaction {rond blanc} possède comme état de croyance l'interaction persistante {carré blanc}.

Avec ces informations et en considérant que l'interaction {rond blanc} est par défaut une interaction persistante, nous pouvons créer le graphe de Pétri de la figure n°10 qui représente les liens entre les interactions persistantes et sporadiques.

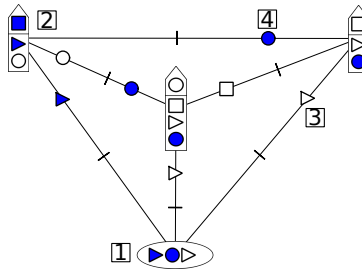


FIGURE 10. Graphe de Pétri que nous souhaitons que l'agent mette en place. L'élément n°1 représente l'état de croyance *unknown* qui regroupe les interactions sporadiques. Les éléments sous la forme du n°2 représentent les interactions persistantes. La partie inférieure du bloc représente les interactions qui pourront être réalisées lorsque l'agent se trouve dans cet état de croyance. Lorsque l'agent est dans l'état de croyance <carré blanc> et qu'il réalise l'interaction {triangle blanc} alors l'état de croyance deviendra *unknown*. Ce lien est représenté par le n°3. Le n°4 représente un état de croyance associé à l'interaction sporadique {rond bleu}.

L'environnement décrit avec les *yeux* de l'agent est le *string problem* [GH13]. Cet environnement est une chaîne de nombres : 1-7-3-2-9-3-5-6-7-8-1-2-4-0-9-8-5-4-6-0, que l'agent doit trier. L'agent est positionné sur un nombre et peut :

- avancer vers un nombre plus grand (resp. plus petit) avec l'interaction {carré blanc} (resp. {carré bleu});
- échanger le nombre sur lequel il se situe avec celui devant lui. Si après l'échange le nombre devant est plus grand (resp. plus petit) alors l'interaction {rond blanc} (resp. {rond bleu}) est enact;
- avancer sur un nombre plus grand (resp. plus petit) via l'interaction {triangle blanc} (resp. {triangle bleu}).

L'agent est alors motivé à trier la chaîne de caractères pour, par la suite, réaliser l'interaction {triangle blanc}. Cet environnement est volontairement simple pour rendre accessible l'analyse de la mémoire de l'agent.

De ce constat, je propose une structure permettant à l'agent d'obtenir une représentation des objets à travers ses interactions primitives.

4 Les phénomènes et les tables d'usage

4.1 Tables d'usages

Les tables d'usage sont inspirées du travail que S. Gay a réalisé lors de sa thèse et qu'il appelle *signature*.

Les types d'interaction sont :

- P : interaction persistante;
- S : interaction sporadique;
- S B : interaction sporadique avec un état de croyance.

Le n°6 représente les pré-interactions. Le n°7 liste les interactions alternative et le n°8 sont les post-interactions.

On peut déduire de la table d'usage de la figure n°11 que cette interaction n'est pas répétable et que l'interaction {carré blanc} n'est pas réalisable avant l'interaction {rond blanc}.

Les interactions {carré bleu} et {rond bleu} sont des interactions qui permettent d'énacter l'interaction {rond blanc}. Ces interactions sont les conditions d'énaction de l'interaction {rond blanc}.

De même que l'interaction {rond blanc} permet d'énacter {carré blanc}, {rond bleu} et {triangle blanc}. On parle alors d'interactions confirmées par l'interaction {rond blanc}.

L'interaction {rond bleu} est une interaction opposée et qui a été énectée trois fois à la place de l'interaction {rond blanc}. Avec cette interaction nous pouvons construire le sous graphe de Pétri de la figure n°?? qui reprend les points ci-dessus.

Les tables d'usage sont mise à jour par lors de la phase d'action et sont utilisées lors de la phase de décision pour le choix de l'interaction suivante.

