



Judgment of Ethics of Behavior for Learning

Jérémy Duval

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

June 19, 2020

Jugement de l'éthique de comportement pour l'apprentissage

Mémoire de Master

Jérémy Duval

Univ. Lyon, Université Lyon 1, LIRIS, UMR5205, F-69622, LYON, France

Encadré par Salima Hassas (LIRIS), Olivier Boissier (LabHC, Mines Saint-Etienne) et Remy Chaput (LIRIS)

Résumé : Les systèmes intelligents sont aujourd'hui partout et nous leur déléguons de plus en plus de tâches qui peuvent remettre en cause des valeurs humaines. Il est important que ces systèmes puissent gagner la confiance des utilisateurs et que l'on puisse cadrer leur comportement. C'est pour cela qu'apparaissent aujourd'hui des systèmes développant des comportements visant à respecter des dimensions éthiques. Plusieurs modèles et projets existent actuellement ; que ce soit avec une approche numérique ou une approche symbolique. Ces deux approches ont chacune leurs avantages et leurs inconvénients, néanmoins complémentaires. C'est dans l'optique d'une approche hybride que nous présentons un modèle de jugement de l'éthique de comportements appris automatiquement. Ce travail s'appuie sur les projets Ethics.ai et Ethicaa pour proposer l'intégration d'un jugement symbolique dans de l'apprentissage par renforcement.

Mot Clefs: Intelligence Artificielle, Éthique, Systèmes Multi-Agents, Apprentissage Par Renforcement, IA symbolique, Grilles électriques intelligentes, Jugement éthique.

Abstract. Today, intelligent systems are everywhere, with always more and more tasks. These systems have to gain the users trust and it is important to set boundaries to them. With this in mind many projects are created today with the aim to develop ethical comportments in intelligent systems. At present, there exist some models and projects about ethical comportments, with top-down approach, but also bottom-up approach. Nevertheless, these have advantages and inconveniences, however complementary. From this perspective of doing, we propose an hybrid model of ethical judgment of comportment for learning, based on the Ethics.ai and Ethicaa projects, which integrate symbolic judgment in reinforcement learning.

Keywords: Artificial Intelligence, Ethics, Multi-Agent System, Reinforcement Learning, Symbolic AI, Smart-grid, ethical judgment.

1 Introduction

Aujourd'hui les systèmes intelligents sont partout, que ce soit dans notre voiture, votre téléphone, ou peut-être même dans votre réfrigérateur. Ces systèmes sont aussi de plus en plus autonomes, nous leur déléguons de plus en plus de tâches. Mais pouvons nous leur faire confiance ? Et quid des situations indésirables ? [2] Prenons l'exemple des *smart grids*. Imaginons un ensemble de bâtiments de différents type organisés autour d'un réseau électrique. Ce réseau possède une quantité d'énergie limitée, à répartir entre tous ces bâtiments. Imaginons que notre habitation fasse partie de cette *smart grid* ; nous aimerions probablement éviter de sacrifier notre confort au profit des autres. Cependant il ne serait pas éthique de surconsommer et de, par exemple, entraîner une panne de courant dans l'hôpital relié à ce réseau et de l'empêcher de fonctionner correctement. Si les divers appareils présents dans l'hôpital cessaient de fonctionner, cela pourrait coûter la vie à de nombreuses personnes. Il serait injuste de mettre en danger des vies humaines pour son propre confort.

Avec cet exemple, nous voyons bien la nécessité de cadrer le comportement des systèmes intelligents. Nous voyons aussi, avec le cas de l'hôpital, des notions d'éthique à prendre en compte. Et c'est ce point qu'il nous semble important de développer dans les système intelligents actuels.

Le travail décrit dans ce rapport s'inscrit dans ce contexte applicatif : une *smart grid* dans laquelle nous cherchons à développer un système intelligent dont le comportement serait éthique. Dans ce mémoire, nous ne ferons pas l'hypothèse que les agents artificiels en eux-mêmes peuvent être considérés comme des agents prenant des décisions éthiques, mais nous utiliserons néanmoins le raccourci "comportement éthique" pour décrire le comportement d'un agent qu'un humain jugerait comme étant éthique selon ses valeurs ou ses principes.

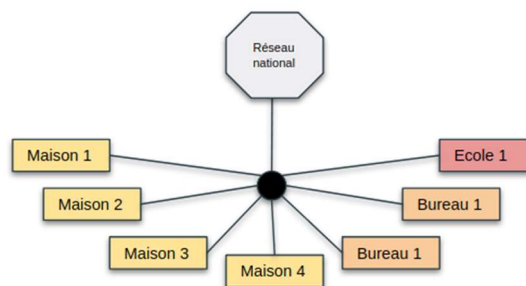


Figure 1: Schéma d'une smart grid

Ce système est composé d'agent organisés en un réseau électrique. Ce réseau est lui-même relié au réseau national (cf Figure 1). Chaque agent représente un bâtiment, associé à une batterie, un système de production d'énergie renouvelable et est relié au réseau national (cf **Erreur ! Source du renvoi introuvable.**). Chaque agent a sa propre consommation d'électricité et peut échanger avec la *smart grid* et le réseau national.

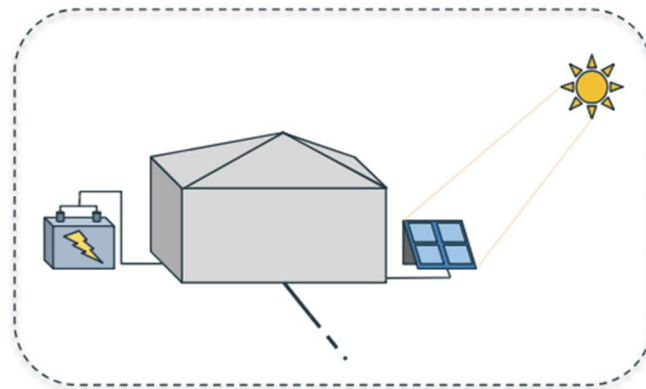


Figure 2 : Schéma d'un agent de la smart grid

Pour simuler cela, nous utiliserons Ethics.ai [9], projet et plateforme d'apprentissage de comportements éthiques dans le cadre de *smart grids*, dont nous détaillerons le fonctionnement plus tard (cf section 2.2). Ce système présente l'avantage de s'adapter facilement à de nouveaux environnements et de nouvelles règles grâce à son approche *bottom-up* (i.e. d'apprentissage). Cependant, une telle approche présente aussi certains désavantages : les raisonnements sont peu compréhensibles par l'humain, il faut un temps d'apprentissage et il est difficile d'inclure des connaissances expertes pouvant être nécessaires, par exemple pour cadrer le comportement des agents et du système.

C'est dans le but de pallier à cela qu'intervient le projet et système Ethicaa [3] [13]. Ethicaa est un modèle d'agent pour le raisonnement sur l'éthique de comportements mis en oeuvre au sein d'un système multi-agent (approche *top-down*). Il est capable de juger le comportement d'un autre en se créant une image de celui-ci au regard de règles éthiques. Ethicaa, de par son côté symbolique, offre la possibilité de correctement cadrer le comportement du système [16]. Le raisonnement des agents est aussi plus compréhensible par l'humain. Néanmoins son adaptation à de nouveaux contextes dont les règles n'ont pas été définis est difficile -problème pouvant être pallié par Ethics.ai- et les échelles de valeurs, discrètes, sont plus complexes à établir que des échelles continues que l'on trouve dans les approches numériques.

Comme nous venons de le voir, Ethics.ai et Ethicaa servent de base pour mon stage. Il est important de noter que ces deux systèmes sont des systèmes multi-agents (SMA). Peu de travaux mêlant éthique et intelligence artificielle concernent les SMA [12] [21]. Leur présence aujourd'hui est pourtant notable ; que ce soit dans la robotique, dans les systèmes distribués en général ou encore bien d'autres domaines. Il serait intéressant, comme pour tout système intelligent d'y inclure de l'éthique. Néanmoins, ces systèmes étant multi-agent, il serait bien plus intéressant de penser l'éthique de manière collective, de par la prise en compte d'autrui, les notions de confiance, de coopération [12], ou de par l'argumentaire que l'on peut mettre en place et les possibilités d'évolution de ses propres valeurs que cela peut amener. Tout ceci a donc fortement motivé mon travail.

De plus, la capacité de jugement symbolique d'Ethicaa présente un réel intérêt, d'autant plus dans le contexte d'apprentissage automatique d'Ethics.ai. Cela ouvre des perspectives afin de cadrer le comportement des agents apprenants -point faible d'Ethics.ai-, de les aiguiller au mieux dans leur apprentissage et ainsi réduire le temps d'apprentissage.

C'est de ces contextes et de ces motivations que j'ai pu élaborer la question de recherche suivante :

Q1 : Comment juger et influencer des comportements éthiques nés d'apprentissage automatique ?

Cette question -qui guidera mon travail tout au long de mon stage- en induit deux nouvelles :

Q2 : Comment juger une action réalisée par un agent ou un groupe d'agent dans un SMA ?

Q3 : Comment attribuer des récompenses en fonction de jugements symboliques ?

Afin de répondre à ces questions, nous avons commencé par établir un état de l'art sur les différents points concernés. Nous avons pris en main les projets Ethics.ai et Ethicaa, sur lesquels se base notre travail. Nous avons étudié l'apprentissage par renforcement mis en place dans Ethics.ai et les différentes variables d'observation du comportement des agents afin de comprendre comment juger eux-ci. Nous avons ensuite pris en main Ethicaa afin d'en comprendre les mécanismes de jugement symboliques. Nous avons dès lors pensé un modèle permettant la liaison entre Ethics.ai et Ethicaa, entre apprentissage automatique et jugement symbolique. Nous avons ensuite implémenté ce modèle et fait différentes expérimentations.

Nos contributions sont : la proposition d'un modèle efficace de jugement éthique de comportement pour l'apprentissage automatique, l'élaboration d'un dashboard permettant l'observation en temps réel du comportement des agents et l'augmentation d'Ethics.ai par l'implémentation de propriétés individuelles d'observation des agents.

Dans la suite de ce mémoire, nous présenterons tout d'abord un état de l'art sur l'éthique dans l'IA et les SMA, les approches d'apprentissage numériques et symbolique, ainsi que les approches hybrides. Nous présenterons ensuite le modèle que nous avons élaborés et l'implémentation actuelle qui en est faite. Nous présenterons les expérimentations qui en ont été faites. Enfin, nous discuterons des résultats, des limites et des perspectives de notre travail.

2 État De l'Art

Dans cette section nous passons en revue des travaux lié à l'éthique en IA (2.1), d'abord de manière générale, puis plus particulièrement dans les systèmes multi-agents. Nous verrons ensuite les approches numériques (2.2) de l'éthique dans l'IA et nous illustrerons cette section par le modèle issu du projet Ethics.ai. Nous parlerons ensuite des approches symbolique (2.3), nous appuyant pour cela sur le modèle Ethicaa. Enfin, nous parlerons de la liaison numérique-symbolique (2.4).

2.1 L'Éthique dans l'IA

Comme nous l'avons mentionné en introduction, les systèmes intelligents se développent de plus en plus et nous leur laissons de plus en plus d'autonomie. Comme l'expose [2], des systèmes enfreignent déjà des règles éthiques, tels que les moteurs de recherche qui collectent les données privées des utilisateurs. Certaines questions apparaissent, par exemple avec les véhicules autonomes [16] et les dilemmes éthiques, tel que le *trolley problem* [2]. Il est donc important de développer l'éthique dans les systèmes intelligents [2] [16]. Cela implique aussi d'intégrer des notions d'explicabilité dans ces systèmes, afin qu'ils soient en mesure de justifier et d'argumenter leurs choix, de manière compréhensible pour l'humain [2] [16]. Comme exposé dans [1], différentes approches peuvent se faire : sous forme d'apprentissage (*bottom-up*), sous forme symbolique avec des systèmes de raisonnement (*top-down*), ou en combinant les deux approches (*hybrid*).

L'exploration des dilemmes consiste à essayer de résoudre des situations où aucune solution au problème n'entraînera pas de regret. C'est-à-dire que chaque solution est, soit indésirable, soit entraîne des conséquences indésirables. Il n'y a donc pas de solutions idéales. On retrouve un certain nombre de travaux à ce sujet, tel que [7].