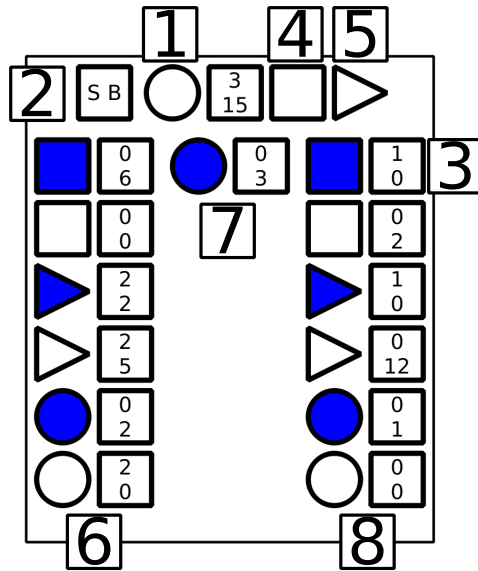


FIGURE 11. Table d'usage de l'interaction {rond bleu} représenté par le n°1. Le n°2 est le type de l'interaction, dans cette figure *sporadique avec croyance*. Les nombres dans les n°3 représentent le nombre de fois que l'interaction a été intended (en haut) et enacted. Le n°4 est l'état de croyance de l'interaction. Le n°5 est la prochaine interaction qui sera réalisée si l'agent énonce cette interaction. La colonne n°6 sont les pré-interaction, la n°7 les interactions opposées et la n°8 les post-interactions

Mise à jour des tables d'usages Lorsque l'*enacter* à réaliser l'interaction intended et a recréé l'interaction enacted, il met à jour les tables d'usages des deux interactions à travers la structure appelé *Phénomènes*.

La mise à jour est effectuée en utilisant le principe suivant : L'interaction enacted doit mettre à jour les poids du succès ou de l'échec des pré-interactions intended et enacted. L'interaction précédemment enacted met à jour les poids de succès ou d'échec des post-interactions intended et enacted. Si l'interaction intended est différent de l'interaction enacted alors l'interaction intended doit mettre à jour le poids d'échec de la précédente interaction enacted. Si l'interaction intended est égale à la précédente interaction intended mais que les interactions enacted sont différentes alors l'interaction intended est considérée comme sporadique.

Table d'usage pour les interactions composites Pour le moment, les interactions composites sont gérées de la même manière que les interactions primitives mais n'apporte pas d'apport dans la représentation des objets de l'environnement. Par conséquence, elles ne seront pas détaillées dans ce mémoire.

4.2 Phénomènes

Nota bene : dans le domaine de la phénoménologie, les objets du monde se représente non pas par l'objet en lui-même mais par les interactions que l'individu peut effectuer sur l'objet. Les composantes du monde, sont inaccessible à l'individu, mais celui-ci arrive à distinguer un objet rond et carré facilement. Le faite de voire un objet, active dans le cerveau une sorte de simulation des interactions que nous pouvons d'effectuer avec l'objet. Cette représentation nous donne la possibilité de comprendre que deux objets sont identiques dans des situations différentes [MFF⁺97].

Pour permettre la gestion des phénomènes, nous avons implémenté 2 algorithmes. Le premier à deux phases. L'une pour l'apprentissage des phénomènes et la deuxième pour l'exploitation et la mise à jour des connaissances. Les interactions sont initialisée comme étant inconnue. L'agent va répéter, jusqu'à un certain seuil, une interaction considérée comme inconnue, pour déterminer si elle est persistante ou sporadique. Le flux d'interactions généré par cet algorithme dans le *string problem* est représenté par la figure n°12.

Le second algorithme part du principe que l'apprentissage peut être réalisé en même temps que l'exploitation. L'agent va apprendre de ces erreurs et essayer de maximiser sa motivation interactionnelle. Les interactions sont initialisées comme étant persistantes et l'agent va alors apprendre au cours de ses expériences les interactions sporadique. La figure n°13 relate l'expérience de l'agent avec cet algorithme dans l'environnement du *string problem*. La figure n°14 montre l'évolution des états de croyance de l'agent.