La prise de décision éthique individuelle consiste en le fait qu'un agent choisisse la meilleure action à adopter dans une situation donnée, au regard de règles éthiques. Cela s'applique aux cas où l'agent est seul ou lorsque l'agent fait partie d'un groupe mais ne prend pas celui-ci en compte dans ses choix. On trouve notamment [5] à ce sujet.

La prise de décision éthique collective est une forme de prise de décision éthique où l'agent se trouve au sein d'un système multi-agent, et, où il va prendre en compte les autres agents. Ce peut être par des notions de confiance, de jugement, etc. Cette forme de prise de décision éthique peut aussi concerner des décisions éthiques prises par un groupe d'agent se coordonnant. On retrouve par exemple [18] à ce sujet.

L'éthique dans les décisions humains-machines, consiste en le fait d'influencer le comportement d'utilisateur humain par le biais d'agent éthique. Cela peut aussi concerner la coopération humain-agent lors de prise de décision. On trouve aussi des travaux tel que [4] visant à permettre à des agents d'apprendre des comportements éthiques en interagissant avec des humains.

Comme vu précédemment, nous nous intéressons ici plus particulièrement à l'éthique collective. D'après [21] on constate que peu de travaux ont abordé la question de l'éthique au sein des systèmes multi-agent, néanmoins, depuis la parution de cet article en 2018, de plus en plus de travaux s'intéressent à la question. On retrouve par exemple [19] présenté lors de la conférence AAMAS en mai 2020. Selon [21], on trouve des approches basées sur l'établissement de règles sociale [20] ou sur le principe de vote [18] pour prendre des décisions. On trouve aussi des approches basées sur la confiance et la coopération entre agents [13].

On retrouve aussi des approches numériques, telle que celle suivie dans le projet Ethics.ai [9], ou symboliques, telle que dans le projet Ethicaa [3] [13], que nous allons détailler maintenant.

2.2 Approche numérique

L'approche numérique, ou approche *bottom-up* consiste, selon [1], au développement de l'agent selon des mécanismes de récompenses -qui peuvent être intrinsèques ou extrinsèques-, de sélection ou par le biais de jeu de données. Ces mécanismes s'apparentent au fait d'être récompensé ou puni pour ses actions. L'agent, cherchant à maximiser les récompenses obtenues pourra ainsi être guidé dans son apprentissage. On retrouve un large panel de méthodes d'apprentissage numériques, telles que l'apprentissage par renforcement, l'apprentissage développemental, l'apprentissage profond, etc. Cette approche possède, comme principal avantage, l'amélioration du système initialement mis en place, et, dans le cas d'apprentissage en ligne, la capacité de s'adapter aux situations nouvelles, aux cas inconnus [1]. Cependant, il est complexe de cadrer le comportement des agents et il est souvent difficile pour l'humain d'en comprendre le raisonnement.

Ethics.ai [6] s'inscrit dans cette approche. Il s'agit d'un système éthique numérique basé sur l'apprentissage par renforcement au sein de systèmes multi-agents. Ethics.ai utilise différents paramètres pour observer une situation : les actions des agents, leurs besoins, ainsi que des propriétés d'observations : l'autonomie, le bien-être, l'énergie perdue, l'équité, l'exclusion et la surconsommation. Celles-ci correspondent au cas d'application principal du projet : les *smart-grids*.

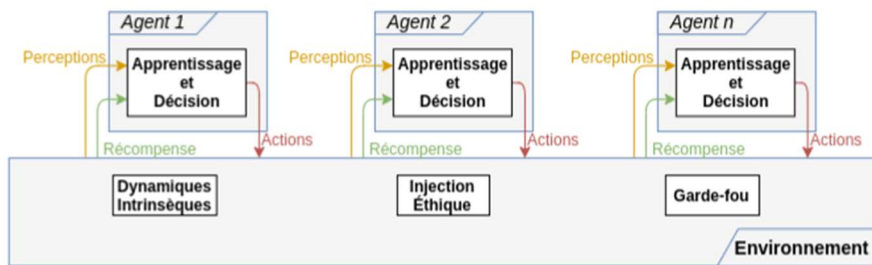


Figure 2: Modèle Ethics.ai d'apprentissage de comportement éthiques dans un contexte multi-agent

Comme on peut le voir dans le schéma ci-dessus -issue de [9]- l'apprentissage de comportement éthiques se fait par le biais de l'environnement, sous forme de récompense.

2.3 Approche symbolique

L'approche symbolique, ou approche *top-down* selon [1] consiste en l'élaboration d'un ensemble de règles éthiques et d'un système de raisonnement sur ces règles. Ces règles sont élaborées par le concepteur du système intelligent et se basent souvent sur un système éthique et moral existant tel que les lois d'Asimov, la déontologie Kantienne, les 10 commandements, etc. Le raisonnement à base de règle, tel que dans cette approche, présente plusieurs avantages : la compréhension du raisonnement des agents par l'humain est facilitée, de même que le cadrage de leur comportement [1]. En effet, ces règles symboliques permettent une meilleure compréhension que des données numériques, grâce aux prédicats et valeurs exprimées de manière intelligible à l'humain. De plus écrire ces règles permet de s'assurer que l'agent agira comme on le souhaite dans les situations ciblées. Néanmoins, l'adaptation à de nouvelles situations est complexe et certaines règles peuvent parfois s'opposer et conduire à des dilemmes éthiques [1].

Imaginons un système composé de trois règles :

R1 : "Il est immoral de mentir"

R2 : "Il est immoral de causer la souffrance directe ou indirect de quelqu'un"

R3 : "Il est moral d'aider quelqu'un dans le besoin"

Plaçons nous dans le contexte de la seconde guerre mondiale. Une personne interrogée cache un résistant. Un soldat demande à cette personne si elle cache chez elle un résistant. Si c'est le cas, le soldat ira perquisitionner le résistant et le torturera pour obtenir des informations. Selon nos deux règles définis ci-dessus, l'intéressé est face à un dilemme éthique. S'il ment, il enfreint R1 mais satisfait R2. S'il dit la vérité, il satisfait R1 mais enfreint R2. Les deux règles étant au même niveau, cette solution est insolvable pour lui.

Il existe cependant certains systèmes de règles hiérarchiques, tel que les lois d'Asimov ; mais même de tels système ne peuvent prévoir toutes éventualité [1]. De surcroît, plus le nombre de règles est élevée, plus la puissance de calcul nécessaire sera importante ; ce qui peut poser problème dans des situation où la prise de décision doit être rapide [1, 3, 13].

Une approche à base de règles ouvre aussi la possibilité d'élaborer un système de jugement d'autrui. Un tel système est utile à l'élaboration d'éthique collective. Cela permet d'introduire des notions de confiance entre agents et de compréhension de l'autre [13].

Reprenons notre exemple précédent. L'interrogé sait que le soldat fait partie du camp ennemi. Il ne lui fait donc pas confiance.

Pendant la seconde guerre mondiale, il y a avait un système de rationnement par habitant. Afin de contrôler cela, chaque habitant recevait régulièrement un certain nombre de ticket pour pouvoir faire ses achats. Reprenons notre personne interrogée. Celle-ci a besoin de plus de nourriture car elle cache un résistant. Cette même personne a surpris plusieurs fois l'épicier de sa rue aider des personnes dans le besoin. Selon R3, notre personne jugera donc l'épicier moral. Elle pourra donc faire confiance à l'épicier et tenter d'obtenir plus de nourriture afin de se nourrir et de nourrir le résistant.

Ethicaa, que nous utiliserons, est un projet répondant bien à ces caractéristiques [3, 13]. Il s'agit d'un système de raisonnement symbolique pour l'éthique collective, élaboré selon le *framework* JaCaMo [6], et basé sur une architecture BDI. Ethicaa ajoute à cette architecture une dimension éthique dans le raisonnement de l'agent. L'agent n'évalue plus seulement la désirabilité et la faisabilité des actions à sa disposition, mais aussi l'acceptabilité de celles-ci au regard de règles morales et de valeurs morales d'une part et de principes éthiques d'autre part. Ethicaa permet aussi aux agents de juger d'autres agents selon ces règles et supports, en se construisant une image du comportement de ceux-ci. L'un des cas d'application d'Ethicaa est la prise de décision éthique dans les échanges d'actions au sein de marchés boursiers. Plusieurs agents sont définis autour de valeurs morales et éthiques (par exemple l'écologie) et interagissent pour l'achat et la vente d'actions concernant des entreprises pouvant être moral, immoral ou neutre au regard de leurs valeurs.

2.4 Approches hybrides

On trouve dans la littérature un certain nombre d'approches hybrides. En effet, comme le souligne [1], l'approche *top-down* apporte un côté explicatif et permet un bon contrôle du comportement de l'agent, tandis que l'approche *bottom-up* permet de l'adaptabilité et l'émergence de nouveaux comportements ; ce sont donc deux approches qui se complètent. On retrouve par exemple, hors du domaine de l'éthique, de l'intégration d'apprentissage par renforcement dans la plateforme SOAR [17], de la représentation et du raisonnement sur des connaissances mélangé avec de l'apprentissage par renforcement [15] ou encore de l'intégration de planification dans de l'apprentissage par renforcement pour de la robotique [14].

Cependant, dans ces approches hybrides, il s'agit toujours de passer du numérique au symbolique, d'ajouter de l'explicatif dans les systèmes numériques, ou encore d'intégrer directement du symbolisme dans le moteur d'apprentissage numérique. Cependant, dans notre cas, il s'agit d'avoir de passer du symbolique (le jugement) au numérique (la récompense et le système d'apprentissage), tout en gardant une séparation distincte entre les deux.

3 Modèle

Dans cette section, nous aborderons le modèle permettant de lier apprentissage numérique et jugement symbolique. C'est aussi ce modèle qui permet la liaison entre Ethicaa et Ethics.ai.

3.1 Vision globale

Comme constaté dans [11], dans une population, les individus, les groupes et les sociétés ont tous leurs propres idées, valeurs, et donc, éthiques ; ainsi que leur propre degré d'influence.

De ce fait, j'ai voulu créer un modèle où seraient présentes ces différentes notions, ces différents degrés. On juge ainsi le comportement d'un agent sur différents degrés : individuellement et collectivement (par groupe et tous ensemble). L'agent peut alors être jugé sur ses propres actions mais est aussi incité à contribuer à la société d'agent entière. De plus les groupes permettent un niveau intermédiaire. En effet, nous pouvons regrouper les agents agent par types ou par proximité d'éthique, etc, et ainsi attribuer un jugement pour le fonctionnement de ce type d'éthique ou d'agents.