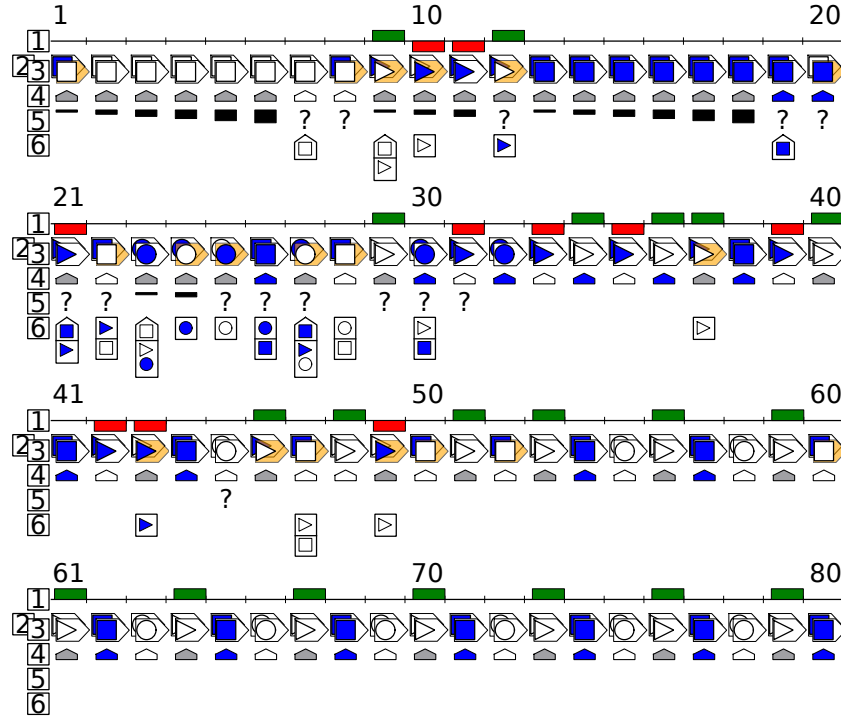


FIGURE 12. Flux d'interaction de l'algorithme d'apprentissage et d'exploitation. La ligne n°4 représente l'état de croyance dans lequel se trouve l'interaction énoncée. En gris, l'état de croyance est inconnue, en blanche elle fait référence à l'état de croyance <carré blanc> et en bleu à l'état de croyance <carré bleu>. La ligne n°6 montre la création et la mise à jour des interactions persistantes et sporadique. L'humeur de l'agent est sur la ligne n°4. Du pas 1 à 6, l'excitation de l'agent augmente jusqu'au seuil (prédéfini). Au pas 7, l'interaction {carré blanc} est considérée comme persistante. Au pas 8, l'agent est curieux car il connaît l'interaction persistante {carré blanc}, mais sa table d'usage est vide. Il va alors essayer les expériences pour pouvoir remplir la table d'usage. Au pas 9, l'agent est excité par l'énoncé de l'interaction {triangle blanc}, mais au pas 10, cette interaction n'est pas répétée. Elle est alors considérée comme étant une interaction sporadique. De même pour l'interaction {triangle bleu} du pas 10 au 12. Au pas 18, l'agent a créé les deux phénomènes attendu à savoir <carré blanc> et <carré bleu>. Mais ces phénomènes sont incomplets. Avec les différents tests, l'agent remplira correctement la table d'usage des phénomène <carré blanc> et <carré bleu> au pas 27. L'exploitation commence, avec des erreurs au pas 32. À partir de ce moment, l'agent se sert de ses croyances pour avancer, mais celui-ci se trompe encore car les tables d'usages ne sont pas correctement remplies. Ce n'est qu'à partir du pas 49 que les différents liens sont correctement établie et que l'exploitation peut se dérouler sans erreur.

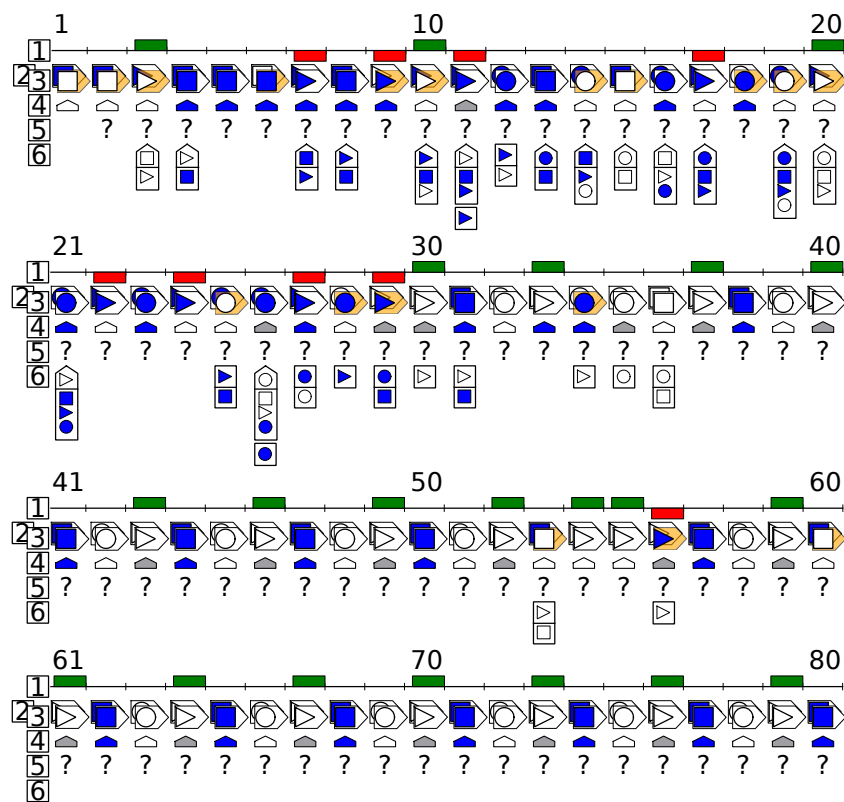


FIGURE 13. 80 premières interactions avec l'algorithme d'apprentissage et d'exploitation en simultané.

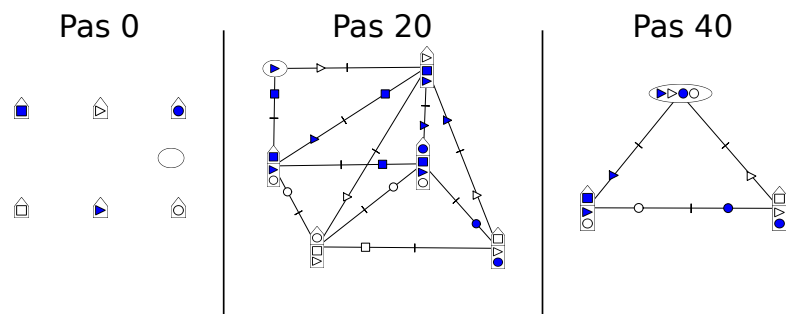


FIGURE 14. Graphe de Pétri trouvé par l'agent après 40 interactions.

5 Environnement de travail

Le logiciel développé durant le stage a été réalisé dans l'optique d'être réutilisé par la suite. Il est composé de 4 bibliothèques et du programme principale qui lie les bibliothèques entre elles. La première bibliothèque est l'implémentation des agents et des mécanismes d'apprentissages et d'exploitations des interactions. La seconde permet de réaliser une simulation d'un agent dans un environnement donné. La troisième permet l'affichage graphique des interactions et des tables d'usages. La dernière sert à lier les modifications de l'agent lors d'une simulation à un serveur KTBS⁵. Le programme principal représenté par la figure n°15 montre un agent utilisant l'algorithme d'apprentissage et d'exploitation simultanée dans le *string problem* au centre.

La partie inférieure du programme montre une partie de la bibliothèque graphique d'interaction.

À droite, la configuration de la simulation et de l'affichage des logs de l'agent. À gauche les logs de l'agent. On peut y voir les interactions activées, proposées, intended, enacted, appris et renforcé et les humeurs de l'agent. En haut, permet de sélectionner l'environnement avec son fichier de configuration. À droite, le fichier de configuration du graphisme.