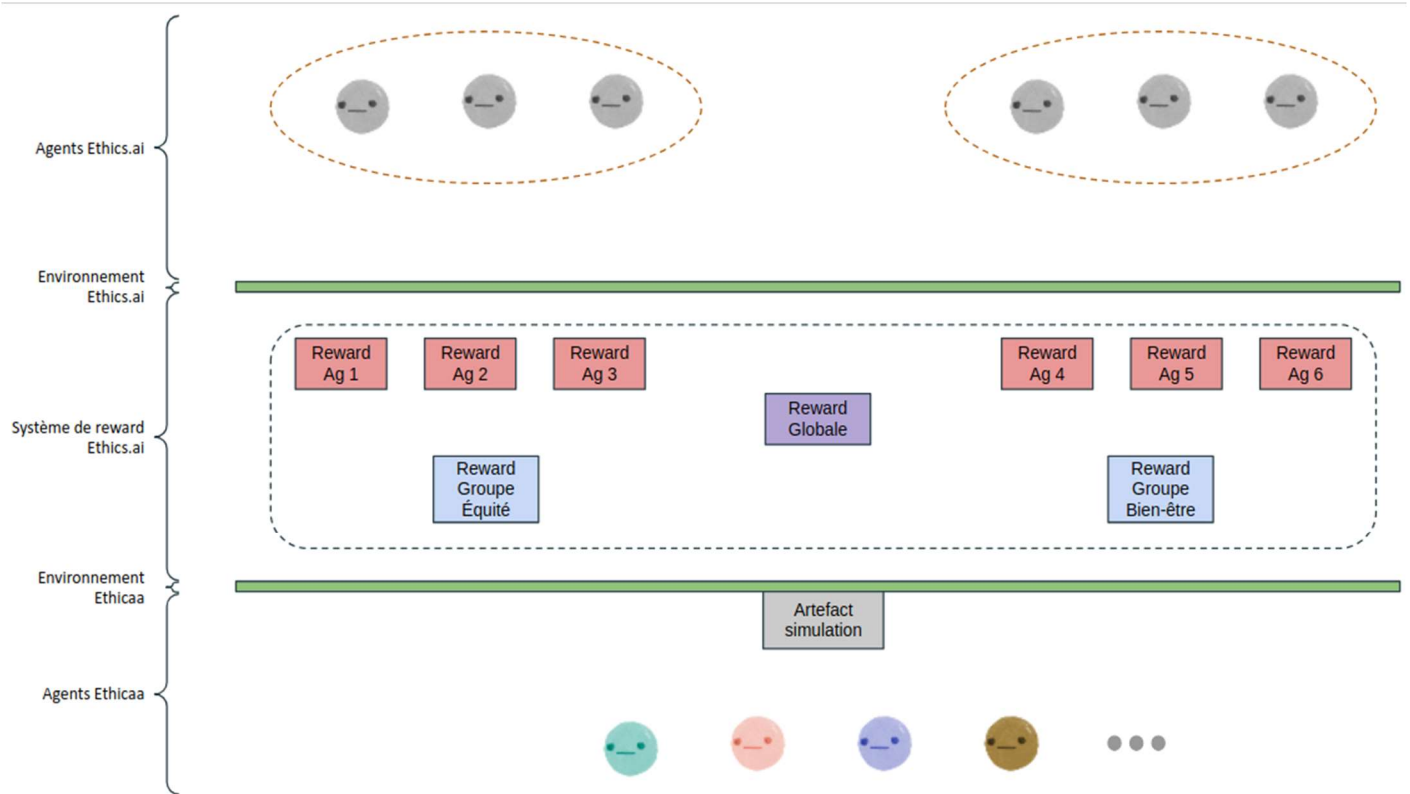


Figure 3: : Schéma du modèle : vue globale

Ci-dessus, l'architecture de base du modèle. On y trouve les agents Ethics.ai en gris. Ceux-ci sont organisés en groupes (pointillés oranges). L'ensemble des agents forme une société. Nous retrouvons aussi les agents Ethicaa (symbolisés par différentes couleurs). Ces agents vont faire office de sage et ont chacun leur propre éthique et leur propre morale. Chaque agent interagit avec l'environnement de son simulateur (Ethics.ai ou Ethicaa). Les agents Ethics.ai reçoivent une récompense selon leurs agissements, attribuée par le simulateur Ethics.ai et prenant en compte les différents jugements des agents Ethicaa. Ces derniers sont informés des agissements des agents Ethics.ai et distribuent leur récompense par le biais du simulateur Ethicaa.

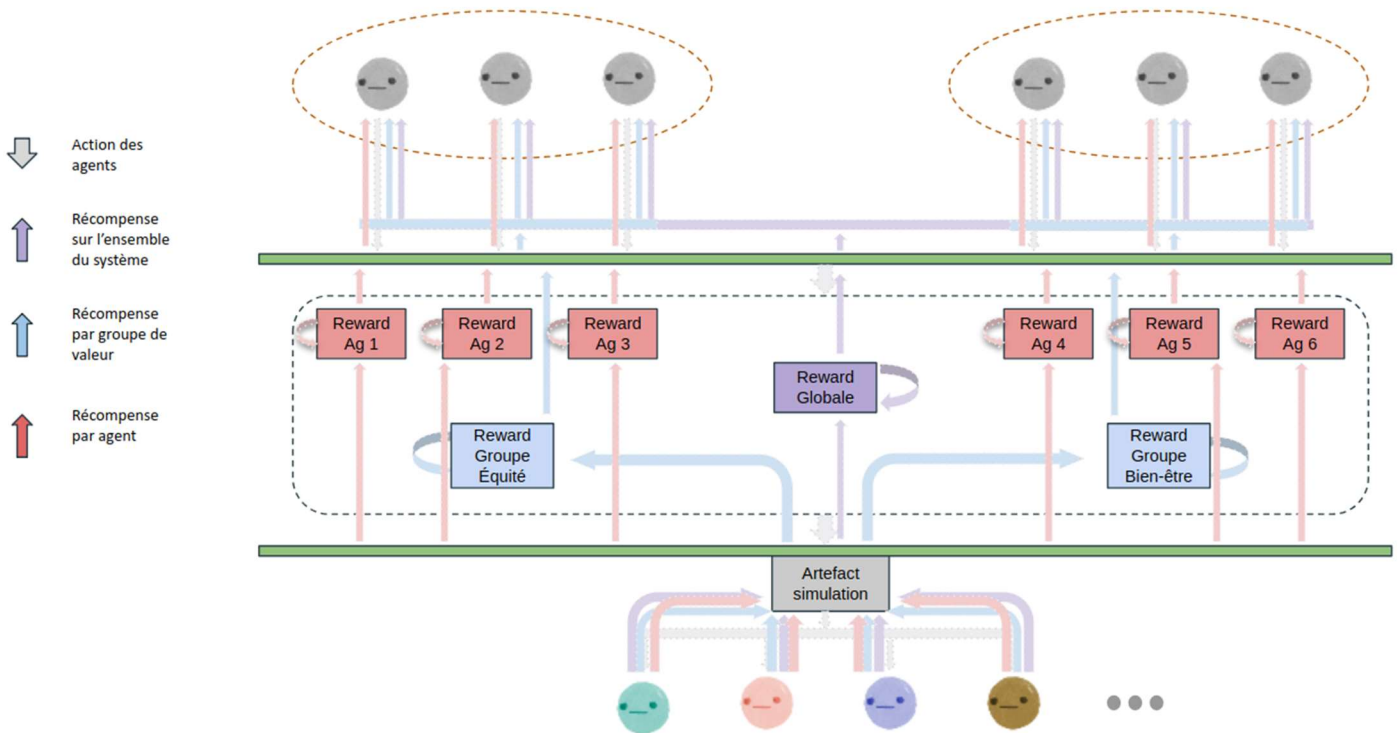


Figure 4: Schéma du modèle : résumé

Ci-dessus, on peut voir les différentes sous-étapes pour une étape de simulation. Dans un premiers temps les agents Ethics.ai agissent sur leur environnement. Ces actions sont transmises aux agents Ethicaa. Ensuite, les agents Ethicaa vont juger ces actions et attribuer une récompense, au regard de leurs règles éthiques, à chaque individu, chaque groupe spécialisé dans l'éthique lui correspondant, et, à la société. Le simulateur Ethics.ai va produire une récompense numérique par rapport à un objectif à maximiser et agréger à chaque fois, sa récompense avec celle d'Ethicaa. Chaque agent Ethics.ai recevra finalement une récompense individuelle, pour son groupe et pour la société entière. Pour le déroulement étape par étape du schéma ci-dessus, le lecteur est invité à consulter les annexes 1 à 6.

3.2 Les agents sages

Nous avons choisi de permettre aux *sages* Ethicaa d'observer le comportement des *agents* par le biais des actions réalisées par ces *agents* à l'instant t et par les valeurs des propriétés globales et individuelles à l'instant $t-1$. Cela diffère donc de ce que nous avons mis en place pour le *sage* humain, mais permet d'avoir un jugement éthique sans avoir à disposition plus d'information que l'*agent* jugé et correspondant au modèle Ethicaa [13].

Lors de l'implémentation, nous avons élaboré deux *sages* Ethicaa -le *sage* de l'indépendance et le *sage* de la justice- selon les définitions suivantes :

Support d'indépendance : "Toute action permettant à la smart grid d'éviter les échanges avec le réseau électrique national promeut la valeur d'indépendance."

Support de justice : "Toute action permettant à la smart grid d'éviter d'exclure un agent et de garantir une équité de confort entre les agents promeut la justice."

D'après ces définitions, nous avons établi les supports de valeur nécessaires au raisonnement des agents Ethicaa. Ces supports sont détaillés dans les tableaux ci-dessous. Il y est décrit pour chaque action les conditions où celles-ci promeuvent (*promote*) ou enrayent (*defeat*) la valeur morale intéressée. Par exemple : Acheter plus de 100 d'énergie enraye la valeur morale Indépendance. Ici on effectue l'action *buy_energy(value)* avec $value > 100$. Ce qui correspond à la colonne *Defeat* de notre tableau. Notons que ce seuil de 100 dans le tableau de l'Indépendance correspond à un seuil de tolérance d'environ 5 à 10% des besoins minimaux en énergie électrique des agents présents dans la *smart grid*. En effet, Ethicai génère un bruit servant à favoriser l'exploration lors du choix des valeurs de chaque action par les *agents*. Nous ne pourrions donc pas réellement obtenir des actions avec $value$ à 0. Ce seuil de 100 permet d'être plus tolérant vis-à-vis de ce bruit d'exploration.

Actions(value) \ Indépendance	Defeat	Promote
buy_energy(value)	value > 100	value < 100
give_energy(value)	\emptyset	value > 100
grid_consumption(value)	\emptyset	value > 100
sell_energy(value)	value > 100	value < 100
storage_consumption(value)	\emptyset	\emptyset
store_energy(value) (from grid)	\emptyset	\emptyset

Figure 6: Tableau de support de valeur d'indépendance

Actions(value) \ Justice	Defeat	Promote
buy_energy(value)	Si $equity_{agent} < \varphi$ ET $well-being_{agent} > (1 + (1-\varphi)) \times well-being_{global}$	Si $exclusion_{agent} > 0.5$ OU (si $equity_{agent} < \varphi$ ET $well-being_{agent} < \varphi \times well-being_{global}$)
give_energy(value)	Si $exclusion_{agent} > 0.5$	Si $exclusion_{agent} < 0.5$ ET si $equity_{agent} < \varphi$
grid_consumption(value)	Si $equity_{agent} < \varphi$ ET $well-being_{agent} > (1 + (1-\varphi)) \times well-being_{global}$	Si $exclusion_{agent} > 0.5$ OU (si $equity_{agent} < \varphi$ ET $well-being_{agent} < \varphi \times well-being_{global}$)
sell_energy(value)	Si $exclusion_{agent} > 0.5$ OU si $equity_{agent} < \varphi$	\emptyset
storage_consumption(value)	\emptyset	\emptyset
store_energy(value) (from grid)	\emptyset	\emptyset

Figure 5: Tableau de support de valeur de justice

Il est important de noter que ces choix sont discutables du fait de ne pouvoir être purement objectif. Il serait intéressant de discuter de ceux-ci avec un expert. Dans le premier tableau, nous avons fixé un seuil de *value* à 100 (environ 10% des besoins minimal d'un agent pouvant être trouvé dans la smart grid). Cela permet d'éliminer la prise en compte du bruit d'exploration. Dans le second tableau, le seuil de la propriété d'exclusion est de 0.5, moyenne de 0 et de 1, les seules valeurs que cette propriété peut prendre.

À l'aide de ces règles de support de valeur, les *sages* vont juger le comportement des agents. Il est aussi nécessaire d'utiliser des règles morales. Dans un souci de simplicité et afin de respecter le temps imparti de ce stage, nous avons choisi de créer deux règles morales simples :

rm1 : "Toute action supportant la valeur morale est jugée morale."

rm2 : "Toute action qui enraye la valeur morale est jugée immorale."

À partir de cela, chaque *sage* Ethicaa va émettre un jugement pour chacune des actions de chacun des *agents* Ethics.ai. Nous obtenons donc une liste contenant un certain nombre de "moral" et de "immoral". Avec cette liste il a ensuite fallu déterminer une récompense numérique à transmettre à l'*agent*. Comme mentionné dans l'état de l'art, nous n'avons pas réellement trouvé de travaux sur ce problème. Ces travaux concernent l'intégration de symbolisme dans des modèles numérique ; c'est-à-dire un passage du numérique au symbolisme. Or, nous cherchons ici à passer du symbolisme au numérique. La solution que nous avons choisi est de remplacer les "moral" par des 1 et les "immoral" par des 0 puis effectuer une moyenne des valeurs de la liste de jugement. Si la liste est vide alors c'est qu'aucune action effectuée n'est morale ou immorale ; on a donc une récompense de 0.5.

3.3 Liaison entre les agents sages et les agents Ethicaa

De cette façon nous établissons donc un lien entre Ethics.ai et Ethicaa ; entre apprentissage automatique [9] [10] et jugement symbolique [13].

Il est important de noter que pour réaliser leurs jugements, les *sage* Ethicaa, vont utiliser trois éléments du simulateur Ethics.ai : les actions réalisées par chaque *agent* (voir l'ensemble *Action* plus haut), les propriétés globales à l'instant $t-1$ (avant action des *agents*) et les propriétés individuelles à l'instant $t-1$ (voir l'ensemble *P* plus haut). Le calcul de ces propriétés est explicité dans le tableau ci-dessous. Les propriétés globales sont issues de [9].

Propriétés	Individuelle	Globale
Autonomy	$1 - \frac{Transaction_{Ag}}{q_{Ag} + p_{Ag} + s_{Ag} + u_{Ag} + Transaction_{Ag}}$	$1 - \frac{\sum_i Transaction_i}{\sum_i (q_i + p_i + s_i + u_i + Transaction_i)}$
Energy Loss	$\min\left(\frac{Besoin_{Ag} + Donne_{Ag} - Pris_{Ag}}{Pris_{Ag}}, 0\right)$	$\min\left(\frac{E + \sum_i Donne_i - \sum_i Pris_i}{\sum_i Pris_i}, 0\right)$
Equity	$\frac{Bien_être_global - Bien_être_global - Confort_{Ag} }{Bien_être_global}$	$1 - Hoover(Conforts)$
Exclusion	$\begin{cases} 1 & \text{si } bien_être_{Ag} < 1^{er} \text{ quartile du } Bien_être_global \\ 0 & \text{sinon} \end{cases}$	$\frac{ Conforts_i < 0.5 \times Bien-Etre }{ Agents }$
Over-consumption	$\min\left(\frac{Pris_{Ag} - Donne_{Ag} - Besoin_{Ag}}{Pris_{Ag}}, 0\right)$	$\min\left(\frac{\sum_i Pris_i - \sum_i Donne_i - E}{\sum_i Pris_i}, 0\right)$
Well-being	confort	médiane du confort des agents

Figure 7: Tableau des formules de calculs des propriétés Ethics.ai

Soit :

- S la société d'agent,
- G l'ensemble des groupes avec N_{groupe} la taille de G ,
- $Agent$ l'ensemble des agents Ethics.ai avec N_{Agent} la taille de $Agent$,
- $Sage$ l'ensemble des agents Ethicaa avec N_{sage} la taille de $Sage$,

- *Action* l'ensemble des actions, $Action = \{buy_energy, give_energy, grid_consumption, sell_energy, storage_consumtion, store_energy\}$,
- *P* l'ensemble des propriétés Ethics.ai servant à observer une simulation avec N_{props} la taille de *P*. $P = \{Autonomy, Energy_loss, Equity, Exclusion, Over-consumption, Well-Being\}$ (cf Figure 7),

$$\langle x_{sage_1}, x_{sage_2}, \dots, x_{sage_{N_{Sage}}}, x_{sys_1}, x_{sys_2}, \dots, x_{sys_{N_{props}}} \rangle$$

le vecteur de poids correspondant à l'importance que la société accorde à chaque sage ethicaa et à chaque évaluation de propriétés faite par Ethics.ai, où x est un poids. x_{sage_i} étant le poid donnée à la récompense attribuée par le sage i . x_{sys_j} étant le poid donnée à la récompense attribuée par le système de récompense Ethics.ai pour la propriété j .

$$\langle x_{sage_1}, x_{sage_2}, \dots, x_{sage_{N_{Sage}}}, \langle x_{sys_1}^{g_1}, x_{sys_2}^{g_1}, \dots, x_{sys_{N_{props}}}^{g_1} \rangle, \dots, \langle x_{sys_1}^{g_{N_G}}, x_{sys_2}^{g_{N_G}}, \dots, x_{sys_{N_{props}}}^{g_{N_G}} \rangle \rangle$$

le vecteur de poids correspondant à l'importance que chaque groupe accorde à son sage ethicaa et à chaque évaluation de propriétés faite par Ethics.ai. On trouve les poids x comme dans le vecteur précédent. Cependant chaque groupe correspondant à un sage mais pouvant tous accorder différentes importances aux jugement Ethics.ai, on retrouve cette fois un vecteur de poids par groupe ; ces poids représentant l'importance accordée par chaque groupe pour chacune des propriétés pour les récompenses Ethics.ai

$$\langle \langle x_{sage_1}^{agent_1}, x_{sage_2}^{agent_1}, \dots, x_{sage_{N_{props}}}^{agent_1}, x_{sys_1}^{agent_1}, x_{sys_2}^{agent_1}, \dots, x_{sys_{N_{props}}}^{agent_1} \rangle, \dots, \langle x_{sage_1}^{agent_{N_{Agent}}}, x_{sage_2}^{agent_{N_{Agent}}}, \dots, x_{sage_{N_{props}}}^{agent_{N_{Agent}}}, x_{sys_1}^{agent_{N_{Agent}}}, x_{sys_2}^{agent_{N_{Agent}}}, \dots, x_{sys_{N_{props}}}^{agent_{N_{Agent}}} \rangle \rangle$$

le vecteur de poids correspondant à l'importance que chaque individu (agent Ethics.ai) accorde à chaque sage ethicaa et à chaque évaluation de propriétés faite par Ethics.ai (ces valeurs sont fixées par le concepteur et non par les agents eux-même). On retrouve ici un vecteur comprenant lui-même des vecteurs du même type que la récompense globale, mais pour chacun des agents (ceux-ci pouvant accordés différentes importantes aux récompenses des *sages* et d'Ethics.ai).

Avec : $g_i \in G, agent_j \in Agent, sage_k \in Sage, a \in Action, Action = \{buy_energy, give_energy, grid_consumption, sell_energy, storage_consumtion, store_energy\}$, $p \in P, P = \{Autonomy, Energy_loss, Equity, Exclusion, Over-consumption, Well-Being\}$, $reward_{sage}$ la récompense attribuée par le sage, $reward_{sys}$ la récompense attribuée par le système Ethicaa.

On a :

$$\forall agent \in Agent, \forall sage \in Sage, \forall p \in P,$$

$$reward_{sage}^{agent} = f(agent, \langle a_1^{agent}, \dots, a_6^{agent} \rangle, \langle p_1^{agent}, \dots, p_6^{agent} \rangle, \langle p_1^{globale}, \dots, p_6^{globale} \rangle)$$

La fonction calculant la récompense attribuée par un *sage* à un agent en fonction de l'agent, de ses actions, de ses propriétés et des propriétés globales.

$$\forall agent \in Agent, \forall p \in P, reward_{sys_p}^{agent} = f(agent, \langle a_1^{agent}, \dots, a_6^{agent} \rangle, p^{agent}, p^{globale})$$

La fonction calculant la récompense attribuée par le système de récompense interne à Ethics.ai, pour un type de propriété donnée, à un agent en fonction de l'agent, de ses actions, et de cette propriété individuelle et globales.

$$reward_individuelle^{agent} = \left(\sum_{sage}^{N_{sage}} reward_{sage}^{agent} \times x_{sage}^{agent} \right) + \left(\sum_{p=0}^{N_{props}} reward_{sys_p}^{agent} \times x_{sys_p}^{agent} \right)$$

$$reward_groupe_g^{sage} = reward_{sage} \times x_{sage} + \left(\sum_{p=0}^{N_{props}} reward_{sys_p}^g \times x_{sys_p}^g \right)$$

$$reward_globale = \left(\sum_{sage}^{N_{sage}} reward_{sage} \times x_{sage} \right) + \left(\sum_{p=0}^{N_{props}} reward_{sys_p} \times x_{sys_p} \right)$$

4 Implémentation du modèle

Dans cette section nous parlerons de la manière dont est implémenté actuellement le modèle ainsi que des choix qui ont été fait.

Comme vu précédemment, par ce modèle, nous lions les systèmes Ethics.ai et Ethicaa. Cependant dans un premier temps, nous avons réalisé un système de *sage* humain, sans Ethicaa, afin de mieux comprendre comment nous pourrions appréhender la situation à un instant t d'une simulation Ethics.ai. Pour cela nous avons remplacé le système de récompense Ethics.ai par un dashboard sur lequel nous pouvons voir en temps réel l'évolution des différentes propriétés d'observation -que ce soit globale ou par agent- ainsi que paramétrer la récompense globale et les récompenses par agent et par propriétés.

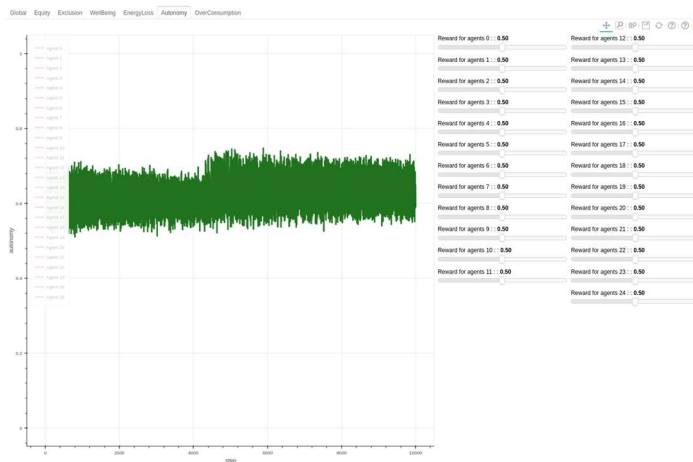


Figure 8: Vue du dashboard pour la propriété globale d'autonomie

Les informations affichées à l'écran sont les résultats des propriétés après chaque action, et donc à $t+1$. Cela nous a permis d'avoir une première version fonctionnelle d'Ethics.ai avec des jugements externes et utilisant les propriétés individuelles.

A partir de cela nous avons pu commencer le travail consistant à lier les systèmes Ethics.ai et Ethicaa. Le premier est réalisé en Python, le second utilise le framework Jacamo. Le lien technique est réalisé par le biais d'un artefact Jacamo communicant avec un serveur REST Python lancé par Ethics.ai et basé sur les travaux de [8].

L'implémentation actuelle comporte quelques différences avec le modèle. Premièrement, nous n'avons pas encore implémenté la notion de groupe dans Ethics.ai, nous n'avons donc pas de récompense de groupe. Ensuite, nous ne prenons pas en compte les récompenses données par Ethics.ai, mais seulement

celles données par Ethicaa. Par ailleurs, la récompense globale est actuellement numérique et sert seulement à calculer le score de fin de simulation ; elle n'est pas issue du jugement des *sages* et n'est pas prise en compte par les *agents*. Enfin, l'image du comportement de chaque *agent* que se construisent les agents sages n'est que ponctuelle sans agrégations sur une fenêtre temporelle.

5 Expérimentations

5.1 Contexte d'expérimentation

Afin de pouvoir comparer au mieux notre modèle avec le modèle initial d'Ethics.ai, nous avons réalisé 4 types de simulations :

- Une simulation sur le modèle initial développé par R. Chaput [9], en ajoutant les propriétés individuelles. Celles-ci ne changent pas l'apprentissage des agents, mais permettent de mieux observer le comportement individuel de ces agents vis-à-vis des différentes propriétés définies. Cette simulation utilise la fonction de récompense "adaptabilité2" qui est obtenant les meilleurs résultats à ce jour parmi les fonctions disponibles dans Ethics.ai
- Une simulation utilisant uniquement le *sage* de l'indépendance
- Une simulation utilisant uniquement le *sage* de la justice
- Une simulation utilisant le *sage* de l'indépendance et le *sage* de la justice

Toutes ces simulations sont configurées afin d'utiliser la même *seed* de génération aléatoire de l'environnement de simulation Ethics.ai : 3934981507. Ainsi chaque évaluation se fait à partir de la même situation initiale. Cette *seed* est une valeur aléatoire respectant les contraintes techniques imposées par le langage Python3 (valeur comprise entre 0 et $2^{32}-1$).