5. KTBS : Kernel for Trace Base System dont les concepts sont disponibles à l'adresse : <https://kernel-for-trace-based-systems.readthedocs.org/en/latest/concepts/overview.html>

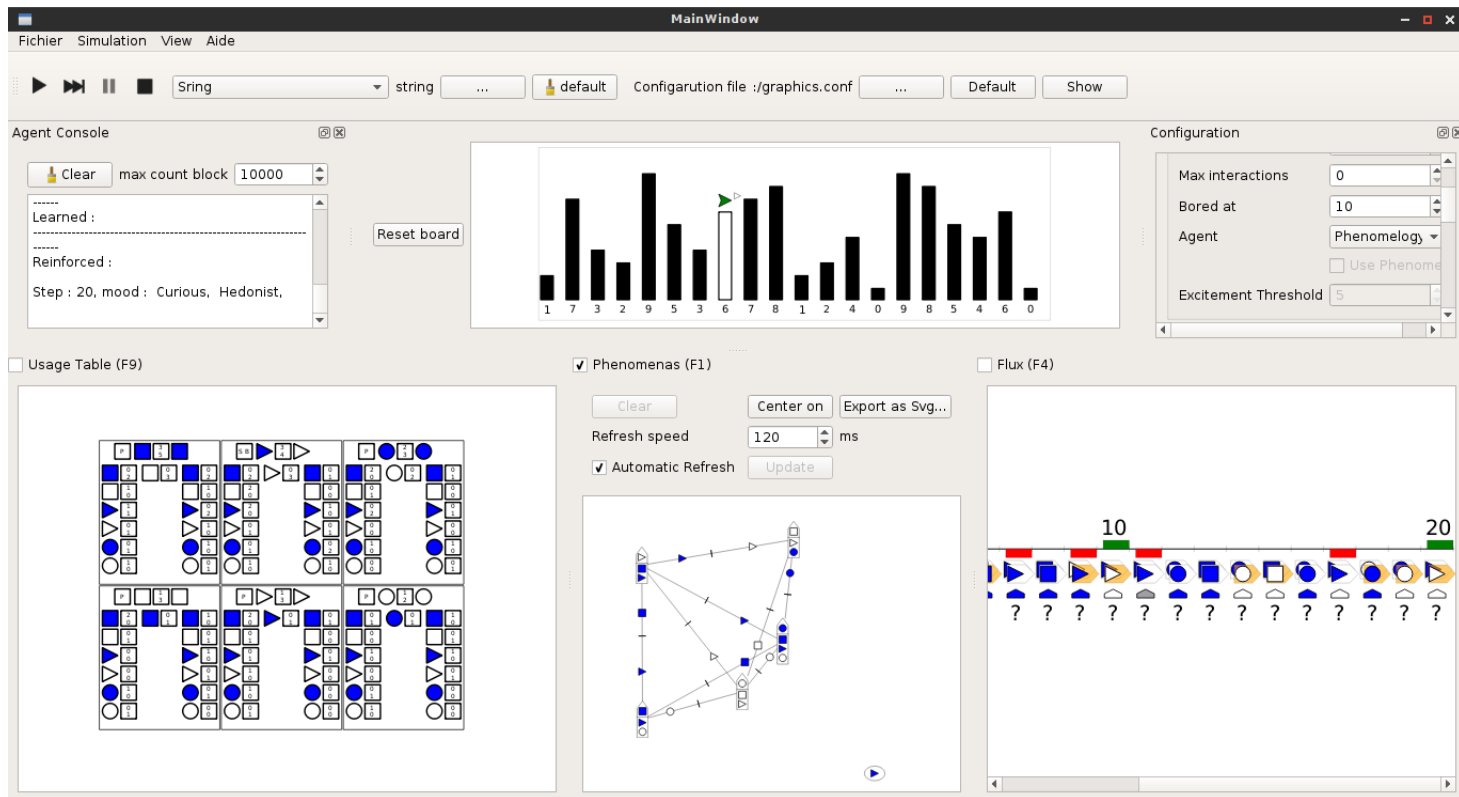


FIGURE 15. Programme principal permettant de simuler différents agents dans différents environnement

6 Travail à venir

Améliorer les usages table des interactions composites. Trouver l'émergence de la construction de l'espace à travers les interactions et les tables d'usage. Faire dans des environnements complexe et avec un agent possédant beaucoup d'interaction.