Nous avons aussi cherché à comparer différents paramètres de simulation pour l'agent justice. Nous avons d'abord fixé un seuil de 0.8 pour l'évaluation de l'équité dans certains cas. Cela veut dire que nous souhaitons atteindre au minimum un score d'équité de 0.8 et que nous avons une tolérance de 0.2. Cette valeur étant arbitraire, nous avons cherché à évaluer différentes possibilités : 0.8, 0.9, 0.95.

De la même manière, nous avons expérimenté ces paramètres en ajoutant un seuil de tolérance sur la valeur de l'action *value*, à l'instar du *sage* d'indépendance, pour les situations où des actions enrayent la valeur de justice. Ce seuil a été fixé à 100 (environ 10% des besoins minimaux d'un agent pouvant être trouvé dans la *smart grid*). Cela permet d'éliminer la prise en compte du bruit d'exploration.

5.2 Résultats

Nous présenterons ici les résultats obtenu grâce à l'implémentation actuelle du modèle, que nous comparerons aux résultats obtenus pour le modèle initial.

5.2.1 Visualisation et jugement humain

Comme mentionné précédemment, nous avons implémenté un dashboard permettant l'observation du comportement des agents par l'intermédiaire des propriétés Ethics.ai. Ce dashboard permet aussi l'attribution de récompenses par le biais de *sliders*.

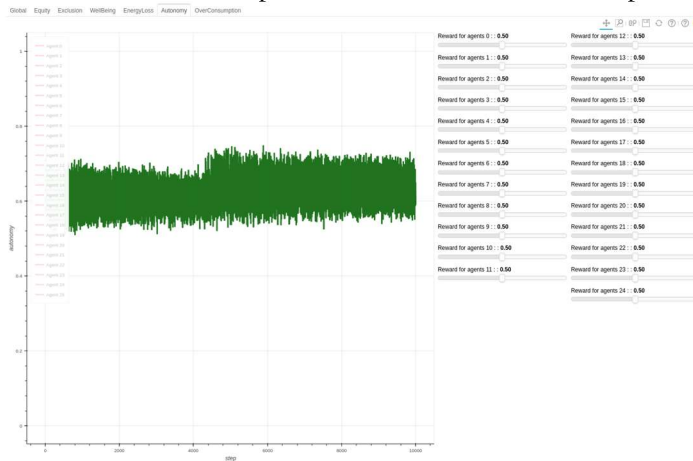


Figure 9: Vue du dashboard pour la propriété globale d'autonomie

Les informations affichées à l'écran sont les résultats des propriétés après chaque action, et donc à $t+1$. Cela nous a donné une première version fonctionnelle d'Ethics.ai avec des jugements externes et utilisant les propriétés individuelles.

5.2.2 Jugement des comportements appris vis-à-vis de l'indépendance

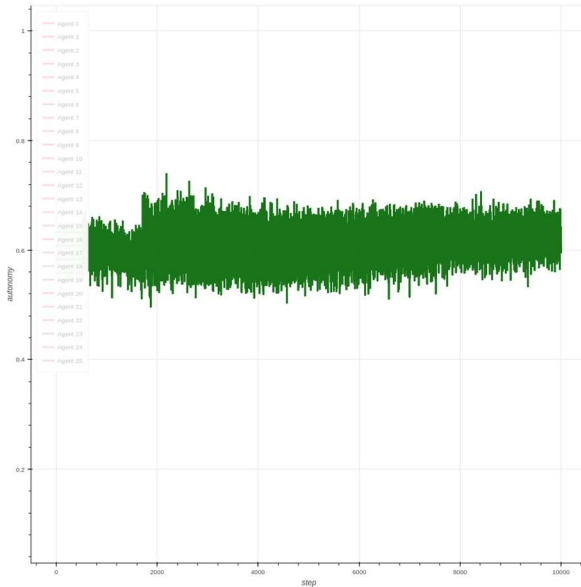


Figure 10: Courbe de la propriété *autonomy* pour le modèle initial

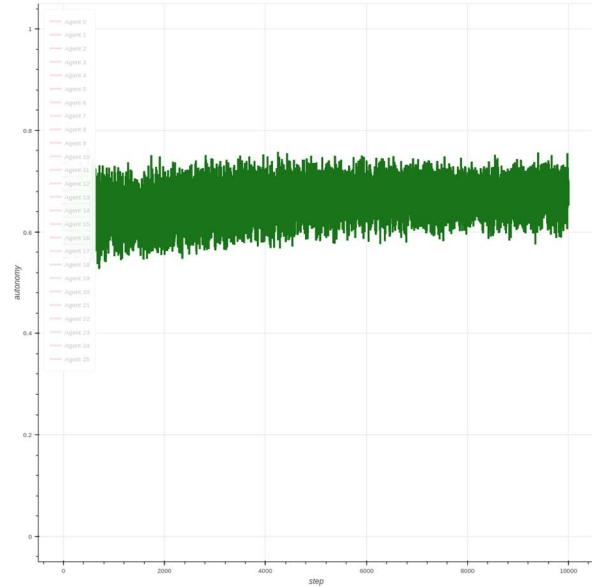


Figure 11: Courbe de la propriété *autonomy* pour notre modèle avec le sage justice seul

Sur ces graphes, on s'intéresse à la propriété *autonomy* sur laquelle se base le *sage* de l'indépendance (les autres graphes sont disponibles en annexes). On constate ci-dessus une meilleure autonomie globale du système Ethics.ai avec notre modèle (à droite) que sans (à gauche).

Cependant, les scores de simulations (évalués à l'aide de la récompense globale et compris entre 0 et 1) montrent que notre modèle avec l'agent indépendance seul, reste bien inférieur au score du modèle initial ; avec 0.6615565984200789 pour notre modèle et 0.8384945693018292 pour le modèle initial. Ceci est notamment dû au fait que ce dernier cherche à maximiser toutes les propriétés alors que notre modèle cherche ici, par le biais du *sage* d'indépendance, à maximiser seulement l'autonomie.

5.2.3 Jugement des comportements appris vis-à-vis de la justice

Trois expérimentations ont eu lieu sur le *sage* de la justice. Premièrement, comme on a pu le voir dans les tableaux de support des valeurs, la propriété d'équité est comparée à un seuil qui définit le minimum acceptable qu'elle doit valoir. Ce

seuil a été sujet à des expérimentations. Nous avons testé le modèle avec trois valeurs de ce seuil : 0.8, 0.9, 0.95. Pour les évaluations, nous nous basons sur les graphiques des trois propriétés utilisées par le *sage* de la justice : *equity*, *exclusion*, *well-being* (l'entièreté des graphes est disponible en annexes).

Hormis pour *well-being* où on constate que le seuil de 0.8 est meilleur, c'est le seuil de 0.9 qui présente les meilleurs graphiques, tel que nous pouvons le voir pour *equity* ci-dessous.

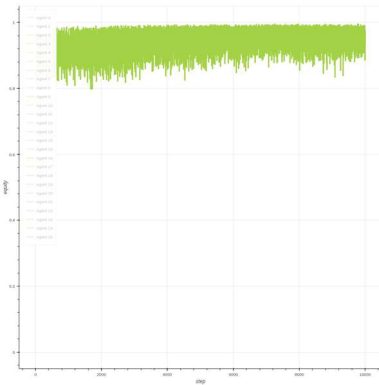


Figure 13: Courbe de la propriété equity pour la sage justice et un seuil de 0.8

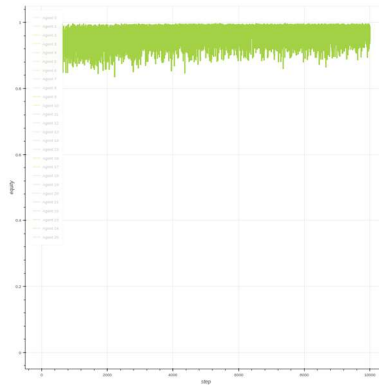


Figure 12: Courbe de la propriété equity pour la sage justice et un seuil de 0.9

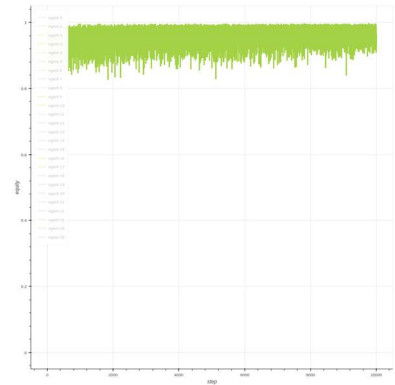


Figure 14: Courbe de la propriété equity pour la sage justice et un seuil de 0.95

Les scores de simulation, 0.910879580389887 pour le seuil de 0.8, 0.9657002326875181 pour le seuil de 0.9 et 0.9468049517060626 pour le seuil de 0.95, viennent donc appuyer le seuil de 0.9, suivi du seuil de 0.95 puis de celui de 0.8. Les résultats, que ce soit pour les graphiques ou pour le score, montrent donc le seuil de 0.9 comme optimal parmi ceux testés.

Nous avons ensuite effectué d'autres expérimentations sur le *sage* justice. À l'instar du *sage* de l'indépendance, nous avons exécuté une simulation en intégrant un seuil de tolérance de 100 (environ 5 à 10% du minimum de la valeur pour tous les agents) sur la valeur que peut prendre l'action afin de voir si le fait d'éliminer le bruit d'exploration pourrait être bénéfique. Nous avons aussi essayé de simplement remplacer la récompense de 0 par 0.0001 dans le cas où une action est jugée immorale (nous rappelons que la récompense finale est la moyenne de la récompense correspondant aux jugements chacune des actions).

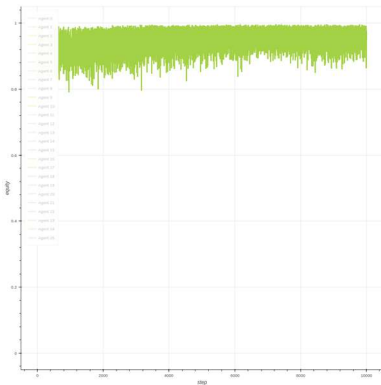


Figure 17: Courbe de la propriété equity pour la sage justice et un seuil de 0.9 plus le seuil de tolérance de 100

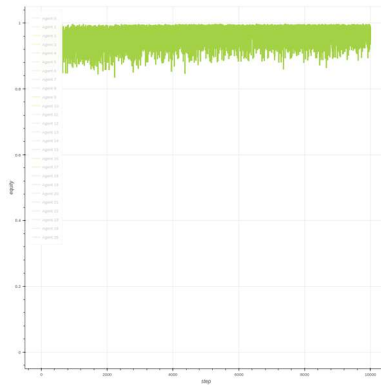


Figure 16: Courbe de la propriété equity pour la sage justice et un seuil de 0.9

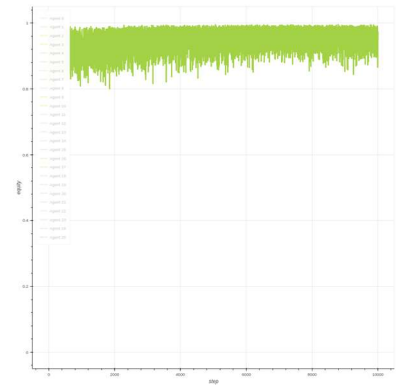


Figure 15: Courbe de la propriété equity pour la sage justice et un seuil de 0.9 plus la récompense de 0.0001 au lieu de 0 pour "immoral"

Ici, les graphiques, tels que ceux concernant *equity* ci-dessus montrent que le *sage* avec un seuil de 0.9 sans autre spécificité obtient les meilleurs résultats.

Cependant, les scores montrent celui-ci comme étant le moins bon : 0.9657002326875181 pour le seuil de 0.9 avec un seuil de tolérance de 100, 0.9503811507112466 pour le seuil de 0.9, 0.9646208099495219 pour le seuil de 0.9 avec une récompense de 0.0001 au lieu de 0. Rappelons cependant que le score se base sur la récompense globale qui est ici une moyenne numérique de la valeur d'équité et de la valeur d'exclusion de chaque agent. Il peut donc il y avoir des divergences entre la valeur numérique et la réalité de la simulation.

Enfin, la dernière expérimentation a été de comparer notre modèle avec le *sage* justice avec un seuil de 0.9, au modèle initial.