Une motivation non testée, mais intéressante, serait de mettre en place la possibilité à l'agent de se concentrer sur une interaction afin de déterminer si celle-ci est optimale ou non, lui permettant d'atteindre ses différents objectifs de manière plus efficace. + simulation

7 Conclusion

L'implémentation d'un agent agnostique utilisant les préceptes de l'apprentissage développemental n'est pas chose facile. L'avancée que j'ai réalisé me semble insignifiante et pourtant ce n'est pas chose facile que d'aller plus loin. L'implémentation de nouveaux algorithmes d'apprentissage et d'exploitation d'interactions

Références

- And03. Michael L Anderson. Embodied cognition : A field guide. *Artificial intelligence*, 149(1) :91–130, 2003.
- Bro91. Rodney A Brooks. Intelligence without representation. *Artificial intelligence*, 47(1) :139–159, 1991.
- GA13. Olivier Georgeon and David Aha. The Radical Interactionism Conceptual Commitment. *Journal of Artificial General Intelligence*, 4(2) :31–36, December 2013.
- Gay14. Simon Gay. *Mécanismes d'apprentissage développemental et intrinsèquement motivés en intelligence artificielle : étude des mécanismes d'intégration de l'espace environnemental*. Thèse de doctorat en informatique, Université Lyon 1, December 2014.
- GC14. Olivier L Georgeon and Amélie Cordier. Inverting the interaction cycle to model embodied agents. *Procedia Computer Science*, 41 :243–248, 2014.
- GH13. Olivier Georgeon and Salima Hassas. Single agents can be constructivist too. *Constructivist Foundations*, 9(1) :40–42, November 2013.
- Gib77. JJ Gibson. *The theory of affordances*, In” *Perceiving, Acting and Knowing*”, Eds. RE Shaw and J. Bransford. Erlbaum, 1977.
- GMG12. Olivier Georgeon, James Marshall, and Simon Gay. Interactional Motivation in Artificial Systems : Between Extrinsic and Intrinsic Motivation. In *International Conference on Development and Learning (ICDL-Epirob)*, November 2012. Article court, communication affichée.
- GS12. Olivier Georgeon and Ilias Sakellariou. Designing Environment-Agnostic Agents. In Peter Vrancx Enda Howley and Matt Knudson, editors, *ALA2012, Adaptive Learning Agents workshop, at AAMAS2012, 11th International Conference on Autonomous Agents and Multiagent Systems*, pages 25–32, June 2012.
- Hus76. Edmund Husserl. *La crise des sciences européennes et la phénoménologie transcendantale*, page 262. Gallimard, 1976.
- MFF⁺97. Akira Murata, Luciano Fadiga, Leonardo Fogassi, Vittorio Gallese, Vassilis Raos, and Giacomo Rizzolatti. Object representation in the ventral premotor cortex (area f5) of the monkey. *Journal of neurophysiology*, 78(4) :2226–2230, 1997.
- O’R11. J.K. O’Regan. *Why Red Doesn’t Sound Like a Bell : Understanding the Feel of Consciousness*. Oxford University Press, USA, 2011.
- Pia59. J. Piaget. *The Construction of Reality in the Child*. The Basic classics in psychology. Basic Books, 1959.
- Ste04. Luc Steels. The autotelic principle. In *Embodied Artificial Intelligence*, pages 231–242. Springer, 2004.
- VT. F Varela and E Thompson. et rosch e.(1991). the embodied mind : cognitive science and human experience.
- WF13. D.B. Webster and R.R. Fay. *The Mammalian Auditory Pathway : Neuroanatomy*. Springer Handbook of Auditory Research. Springer New York, 2013.