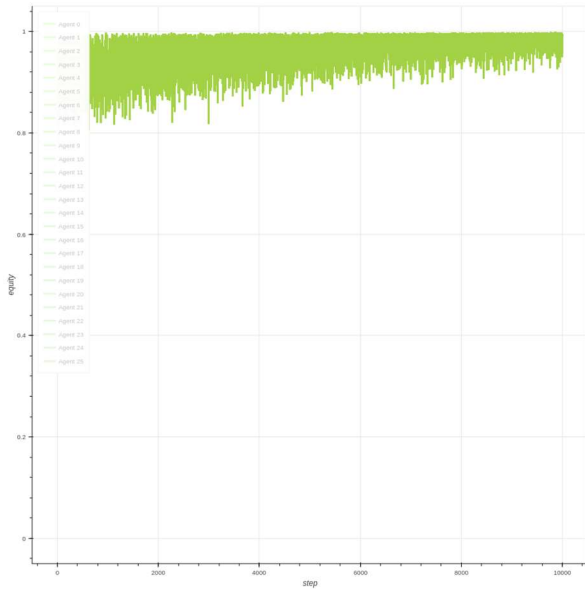


Figure 18: Courbe de la propriété equity pour le modèle initial

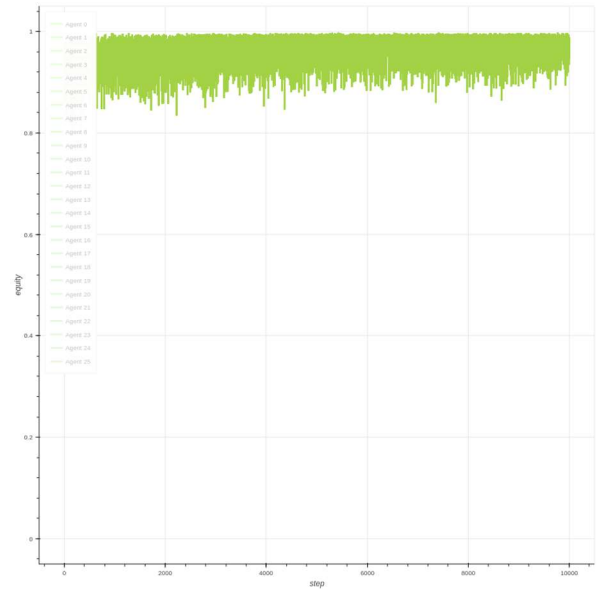


Figure 19: Courbe de la propriété equity pour l'agent justice seul et un seuil de 0.9

Comme on peut le voir sur les graphes ci-dessus (*equity*), notre modèle est meilleur pendant les 4000 premières étapes de simulation et s'améliore doucement, contrairement au modèle initial qui progresse rapidement et dépasse notre modèle après 6000 étapes. Cette évolution se retrouve aussi pour les propriétés *exclusion* et *well-being*.

Concernant les scores de simulations, le modèle initial obtient 0.8384945693018292, inférieur à notre modèle qui obtient 0.9468049517060626.

5.2.4 Jugement des comportements appris vis-à-vis de l'indépendance et de la justice réunis

Nous avons ensuite réalisé des expérimentations de notre modèle avec deux *sages* en même temps : l'indépendance et la justice (seuil de 0.9).

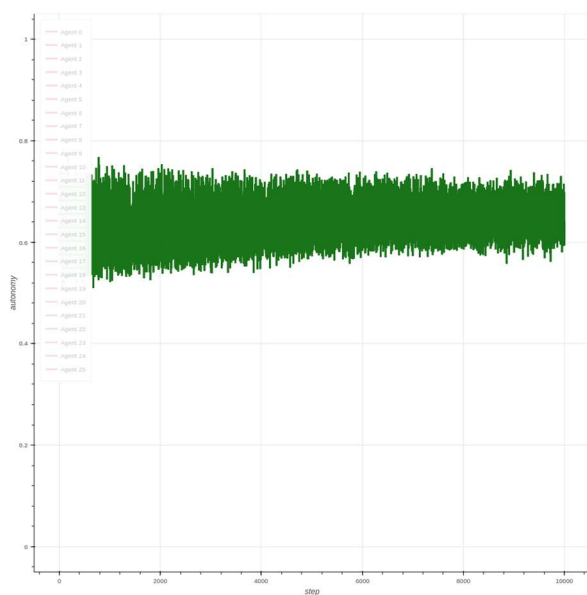


Figure 20: Courbe de la propriété *autonomy* pour notre modèle avec l'agents indépendance seul

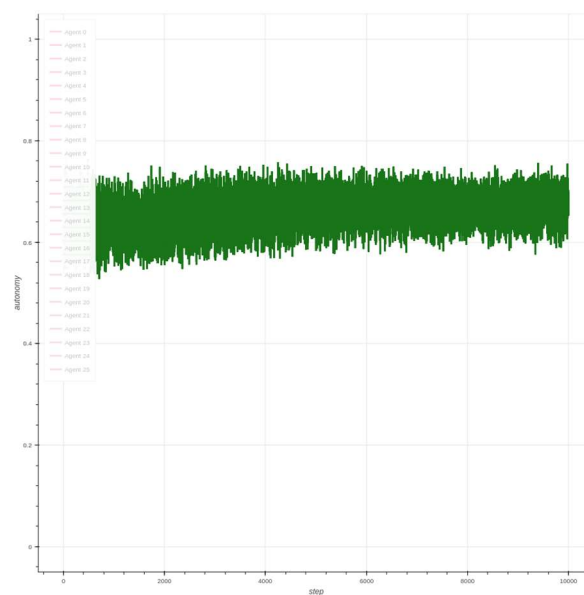


Figure 21: Courbe de la propriété *autonomy* pour notre modèle avec les agents indépendance et justice

Comme on peut le voir ci-dessus, le score de la simulation avec le *sage* indépendance seul reste meilleur que la simulation avec les deux *sages*.

Concernant la comparaison avec la simulation de l'agent justice seul, ce dernier présente de meilleurs résultats pour *equity* mais est moins bon que la simulation avec les deux agents pour *exclusion* et *well-being* (graphes en annexes).

Enfin, nous avons comparés les graphiques issus du modèle initial et de notre modèle. Il s'avère que pour *autonomy*, *equity* et *exclusion*, nous obtenons de meilleurs résultats avec notre modèle comme on peut le voir ci-dessous.

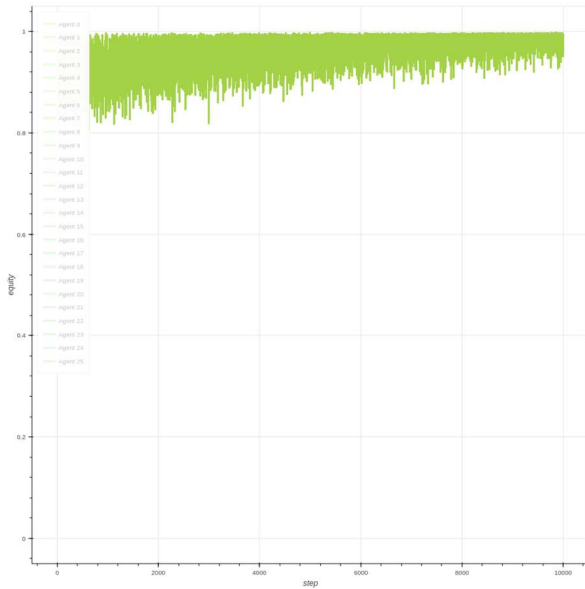


Figure 21 : Courbe de la propriété equity pour le modèle initial

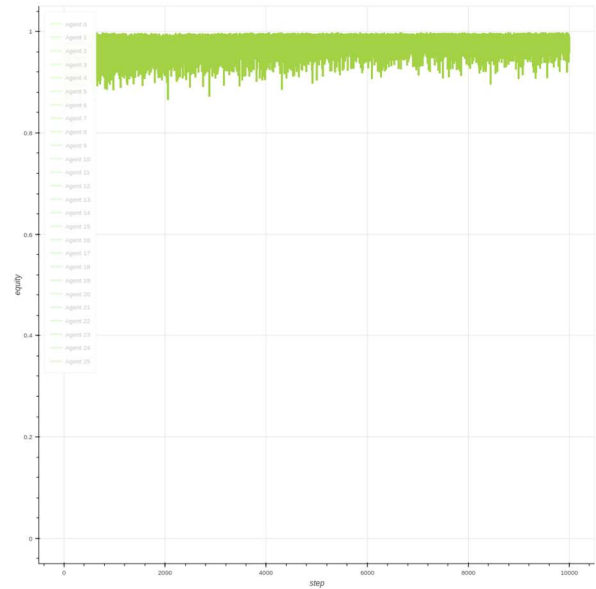


Figure 22: Courbe de la propriété equity pour notre modèle avec les agents indépendance et justice

Pour *well-being*, notre modèle est meilleur au début mais est rattrapé à la fin. Concernant les propriétés non ciblées par nos *sages*, nous sommes à égalité pour *energy-loss* et meilleur dans les dernières étapes pour *over-consumption*.

Néanmoins, en ce qui concerne les scores, notre modèle atteint 0.8035280420481312 alors que le modèle initial arrive à 0.8384945693018292.

6 Discussions

6.1 Discussion des résultats

Comme nous venons de le voir, notre modèle permet l'intégration de jugement symbolique dans le cycle d'apprentissage numérique qu'offre Ethics.ai. Nous avons créé deux *sages* autour de l'indépendance et de la justice. Ceux-ci permettent de cadrer et d'orienter le comportement des agents Ethics.ai durant l'apprentissage.

Nous avons ensuite expérimenté ce modèle de plusieurs manières. Les résultats présentés ci-avant nous permettent tout d'abord de montrer selon les graphiques que le seuil d'équité fixé à 0.9 pour l'agent justice est meilleur que les autres. Ensuite, nous pouvons observer que le *sage* d'indépendance, fixé sur la seule propriété *autonomy* obtient le meilleur graphique pour celle-ci parmi toutes les simulations effectuées. Cependant, pour le *sage* de la justice, qui se concentre sur *equity*, *exclusion* et *well-being*, on obtient de meilleurs résultats pendant la première moitié de la simulation, puis le modèle initial devient meilleur. Enfin, la simulation utilisant les deux *sages* ensemble permet d'obtenir globalement de meilleurs résultats que le modèle initial si l'on se base sur les graphes. Cependant ceci est contredit par les scores de simulation.

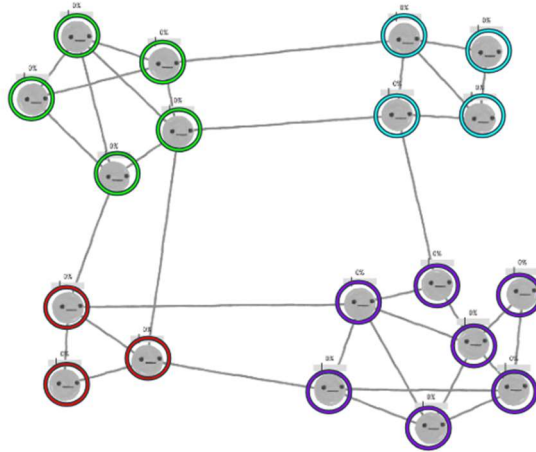
Ces contradictions entre les graphiques et les scores peuvent potentiellement s'expliquer par le fait que ces scores s'appuient sur le système de récompenses globales qui n'est pas vraiment implémenté dans notre modèle. Nous n'effectuons qu'une moyenne des propriétés que nous étudions.

6.2 Limites et perspectives

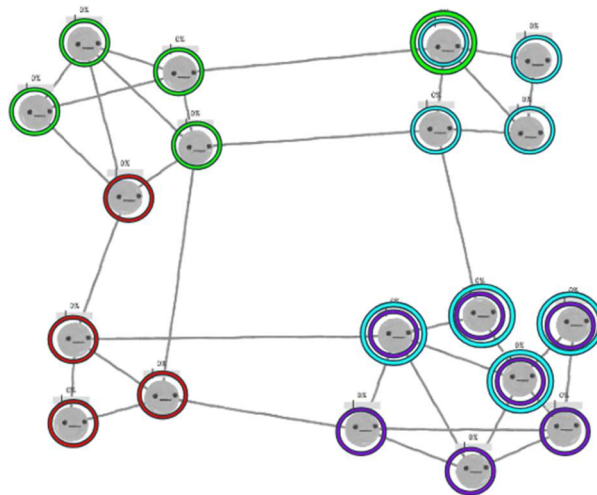
Comme mentionné dans la partie 4, l'implémentation actuelle comporte quelques différences avec le modèle. Pour rappel, nous n'avons pas encore implémenté la notion de groupe dans Ethics.ai, nous n'avons donc pas de récompense de groupe. Ensuite, nous ne prenons pas en compte les récompenses données par Ethics.ai, mais seulement celles données par Ethicaa. Par ailleurs, la récompense globale est actuellement numérique et sert seulement à calculer le score de fin de simulation. Enfin, les *sages* sont censés se construire une image du comportement de chaque *agent* mais actuellement nous n'effectuons qu'un jugement ponctuel, sans construction d'image dans le temps. Il serait donc intéressant de continuer l'implémentation jusqu'à correspondre réellement au modèle.

Il serait aussi intéressant de pousser les explorations sur la valeur des seuils de supports de valeurs et plus largement sur la définition des *sages* Ethicaa.

Enfin, notre modèle soulève une dimension qu'il serait intéressant de développer, celle de la diversité des idées. Nous pourrions notamment l'ouvrir sur la science des réseaux [11].



Nous pourrions constituer, comme ci-dessus, un certain nombre de groupes d'*agents* (en gris) interconnectés. Ces groupes seraient organisés autour d'un *sage* Ethicaa qui jugerait le comportement des *agents* de son groupe et enseignerait ainsi sa "vision" de l'éthique. Chaque lien représentant une connexion entre *agents*, ceux-ci pourraient discuter et argumenter. Cela permettrait une transmission et une cohabitation d'idées, comme sur le schéma ci-dessous.



Cependant, cela ne signifie pas une uniformisation des idées, des valeurs, des éthiques. Comme le souligne [11] il est important d'avoir une diversification des

idées sans avoir d'uniformisation ou qu'un groupe forme une bulle imperméable aux idées externes. C'est un équilibre des liens inter et intra groupes à trouver ; le principe de *bonding and bridging* [11].

Un tel modèle pourrait aussi inclure des notions telles que la confiance, l'amitié (accorder plus ou moins de poids aux liens avec les autres agents), les relations hiérarchiques, etc. Il serait ainsi envisageable de prendre en compte une grande diversité d'utilisateurs, d'utilisations, d'idées et d'éthiques.

7 Financements

Ce travail a été financé par la région Auvergne-Rhône Alpes, dans le cadre du projet Ethics.ai.

8 Remerciements

J'aimerais remercier Olivier Boissier, Rémy Chaput et Salima Hassas, mes encadrants, pour leur bienveillance, leur aide, les connaissances et le savoir-faire qu'ils m'ont transmis tout au long de ce stage.

J'aimerais remercier le laboratoire LIRIS et mes encadrants qui m'ont permis d'effectuer un stage épanouissant dans un cadre agréable.

J'aimerais remercier Amélie Cordier, Élise Lavoué, Salima Hassas et Jean-Charles Marty qui m'ont permis de m'intéresser au métier d'enseignants-chercheur et d'y trouver ma voie.

J'aimerais remercier mes collègues pour les moments de partage, autant sur le plan intellectuel que social.

J'aimerais remercier ma famille pour son soutien tout au long des mes études et de mon stage.

Enfin, j'aimerais particulièrement remercier Axelle, ma compagne, ainsi qu'Adrien et Béatrice, mes amis les plus proches, pour leur soutien sans faille dans toutes les épreuves de la vie.

9 Références

1. Allen, C. et al.: Artificial Morality: Top-down, Bottom-up, and Hybrid Approaches. *Ethics Inf. Technol.* 7, 3, 149–155 (2005).

- <https://doi.org/10.1007/s10676-006-0004-4>.
2. Allen, C. et al.: Why Machine Ethics ? IEEE Intell. Syst. 7 (2006).
 3. Balbo, F. et al.: Ethique et agents autonomes.
 4. Battaglino, C., Damiano, R.: Coping with Moral Emotions (Extended Abstract). 2.
 5. Berreby, F. et al.: A Declarative Modular Framework for Representing and Applying Ethical Principles. 10.
 6. Boissier, O. et al.: Dimensions in programming multi-agent systems. Knowl. Eng. Rev. 34, e2 (2019). <https://doi.org/10.1017/S026988891800005X>.
 7. Bonnemains, V.: Formal ethical reasoning and dilemma identification in a human-artificial agent system. 228.
 8. Bosello, M., Ricci, A.: From Programming Agents to Educating Agents – A Jason-based Framework for Integrating Learning in the Development of Cognitive Agents. 17.
 9. Chaput, R. et al.: Apprentissage adaptatif de comportements éthiques. 8.
 10. Chaput, R.: Apprentissage évolutif de comportements éthiques - Mémoire de Master.
 11. Christakis, N.A., Fowler, J.H.: Connected: The Surprising Power of Our Social Networks.
 12. Cointe, N. et al.: De l'intérêt de l'éthique collective pour les systèmes multi-agents. 6.
 13. Cointe, N.: Jugement éthique pour la décision et la coopération dans les systèmes multi-agents. École des Mines de Saint-Étienne (2017).
 14. Eppe, M. et al.: From semantics to execution: Integrating action planning with reinforcement learning for robotic causal problem-solving. ArXiv190509683 Cs. (2019).
 15. Lu, K. et al.: Robot Representation and Reasoning with Knowledge from Reinforcement Learning. ArXiv180911074 Cs. (2018).
 16. Moor, J.H.: The Nature, Importance, and Difficulty of Machine Ethics. IEEE Intell. Syst. 4 (2006).
 17. Nason, S., Laird, J.E.: Soar-RL: integrating reinforcement learning with Soar. Cogn. Syst. Res. 6, 1, 51–59 (2005). <https://doi.org/10.1016/j.cogsys.2004.09.006>.
 18. Noothigattu, R.: A Voting-Based System for Ethical Decision Making. 8.
 19. Rodriguez-Soto, M. et al.: A Structural Solution to Sequential Moral Dilemmas. N. Z. 9 (2020).
 20. Singh, M.P.: Norms as a basis for governing sociotechnical systems. ACM Trans. Intell. Syst. Technol. 5, 1, 1–23 (2013). <https://doi.org/10.1145/2542182.2542203>.
 21. Yu, H. et al.: Building Ethics into Artificial Intelligence. ArXiv181202953 Cs. (2018).

10 Annexes

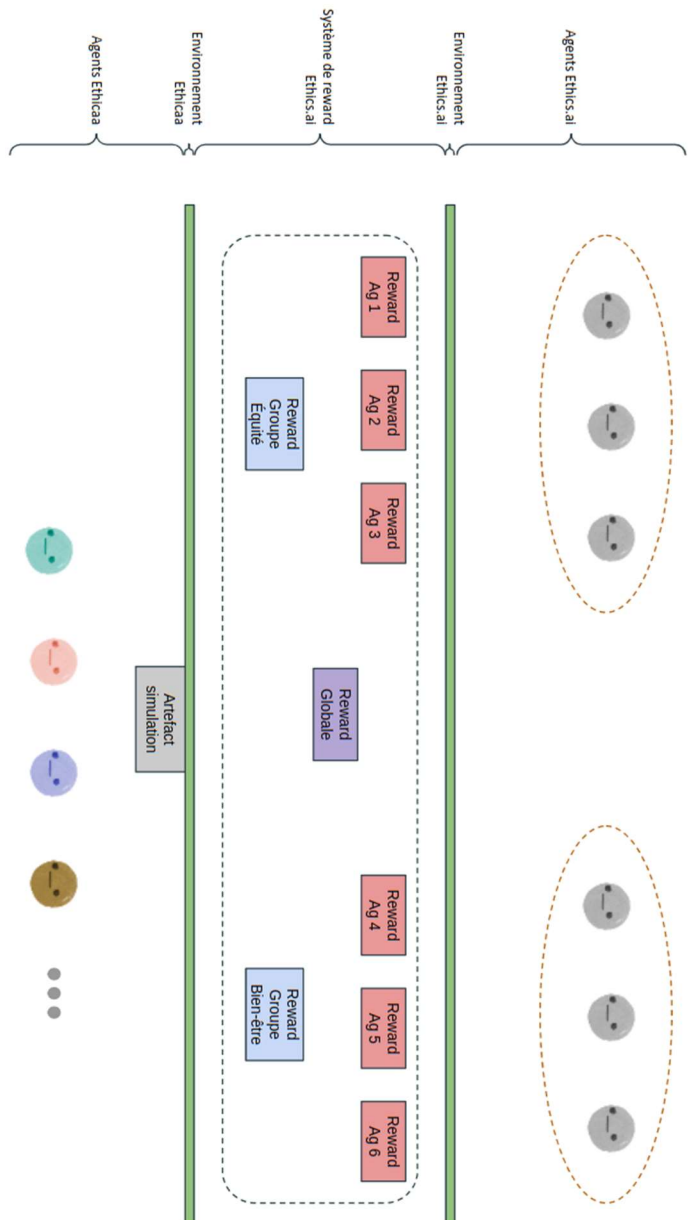


Figure 22: Schéma du modèle : vue globale

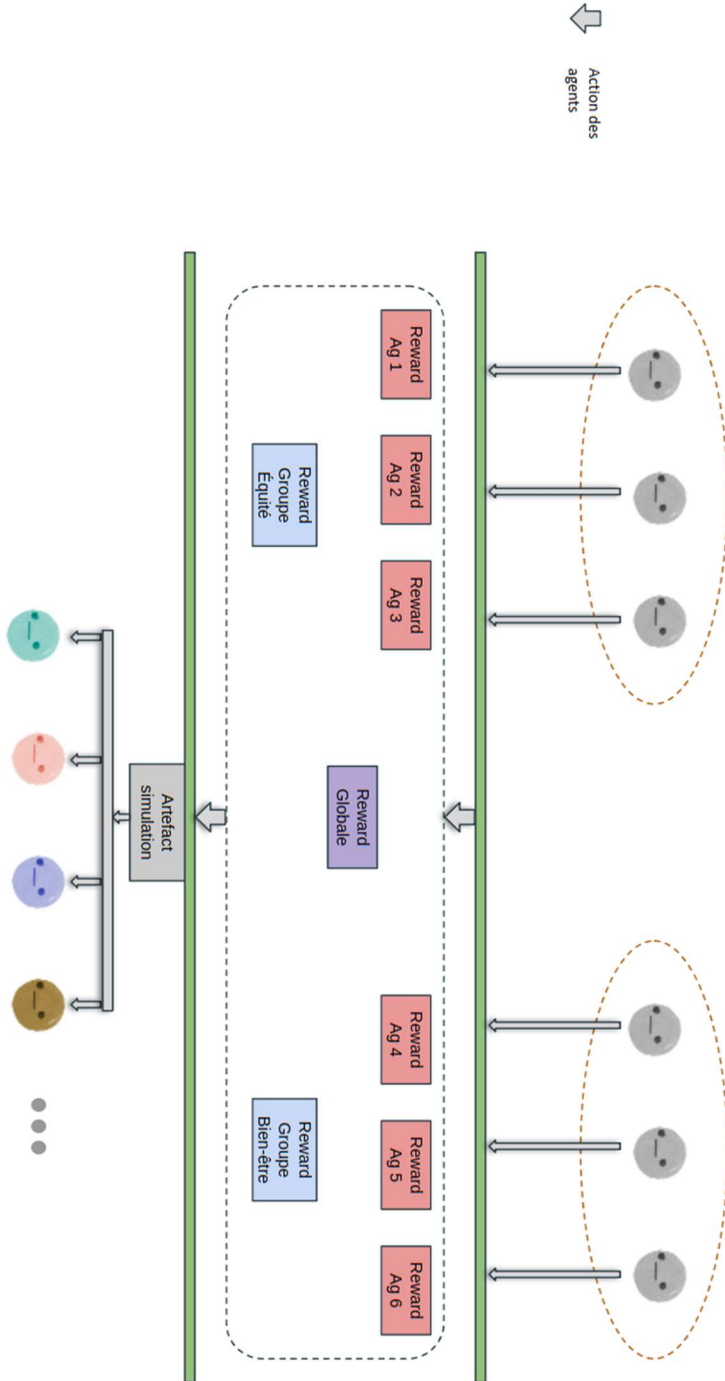


Figure 23: Schéma du modèle : étape 1 : les agents agissent

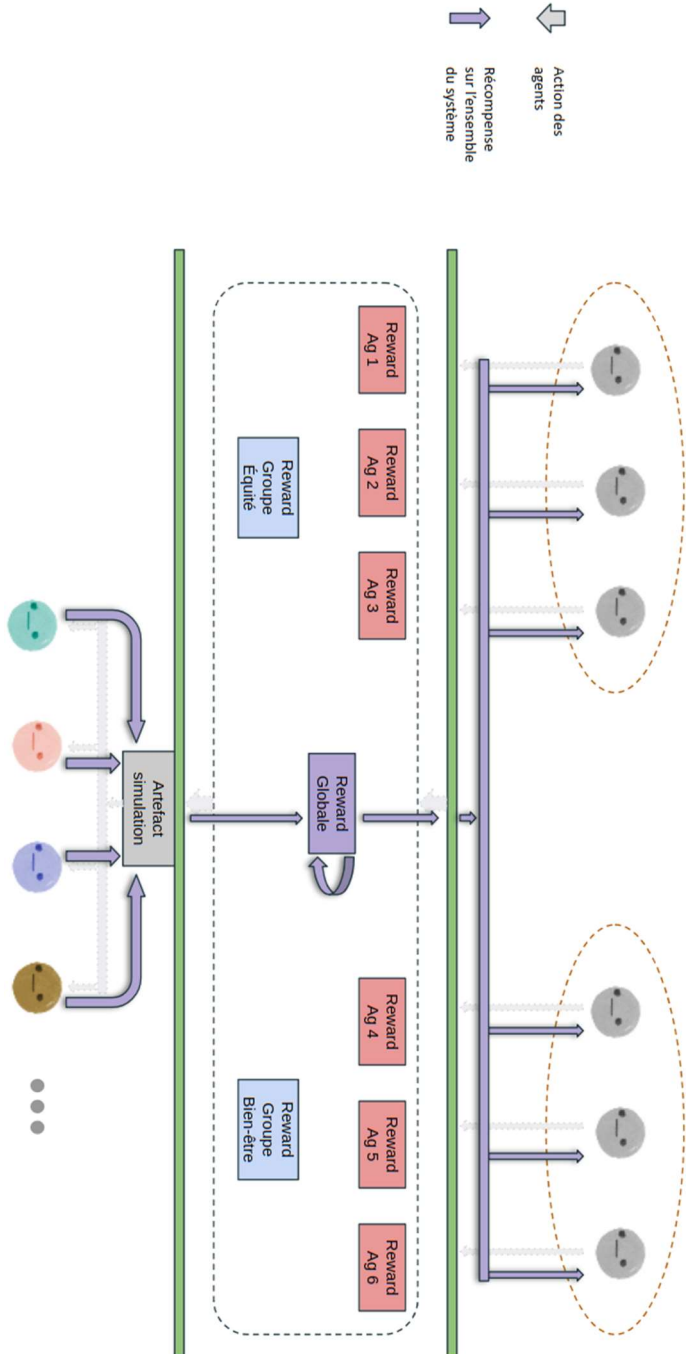


Figure 24: Schéma du modèle : étape 2 : Calcul de la récompense globale

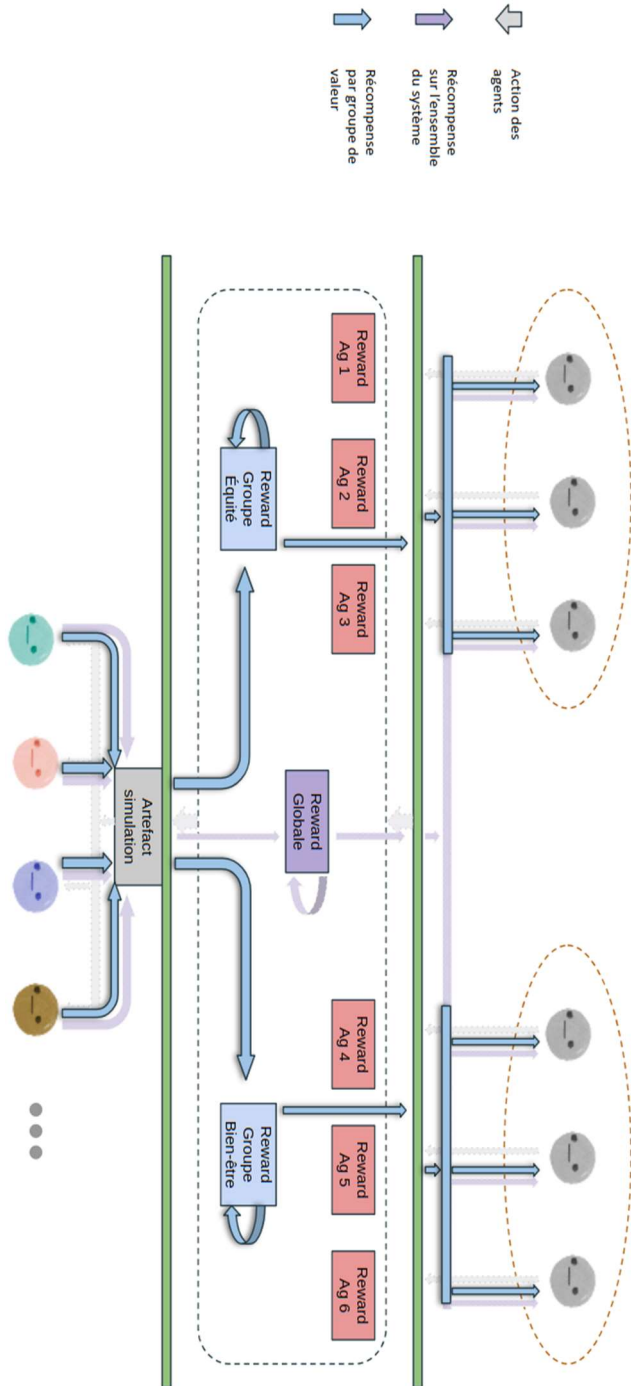


Figure 25: Schéma du modèle : étape 3 : Calcul de la récompense de groupe

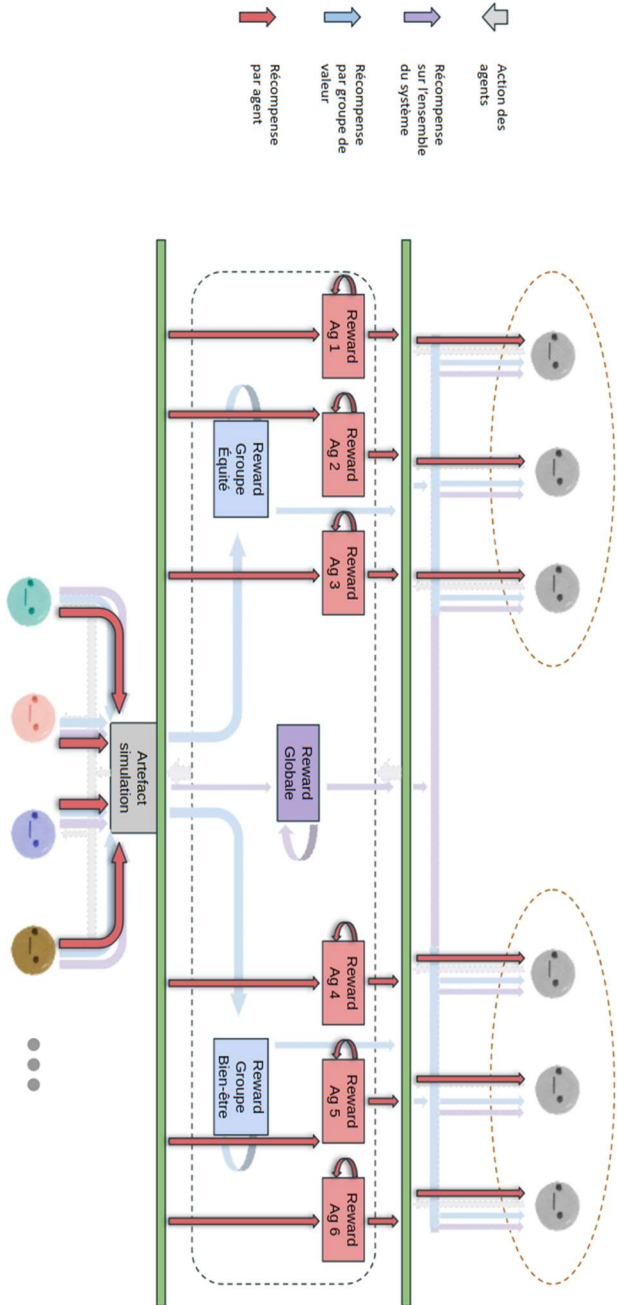


Figure 26: Schéma du modèle : étape 3 : Calcul de la récompense individuelle

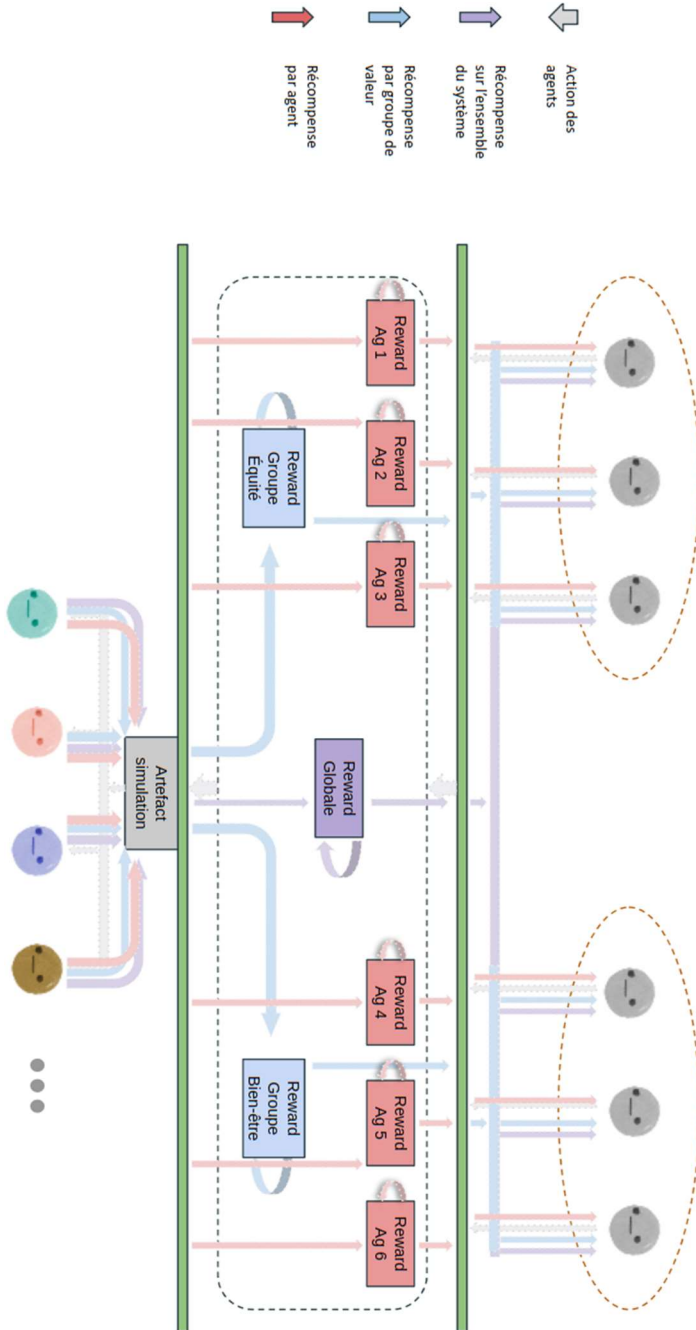


Figure 27: Schéma du modèle : résumé